Selective Cleavage of Variant Surface Glycoproteins from *Trypanosoma brucei*

By JAMES G. JOHNSON* and GEORGE A. M. CROSS†

MRC Biochemical Parasitology Unit, Moltano Institute, University of Cambridge, Downing Street, Cambridge CB2 3EE, U.K.

(Received 22 September 1978)

Two conformationally distinct regions were revealed by tryptic cleavage of six undenatured variant surface glycoproteins purified from clones of *Trypanosoma brucei*. Within 5 min, the native glycoproteins (65 000 mol.wt.) were cleaved, yielding a large *N*-terminal fragment (48 000–55 000 mol.wt. depending on the variant) together with one or more *C*-terminal fragments. After 30–60 min incubation, further breakdown of the large fragment occurred in some variants. The ultimate large product (40 000–52 000 mol.wt.) was very resistant to further degradation by trypsin (in the absence of denaturation). The distinction between *N*-terminal and *C*-terminal domains may be significant in relation to the organization and function of these glycoproteins on the trypanosome surface.

Three species, *Trypanosoma brucei*, *Trypanosoma vivax* and *Trypanosoma congolense*, are the major agents of human and animal trypanosomiasis in Africa. The pathogenicity of these organisms in susceptible hosts and their persistence in apparently healthy reservoir hosts is due to their ability to evade the host's immune responses. Antigenic variation is the primary mechanism for evasion (for reviews see Doyle, 1977; Cross, 1978*a,b*; Vickerman, 1978). Antigenic variation is mediated through sequential expression of an extensive series of variant surface glycoproteins, which confer serological individuality to successive isolates of trypanosomes taken from a single persistent infection. The plasma membrane of each serologically homogeneous population of *T. brucei* is covered by a single characteristic glycoprotein forming a closely packed layer (Cross, 1975; Cross & Johnson, 1976) visible in the electron microscope as a surface coat of uniform depth and density (Vickerman, 1969). The variant glycoproteins are firmly attached to the surface of healthy living trypanosomes, but appear to be largely and instantaneously released in soluble form after cell disruption by mechanical or osmotic means. Their solubility has facilitated the purification of individual

glycoproteins from many variants of several clones or isolates of *T. brucei* (Cross, 1975, 1977).

Variant surface glycoproteins from *T. brucei* consist of a single polypeptide having an apparent mol.wt. of 65 000 on SDS/polyacrylamide gels. The carbohydrate content varies from 7 to 17% (by wt.) in a range of glycoproteins studied (Johnson & Cross, 1977). Isoelectric points and amino acid compositions vary widely and *N*-terminal sequences so far studied show no obvious homology (Bridgen *et al.*, 1976). Although the variant glycoproteins are antigenically distinct, especially *in situ*, they have been shown to have cross-reacting determinants (Barbet & McGuire, 1978).

The repertoire of variant glycoproteins that may be expressed by a single trypanosome clone is so far undetermined, but probably exceeds 100. Their main apparent function is to disguise the parasite surface to avoid detection and destruction by antibodies generated against previously expressed variants. The next step is to explore the structure of variant glycoproteins in order to illuminate the structural and genetic basis of antigenic variation.

Undenatured proteins may be quite resistant to proteolysis, and the native conformation may selectively expose specific regions that are susceptible to enzymic cleavage. Exploitation of this concept was conspicuously successful in distinguishing functional domains of the immunoglobulins (Porter, 1973). Although the main requirement of variant glycoproteins seems to be to differ antigenically, we reasoned that there might be constraints on conformational variations for reasons of synthesis, processing or the need for a common region for membrane attachment. These considerations led us to try

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

* Present address: Malaria Section, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20014, U.S.A.

† To whom reprint requests should be addressed. Present address: Department of Immunochemistry, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, U.K.

the approach of selective cleavage to determine whether distinct structural and functional domains could be distinguished in variant glycoproteins. The results were not as striking as with immunoglobulins. However, a common feature of all variant glycoproteins was a cleavage pattern that suggested that the *N*-terminal two-thirds and the *C*-terminal one-third of the polypeptide folded to form two distinct and independent domains.

