Reprinted from *J. Mol. Biol.* (1991) **218**, 679–683

Crystallization and Preliminary X-ray Analysis of an Intact Soluble-form Variant Surface Glycoprotein from the African Trypanosome, *Trypanosoma brucei*

James A. Down, Scott C. Garman, Anne M. Gurnett, Mervyn J. Turner and Don C. Wiley

J. Mol. Biol. (1991) **218**, 679–683

Crystallization and Preliminary X-ray Analysis of an Intact Soluble-form Variant Surface Glycoprotein from the African Trypanosome, *Trypanosoma brucei*

James A. Down†, Scott C. Garman, Anne M. Gurnett, Mervyn J. Turner and Don C. Wiley

Department of Biochemistry and Molecular Biology
Harvard University, and Howard Hughes Medical Institute
7 Divinity Avenue, Cambridge, MA 02138, U.S.A.
and Merck, Sharp and Dohme Research Laboratories
Box 2000, Rahway, NJ 07065, U.S.A.

(Received 30 November 1990; accepted 2 January 1991)

The intact variant surface glycoprotein (VSG) ILTat 1·24 from *Trypanosoma brucei* has been crystallized. An amino-terminal domain of the protein comprising two thirds of the sequence had been crystallized previously after proteolytic digestion. Now intact VSG crystals have been grown from 50 mM-Mes (pH 6·5) containing 62% (w/v) saturated ammonium sulfate. The crystals are demonstrated to contain the intact VSG by h.p.l.c. gel filtration and reaction with an antibody to the inositol phosphate oligosaccharide on the VSG carboxy terminus. The space group of the crystals is $P6_222$ (or $P6_422$) with unit cell dimensions $a = b = 184$ Å and $c = 214$ Å. Preparative isoelectric focusing may have facilitated crystallization.

Trypanosoma brucei is a parasitic protozoan that inhabits the bloodstream and tissue spaces of its mammalian host. *T. brucei brucei* causes nagana, a disease that is endemic in cattle throughout a large area of equatorial Africa. The related subspecies, *T. brucei rhodesiense* and *T. brucei gambiense* cause African sleeping sickness in humans. Both diseases are characterized by a relapsing parasitemia, associated with the expression of antigenic variation within the parasite population (for a review, see Turner, 1984). Different antigenic variants express distinct variant surface glycoproteins, or VSGs‡, which form a close packed matrix visible in transmission electron microscopy as an electron-dense layer, 10 to 15 nm thick (Vickerman & Luckins, 1969).

The biochemistry and molecular biology of the VSGs has been extensively reviewed (Cross, 1990; Turner, 1985; Pays & Steinert, 1988; Donelson, 1987) and is only briefly summarized here. VSGs are glycoprotein dimers containing about 450 amino acids in the monomeric subunit. Sequence diversity is extensive, producing anything between 10% and 90% homology. Despite this remarkable sequence diversity, several features of the molecules are conserved. Sequence variability is highest in the amino-terminal two-thirds of the molecule, while sequence similarity in the carboxy-terminal segment of the mature polypeptide has been used to establish the existence of at least two VSG classes, I and II. The carboxy- and amino-terminal portions are separated by a protease-susceptible region that allows cleavage into two domains by exogenous or endogenous protease. VSGs are anchored to the membrane through glycophosphatidyl-inositol attached to the C-terminal amino acid (Asp or Asn for class I; Ser for class II). The molecule can be solubilized by activation of an endogenous phospholipase C, which removes dimyristoylglycerol and exposes an antigenic determinant within the glycolipid that is common to all VSGs, known as the cross-reacting determinant (CRD). Southern

† Current address: Becton Dickinson and Co., Research Center, P.O. Box 12016, Research Triangle, NC 27709, U.S.A.

‡ Abbreviations used: VSGs, variant surface glycoproteins; SDS/PAGE, sodium dodecyl sulfate/polyacrylamide gel electrophoresis; ILTat, International Laboratory for Research on Animal Diseases trypanosome antigen type; MITat, Molteno Institute trypanosome antigen type; CRD, cross-reacting determinant; TBS, tris-buffered saline; IEF, isoelectric focusing; PVDF, polyvinylidene difluoride; Mes, 2-(*N*-morpholino)ethane sulfonic acid; p*I*, isoelectric point; glu-c, *Staphylococcus aureus* V-8 endoproteinase glu-c; h.p.l.c., high pressure liquid chromatography.