Methods

Trypanosomes and isolation of variant surface glycoprotein

Clones of *T. brucei* strain 427 were prepared from single trypanosomes as previously (Cross, 1975), with the use of X-irradiated mice when possible. Infected mouse blood was stored in liquid N₂, the preserved samples being sequentially numbered when frozen. As there is, so far, no rational basis for classification of variant antigens they are, in this paper, referred to simply by the sample reference numbers of the clone from which they were obtained. These numbers have no further significance. The variant clones studied in this paper were selected randomly. A variant (221) additional to those whose isolation was previously described (Cross, 1975) was obtained as the predominant variant present after several subcultures in the bloodstream-form trypanosome cultures initiated by Hirumi *et al.* (1977). Variant glycoproteins were purified as described previously (Cross, 1975, 1977), but without surface labelling, and were homogeneous with respect to molecular weight and isoelectric point.

Trypsin digestion and purification of products

Small-scale analytical digestions were performed by incubation of variant glycoprotein at 2mg/ml in 0.050M-Hepes, pH 7.4, at 37°C for various times with trypsin at concentrations ranging from 0.028 to 0.10mg/ml. The specific concentration of trypsin is indicated for each experiment. Reactions were terminated by heating for 3min at 100°C, and the products analysed by SDS/polyacrylamide-gel electrophoresis. Larger-scale digestions for purification of products were performed under similar conditions but at higher concentrations of variant glycoprotein (5mg/ml) with proportionate increases in trypsin concentration. Preparative digestions were terminated by addition of phenylmethanesulphonyl fluoride (1mM) or by addition of a 20% excess of soya-bean trypsin inhibitor or, more recently and most conveniently, by passing the mixture through a small column of soya-bean trypsin inhibitor coupled to CNBr-activated Sepharose (Pharmacia) at 20°C. Coupling of the inhibitor to Sepharose was achieved essentially according to the general instructions supplied by the manufacturer. The column was equi-

librated with 0.05M-Tris/0.1M-NaCl, pH 7.6. For re-use of the column, trypsin was eluted with 0.05M-sodium citrate, pH 3.0. Variant glycoprotein-cleavage products were purified by gel filtration on Sephadex, Sephacryl or Ultrogel, on columns having dimensions 16mm×900mm or 15mm×800mm. Columns were eluted with 0.050M-NH₄HCO₃ together with 0.5M-NaCl in the case of Sephacryl (Belew *et al.*, 1978). Absorbance of the effluent was monitored simultaneously at 206 and 278nm. Preparative (column) isoelectric focusing was performed as previously described (Cross, 1975). Polyacrylamide-gel isoelectric focusing was performed by using the apparatus and methods of LKB.

Other analytical methods

Products of trypsin digests were analysed either on 10% polyacrylamide gels as previously described (Cross, 1975) or by other published methods (Laemmli, 1970). Low-molecular-weight fragments purified from large-scale digests were also analysed on 15% polyacrylamide gels run in silicon-treated tubes. Amino acid and sugar analyses were performed as previously described (Cross, 1975; Johnson & Cross, 1977). *N*-Terminal amino acid-sequence analysis of fragments from variant glycoproteins 52 and 99 were performed by using automatic solid-phase techniques as previously described (Bridgen *et al.*, 1976). *N*-Terminal sequence analysis of fragments from variant glycoproteins 55 and 221 were obtained through the generous assistance of Dr. D. Stone and Mrs. S. Paterson (Wellcome Research Laboratories), by using a Beckman 890C sequencer and making identifications by two-dimensional t.l.c. of phenylthiohydantoin derivatives and amino acid analysis after back-hydrolysis.

Materials

Trypsin (treated with chymotrypsin inactivators) was purchased from Sigma or Worthington.

Results

In a brief series of initial experiments, native variant glycoproteins were treated with several concentrations of Pronase, trypsin, chymotrypsin, papain and pepsin, under conditions appropriate to the pH optimum for each enzyme. Only trypsin showed a clearly controlled effect that encouraged further study. Each variant glycoprotein was cleaved to produce a large fragment (41 000–52 000 mol.wt.), which was relatively or substantially resistant to further degradation, together with one or several smaller peptides which were rather susceptible to further fragmentation. The small fragments often stained poorly with Coomassie Blue and cannot be seen on some of the densitometer tracings illustrating this paper. Large and small tryptic fragments

were purified for further characterization from digests of about 15 mg of variant glycoprotein.

Fragmentation of variant glycoprotein 52

The results illustrated in Fig. 1 show the rapid initial cleavage of the 65 000-mol.wt. glycoprotein, yielding a 48 000-mol.wt. fragment (f48) which, by 30 min, was almost completely degraded to a 41 000-mol.wt. fragment (f41) which was resistant to further cleavage for at least 16 h.

Fig. 1. Fragmentation of variant glycoprotein 52. Densitometer scans at 265 nm of Coomassie Blue-stained polyacrylamide gels of variant glycoprotein 52 untreated (a) or treated with trypsin (one-seventieth by weight) for 7 min (b), 15 min (c) or 60 min (d). The vertical bar represents 0.2 A unit. Three vertical arrows indicate mol.wts. of 65 000, 48 000 and 41 000. Top of gel is at left.

Fig. 2 shows the initial separation, on Sephadex G-100, of the products of digestion of 14 mg of glycoprotein by 0.2 mg of trypsin for 60 min at 37°C. A second digestion produced a similar profile, except that the area of peak D was equal to that of peak C. By polyacrylamide-gel analysis it was shown that 'peak' A (a shoulder on the leading edge of peak B) contained pure fragment f48. Peak B consisted of a mixture of approximately equal amounts of fragments f48 and f41. Re-running peak B on Sephadex G-150 produced a similar shoulder and peak pattern to that of the original profile (Fig. 2). Fragment f48 could be obtained pure, but fragment f41 could not be purified from the mixture by gel filtration. There was no obvious reason why transformation of fragment f48 to f41 was slower in this experiment than in the analytical experiment shown in Fig. 1. On subsequent occasions there has been no difficulty in obtaining complete transformation into, and purification of, fragment f41. It seems likely that the transformation of fragment f48 into f41 might be very sensitive to small variations in incubation conditions, such as salt concentration, or to the purification and storage history of variant glycoprotein, which changed slightly during the course of these studies. Early preparations of variant glycoprotein yielding mainly fragment f48 probably contained some salts (particularly NH_4HCO_3) and were stored at -20°C after freeze-drying. Recent preparations yielding fragment f41 were stored at -196°C in essentially salt-free solution, which might have favoured unfolding of the

Fig. 2. Sephadex G-100 chromatography of tryptic cleavage products from variant glycoprotein 52.

Fractions of volume 2.2 ml were collected. Absorbance was monitored at 220 nm. Peaks A, B, C and D were pooled as indicated. Void volume was 57 ml, salt volume 163 ml.

molecule, thereby facilitating the further cleavage of fragment f48.

After removal of a small amount of lower-molecular-weight material on a Sephadex G-50 column (Fig. 3a), peak C yielded a single polypeptide. Peak D

Fig. 3. Sephadex G-50 chromatography of peaks C (a) and D (b) from Fig. 2

The vertical bars represent 0.2 A unit at 206 nm. Peaks C, G, H and J were individually rechromatographed (see the text). Void volume was 47 ml, salt volume 134 ml.

(Fig. 2) was also re-run on Sephadex G-50 (Fig. 3b), yielding three peaks (G, H and J), which were resolved completely by re-running each individually on the same Sephadex G-50 column. Peak J was present in a smaller amount than peaks G and H, which were present in equal amounts. Component C could be further degraded by trypsin to products indistinguishable from G and H. None of components C, G and H were degraded by treatment with CNBr, indicating the absence of methionine residues. By periodate/Schiff staining of gels and by colorimetric, gas-chromatographic and amino sugar (amino acid analyser) analysis, carbohydrate was shown to be present in components C, G and H but was absent from fragments J, f41 and f48. On polyacrylamide gels stained with Coomassie Blue, component C appeared distinctly pink compared with the blue colour of other bands. This pink colour was exhibited by glycopeptides of similar size from other variant glycoproteins.

When the proportions in which the various fragments were recovered was also considered, these results suggested that the initial cleavage of variant glycoprotein 52 yields fragment f48 together with component C, which is further degraded to components G and H. Mol.wt. values of 17000 (f17), 9000 (f9) and 8000 (f8) were assigned to fragments C, G and H respectively. From their respective K_D values of 0.15, 0.29, 0.33 and 0.44 on Sephadex G-50, the apparent mol.wts. of fragments C, G, H and J were calculated to be 20000, 12000, 10000 and 8000. On 15% polyacrylamide/SDS gels, the apparent mol.wts. of fragments C, H and J were 20200, 12900 and 9600. Coomassie Blue staining failed to reveal component G. This may be a reflection of its amino acid composition (see below). These molecular-weight values obtained by use of Sephadex and polyacrylamide would be expected to be anomalously high owing to conformational reasons and carbohydrate content (about 20%, w/w). Furthermore, if these experimentally determined molecular weights are used, the amount of glucosamine recovered from the fragments exceeds what is present in the original variant glycoprotein.

Amino acid analyses (Table 1) of the recovered fragments substantiated the deduced cleavage pattern. Substantial variations in the proportions of certain amino acids confirm the individuality of fragments f8 and f9, and the sum of their compositions equals the composition of fragment f17. Addition of the compositions of fragments f17 and f48 gives values agreeing well with an analysis of the intact variant glycoprotein. Component J was assigned a mol.wt. of 7000 (f7). It clearly cannot be derived from fragment f17, as it contains histidine and arginine, which fragments f17, f9 and f8 do not. Fragment f7 seems likely to result from the cleavage of fragment f48 to f41. This inference cannot be

Table 1. Amino acid and glucosamine analyses for variant glycoprotein 52 and tryptic fragments

Amino acid	Component ...	Composition (mol/mol of polypeptide)							Intact glycoprotein
		f7	f8	f9	Sum of f8+f9	f17	f48	Sum of f17+f48	
Asx		10.7	11.0	4.6	15.6	15.5	58.5	74.0	79.3
Thr		6.4	7.3	15.2	22.5	21.9	47.3	69.2	67.5
Ser		6.2	3.7	3.5	7.2	6.4	27.6	34.0	32.5
Glx		5.2	9.6	15.7	25.3	25.5	42.6	68.1	65.4
Pro		2.6	0	4.3	4.3	3.9	10.3	14.2	15.8
Gly		4.6	5.7	12.4	18.1	19.0	24.5	43.5	37.4
Ala		10.7	9.2	8.6	17.8	17.4	59.3	76.7	75.8
Cys		1.84	5.1	4.7	9.8	11.6	8.0	19.6	15.8
Val		0.82	1.7	4.1	5.8	5.1	13.4	18.5	18.7
Met		0	0	0	0	0	3.4	3.4	4.1
Ile		2.2	0	0	0	0	15.8	15.8	17.9
Leu		4.4	1.7	0.60	2.3	2.3	44.6	46.9	51.2
Tyr		0.82	1.6	0.34	1.9	1.3	16.1	17.4	17.0
Phe		1.00	1.8	0.23	2.0	1.8	9.0	10.8	11.8
His		1.16	0	0	0	0	6.0	6.0	6.4
Lys		7.4	15.9	13.6	29.5	30.7	42.5	73.2	69.9
Arg		1.5	0	0	0	0	15.0	15.0	16.7
GlcN		0	3.5	2.4	5.9	5.3	0	5.3	4.4
Column peak identification		J	H	G		C	A		
Molecular weight		7000	8000	9000		17000	48000		65000

verified by comparison of the amino acid compositions of fragments f48 and f41, as neither they nor f7 possess unique compositional features. Fragment f7 is unlikely to be a C-terminal extension of f17 because only about 0.1 mol of f7 was recovered/mol of f17.

N-Terminal-sequence analysis (residues 2–15) of fragment f48 corresponded to that of the intact variant glycoprotein (variant III in Bridgen *et al.*, 1976). Taken together, these results suggest that the fragments occur in the order (f41, f7)–(f8, f9) in the intact variant glycoprotein. In agreement with these results is the additional finding that CNBr cleavage or trypsin digestion of denatured citraconylated variant glycoprotein 52 both produce fragments containing all the glucosamine of the intact variant glycoprotein but lacking any histidine. These fragments are larger than f17 by about 4000 mol.wt.

Fragmentation of variant glycoprotein 55

When incubated with trypsin, variant glycoprotein 55 was cleaved rapidly, yielding a single large fragment (f52) of 52000 mol.wt. Cleavage was complete within 30 min and fragment f52 was stable for at least 16 h in the presence of trypsin. This is the largest stable fragment obtained from any variant glycoprotein. Preparative digests were performed with 5 mg of variant glycoprotein/ml and 0.1 mg of trypsin/ml for 40 min at 37°C. Fragment f52 was purified on Sephadex G-150 with or without prior removal of trypsin. N-Terminal-sequence analysis in the Beck-

man sequencer showed that residues 1–13 were identical with the previously published N-terminal sequence of the intact antigen (variant IV in Bridgen *et al.*, 1976).

Fragment f52 from variant glycoprotein 55 was examined by electrofocusing on polyacrylamide gels. Three closely spaced equimolar bands were seen with isoelectric points of 8.3, 8.0 and 7.8, suggesting sequence heterogeneity in the C-termini due to slight variation in the initial point of tryptic cleavage or subsequent fraying of the C-termini. Similar charge dispersity was observed in the large tryptic fragments of all antigens examined. The amino acid and glucosamine composition of fragment f52 is shown in Table 2. Most of the glucosamine present in the intact variant glycoprotein is missing from this fragment. Small fragments representing the C-terminal products of tryptic cleavage could not be recovered from variant glycoprotein 55, suggesting that they were rapidly degraded to small peptides. The C-terminus of variant glycoprotein 55 seems to be especially sensitive to proteolytic degradation.

Fragmentation of variant glycoprotein 99

Analytical-scale digests of variant glycoprotein 99 (Fig. 4) produced fragments with apparent mol.wts. of 48000 (f48), 40000 (f40) and 20000. This pattern was similar to that found for variant glycoprotein 52, but the small fragment [assigned a mol.wt. of 17000 (f17) rather than 20000] was clearly visible on gels and seemed more resistant to further degradation than

Table 2. *Amino acid and glucosamine analyses for several variant glycoproteins and tryptic fragments*

Variant glycoprotein ...		Composition (mol/mol)														49	
		55					99					60				Sum of f48+ f49	
Amino acid	Component	f52	Intact glycoprotein	f48	f17	Sum of f48+f17	Intact glycoprotein	f52	f13	Sum of f52+f13	Intact glycoprotein	f45	2×f10	2×f10	Intact glycoprotein		
Asx	...	44.1	60.0	49.4	27.7	77.1	75.8	52.0	15.8	67.8	61	36.8	24.8	61.6	55.0		
Thr		55.0	73.0	42.7	10.7	53.4	54.8	56.9	13.8	70.7	70	39.6	20.8	60.4	55.0		
Ser		36.0	41.5	21.6	2.3	23.9	24.7	20.8	6.6	27.4	26	23.8	13.8	37.6	36.1		
Glx		50.6	63.0	40.3	29.7	70.0	70.4	65.0	19.4	84.4	80	57.3	31.2	88.5	84.1		
Pro		27.0	26.0	7.8	8.0	15.8	15.4	14.3	3.4	17.7	19	10.0	2.7	12.7	16.6		
Gly		25.0	32.4	41.8	20.3	62.1	62.6	35.4	9.8	45.2	44	40.7	15.0	55.7	52.7		
Ala		85.0	101.0	89.9	12.0	101.9	100.0	61.5	10.6	72.1	75	64.5	14.0	78.5	78.1		
Cys		10.0	14.0	6.0	9.2	15.2	16.0	—	—	—	10	4.5	11.8	16.3	15.6		
Val		14.6	18.2	17.0	2.4	19.4	20.2	15.3	4.1	19.4	17	15.4	8.0	23.4	22.9		
Met		1.2	1.33	11.2	0	11.2	10.9	7.9	1.1	9.0	11	7.1	0.4	7.5	8.5		
Ile		21.5	24.4	24.6	2.3	26.9	26.7	14.2	3.3	17.5	16	13.0	3.3	16.3	17.4		
Leu		46.1	53.0	40.7	1.4	42.1	43.2	54.0	10.0	64.0	62	42.2	5.5	47.7	52.5		
Tyr		12.5	13.3	4.1	0	4.1	4.2	12.6	1.8	14.4	15	9.8	0.7	10.5	12.9		
Phe		7.7	9.0	10.4	2.0	12.4	12.9	5.0	1.9	6.9	6.5	7.6	1.6	9.2	11.1		
His		10.8	12.0	9.4	1.7	11.1	12.5	7.3	1.4	8.7	8.0	8.6	2.4	11.0	12.4		
Lys		35.0	53.0	40.2	23.8	64.0	61.4	47.0	13.4	60.4	62	43.1	28.8	71.9	74.0		
Arg		16.0	17.0	10.5	5.0	15.5	16.0	13.0	2.0	15.0	15	5.3	0	5.3	7.6		
GlcN		2.1	10.1	9.4	2.5	11.9	13.0	1.9	3.1	5.0	7	0	4.4	4.4	4.2

Fig. 4. Fragmentation of variant glycoprotein 99. Densitometer scans at 265nm of Coomassie Blue-stained polyacrylamide gels of variant glycoprotein 99 were treated with trypsin (one-seventieth by weight) for 5 min (a), 15 min (b) or 60 min (c), or untreated (d). Three vertical arrows indicate mol.wts. of 65000, 40000 and 20000. Top of gel is at left.

the f17 of variant glycoprotein 52. After 5 min incubation, all the intact variant glycoprotein had been degraded to fragments f48 (about 65%), f40 (about 10%) and f17 (about 25%). During continued digestion for 60 min the proportion of fragment f17 was unchanged but fragment f40 increased in a reciprocal manner to the decrease in fragment f48. In a larger-scale experiment, variant glycoprotein 99 at 5 mg/ml was digested with trypsin (0.14 mg/ml) for 20 min at 37°C. The mixture was rapidly cooled to 0°C and applied directly, in the light sucrose solution, to an electrofocusing column. Six detectable peaks were resolved with isoelectric points of 4.94, 5.67, 8.20, 8.56, 8.66 and 8.73. The three major components with pI values 5.67, 8.20 and 8.66 gave single bands on polyacrylamide gels with apparent mol.wts. of 20000, 40000 and 48000, corresponding to fragments f17, f40 and f48 respectively.

The *N*-terminal amino acid sequence of fragment f48 (residues 2–15) was identical with that of the intact variant glycoprotein (variant I in Bridgen *et al.*, 1976). The amino acid compositions of fragments f48 and f17 are shown in Table 2, and their sum agrees well with the composition of intact variant glycoprotein 99. In contrast with the situation with variant glycoproteins 52 and 55, cleavage of variant glycoprotein 99 leaves most of the glucosamine in the large fragment.

Fragmentation of variant glycoprotein 60

Analytical studies again showed complete breakdown of intact antigen to a 52000-mol.wt. fragment (f52) within 5 min. In initial experiments, fragment f52 survived for at least 60 min without further breakdown. As with other variant glycoproteins, however, a proportion of fragment f52 was sometimes degraded further to 45000 mol.wt. Fragment f52 and a small fragment (f13) of assigned mol.wt. 13000 but apparent mol.wt. 20000 on polyacrylamide gels were purified and analysed (Table 2). Of the total protein recovered from a Sephadex G-150 column, fragment f52 represented 70% and fragment f13 20%. Most of the glucosamine was in the smaller fragment, which had a pI of 4.4 on electrofocusing gels.

Fragmentation of variant glycoprotein 49

A preliminary experiment suggested that variant glycoprotein 49 was more resistant to cleavage than the other variants. Cleavage of the glycoprotein *in situ* on living trypanosomes was also inefficient. Cleavage of the isolated variant glycoprotein yielded two large fragments of mol.wts. 55000 (f55) and 45000 (f45). The kinetics of cleavage suggested stepwise removal of two 10000-mol.wt. fragments ($2 \times f10$). Even after 16 h incubation with trypsin, small amounts of fragment f55 remained. After digestion of 15 mg of variant glycoprotein with

0.75 mg of trypsin for 90 min at 37°C, the major peak on Sephadex G-150 contained a mixture of fragments f45 (about 75%) and f55 (about 20%) and intact variant glycoprotein (about 5%). Rechromatography of the main peak, excluding the leading edge, yielded fragment f45 slightly contaminated with fragment f55. Material from a lower-molecular-weight peak on Sephadex G-150 ran at the front of 10% polyacrylamide gels. This fraction contained all the glucosamine and none of the arginine present in the intact variant glycoprotein (Table 2). The lack of arginine suggested that trypsin digestion of denatured citraconylated variant glycoprotein 49 should yield a polypeptide of mol.wt. >20000, containing all the sugar of the parent variant glycoprotein. This suggestion was confirmed.

Fragmentation of variant glycoprotein 221

Variant glycoprotein prepared from variant 221 on three occasions showed a single band on SDS/polyacrylamide gels but in a position corresponding to a mol.wt. of 59000. This may be a genuine difference from other variant glycoproteins, but more probably represents extreme susceptibility to C-terminal degradation, as the isoelectric point of the major peak of variant glycoprotein 221 has differed on each occasion of its isolation.

Variant glycoprotein 221 (15 mg) was treated with 0.75 mg of trypsin for 20 min at 37°C. Trypsin was removed by passing the digest through a column of soya-bean trypsin inhibitor attached to Sepharose. The eluate was passed through a column of concanavalin A, also immobilized on Sepharose. Unretarded components and those eluted by 0.1 M- α -methyl D(+)-glucoside were separately chromatographed on Sephacryl S-200. Material eluted with α -methyl glucoside yielded a homogeneous component (f44) of mol.wt. 44000. This fragment was sequenced through the first 15 residues, and the sequence was identical with that of the intact variant glycoprotein. Neither fragment f44 nor the two purified small fragments have been further characterized. A point of interest, however, is that fragment f44 is retained by concanavalin A-Sepharose and two smaller C-terminal fragments are not. This is the converse situation of that obtained with variant glycoprotein 52.

The N-terminal amino acid sequence of variant glycoprotein 221 was determined by Dr. D. Stone and Mrs. S. Patterson to be as follows: Ala-Ala-Glu-Lys-Gly-Phe-Lys-Gln-Ala-Phe-Trp-Gln-Pro-Leu-Cys-Gln-Val-Ser-Glu-Glu-Leu-Asp-.

Discussion

As far as variant glycoprotein structure has been studied, the available evidence implies the existence of extensive amino acid-sequence variation between

antigenically distinct glycoproteins. The experiments described in this report were performed as part of a search for potentially conserved aspects of variant glycoprotein structure, which might lead us towards an understanding of the genetic basis of antigenic variation. Our results failed to reveal an obvious division of the variant glycoprotein into variable and constant regions, as is found for the immunoglobulins. However, our results suggest that tryptic cleavage of native variant glycoproteins distinguishes conformationally independent N-terminal and C-terminal domains. These studies were performed on six variant glycoproteins chosen at random from available clones of *T. brucei* strain 427. In each case, tryptic cleavage initially produced a large fragment (48000–55000 mol.wt.) which was, in some variant glycoproteins, completely or partially susceptible to further breakdown. For every variant glycoprotein, a large fragment (40000–52000 mol.wt.) was finally produced that was highly resistant to further degradation by trypsin. Large tryptic fragments from four variant glycoproteins were shown, by sequence analysis, to represent the N-terminal sections of the variant glycoproteins. Restricted access to sequencing facilities prevented analysis of fragments from the remaining two variant glycoproteins.

In every variant glycoprotein studied, the large N-terminal and smaller C-terminal fragments dissociated readily from each other, in the absence of denaturing agents, suggesting that they are independently folded and represent conformationally distinct domains. Each fragment contains disulphide-linked cysteine residues. There is no disulphide linkage between the trypsin-cleaved domains. The precise point (or points) of cleavage differs appreciably between each variant glycoprotein, suggesting an absence of total sequence and conformational homology in the cleavage zone.

The smaller C-terminal fragments arising from tryptic cleavage have not been characterized in detail. From their compositions it is clear that they cannot have an entirely common sequence, although a 50% homology of the approx. 150 residues represented in the C-terminal fragments would have been undetectable. The possible existence of homology in this region needs to be examined by detailed sequence studies. 'Fingerprinting' evidence alone was unconvincing.

No common features are persistently associated with each of the cleavage domains. One observation, which may be significant if upheld by studies of additional variant glycoproteins, is that, in those variants containing the lowest amounts of carbohydrate (49 and 52), it is all found in the C-terminal region. This fact has been exploited in other studies (Cross & Johnson, 1976; J. G. Johnson & G. A. M. Cross, unpublished work) to deduce the orientation of variant glycoprotein on the cell surface. Observations

show that the carbohydrate is present in the innermost layer of the surface coat, suggesting that the C-terminal region contains the site for interaction of variant glycoprotein with the plasma membrane. At the present time we can only speculate about the mode of attachment of variant glycoprotein to plasma membrane. The ease of release and solubility of the variant glycoprotein suggest that attachment could be by ionic interaction with a receptor (possibly a phospholipid or glycolipid) in the membrane. Alternatively, the variant glycoprotein could be covalently linked to a hydrophobic peptide tail which is inserted into the lipid bilayer. This possibility, which cannot yet be discounted, would demand the existence of a specific and rapid cleavage mechanism, causing almost instantaneous release of soluble glycoprotein. The possibility that such a mechanism might be involved is made more credible by the apparent susceptibility of variant glycoproteins to further proteolytic degradation at the C-terminus. It should be mentioned that C-terminal heterogeneity is a common occurrence in semipurified variant glycoprotein preparations.

The demonstration of strongly cross-reacting antigenic determinants on variant glycoproteins (Barbet & McGuire, 1978) may add further significance to the tryptic cleavage results. Recent studies (Cross, 1979) show that the cross-reacting determinants are localized in the C-terminal domain.

Comparison of the N-terminal amino acid sequence of variant glycoprotein 221 with those previously published (Bridgen *et al.*, 1976) shows no significant homology, except that, in all five glycoproteins studied, only alanine or threonine is found at the N-terminus.

J. G. J. was supported by the Medical Research Council of Canada and by the H. E. Durham Fund, King's College, Cambridge. We thank P. J. Bridgen, J. Bridgen, L. S. Davey, D. Stone and S. Patterson for their contributions to this work.

References

- Barbet, A. F. & McGuire, T. C. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1989–1993
- Belew, M., Porath, J., Fohlman, J. & Janson, J.-C. (1978) *J. Chromatogr.* **147**, 205–212
- Bridgen, P. J., Cross, G. A. M. & Bridgen, J. (1976) *Nature (London)* **263**, 613–614
- 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. (1978a) *Trends Biochem. Sci.* **3**, 49–51
- Cross, G. A. M. (1978b) *Proc. R. Soc. London Ser. B* **202**, 55–72
- Cross, G. A. M. (1979) *Nature (London)* in the press
- Cross, G. A. M. & Johnson, J. G. (1976) in *Biochemistry of Parasites and Host-Parasite Relationships* (van den Bossche, H., ed.), pp. 413–420, Elsevier/North-Holland, Amsterdam
- Doyle, J. J. (1977) in *Immunity to Blood Parasites of Animals and Man* (Miller, L. H., Pino, J. A. & McKelvey, J. J., eds.), pp. 31–63, Plenum Press, New York and London
- Hirumi, H., Doyle, J. J. & Hirumi, K. (1977) *Science* **196**, 992–994
- Johnson, J. G. & Cross, G. A. M. (1977) *J. Protozool.* **24**, 587–591
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Porter, R. R. (1973) *Science* **180**, 713–716
- Vickerman, K. (1969) *J. Cell Sci.* **5**, 163–194
- Vickerman, K. (1978) *Nature (London)* **273**, 613–617