0022-2836/91/080679-05 $03.00/0

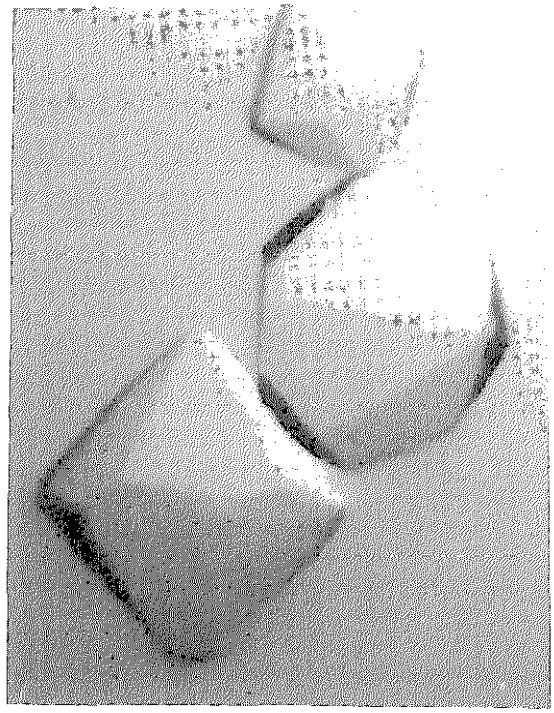

Figure 1. Photomicrograph of crystals of ILTat 1·24. Crystals were initially produced in hanging drops (McPherson, 1985), set up in 24-well tissue culture plates (Linbro), by mixing 1 μl of protein solution (40 mg/ml) and 1 μl of reservoir buffer containing 50 mM-Mes (pH 6·5), 60 to 62% saturated with $(NH_4)_2SO_4$ and including 0·01% (w/v) NaN_3. Single drops were equilibrated over 1 ml reservoirs on silanized glass coverslips (Aquasil; Pierce). Larger, better diffracting crystals have been grown by slowing the rate of growth through the addition of 0·25 M-NaCl or Tris·HCl to a final concentration of 0·07 M. The crystals shown are 0·1 mm in their longest dimension.

hybridization experiments suggest that the trypanosome may contain about 1000 different VSG genes.

Crystal structures of the N-terminal domains of two VSGs have been reported at 6 Å (1 Å = 0·1 nm) resolution (Freymann *et al.*, 1984; Metcalf *et al.*, 1986) and have been determined to 2·9 Å resolution (Freymann *et al.*, 1990), revealing an unusual 100 Å long rod-like molecule with a central core of long α-helices. A number of other VSG N-terminal domains have also been crystallized (Metcalf *et al.*, 1988). Crystals of an intact VSG, including the C-terminal one-third of the sequence have now been produced from the class I VSG ILTat 1·24.

To prepare VSG, trypanosomes of ILTat 1·24 serodeme were purified from the blood of infected rats according to the method of Lanham & Godfrey (1970). Soluble VSG was prepared as described by Cross (1975) and used for crystallization directly after the DEAE-cellulose purification or following additional purification by preparative isoelectric focusing (IEF) in a Rotofor IEF apparatus (Biorad) according to the manufacturers instructions.

The crystals of whole ILTat 1·24 were obtained under conditions similar to those previously found to crystallize the amino-terminal domain (20 mM-Tris·HCl (initial pH 7·4); 50 mM-Mes (initial pH 6·5); saturated to 60 to 64% (w/v) with $(NH_4)_2SO_4$) after a VSG preparation was subjected to fractionation on DE52 cellulose. In contrast to the amino-terminal domain of ILTat 1·24, which crystallized as octahedra in clear liquor (Metcalf *et al.*, 1988), crystals of the intact VSG were hexagonal bipyramids (Fig. 1) and typically grew to about 100 μm^3 in wells containing large amounts of precipitated protein. The crystals were found by SDS/PAGE to contain predominantly a 53,000 M_r protein, identical in size to the original intact VSG rather than the 42,000 M_r amino-terminal domain. (These M_r values were previously reported as 56·2 and 40·7 kDa, respectively, corresponding to sizes observed on larger gels of lower percentage polyacrylamide (Metcalf *et al.*, 1988)). Along with the whole 53,000 M_r VSG, polypeptides of 49,000 and 37,000 were observed in lesser quantities. Amino-terminal sequences obtained by gas-phase Edman degradation after electrophoretic transfer to PVDF membrane (Matsudira, 1987) showed the amino termini of the 53,000 and 49,000 M_r protein bands corresponded to that of intact ILTat 1·24 (THFGV . . .: Metcalf *et al.*, 1988). For a single crystal approximately 0·5 mm^3 in size, the relative yields of protein were 25 pmol for the intact species and 1 to 2 pmol for the 49,000 M_r fragment (the quantity of the 37,000 M_r protein was below the level of detection) suggesting that the crystals contained intact 53,000 M_r protein and the 49,000 M_r fragment in a ratio of approximately 10 to 1. Because previous results have shown that the carboxy-terminal portions of VSGs are susceptible to proteolytic cleavage by endogenous proteases (Metcalf *et al.*, 1988) and since the crystallization drops initially contained only intact 53,000 M_r ILTat 1·24 the smaller polypeptides were considered to be proteolytic products.

It was subsequently found that purification of ILTat 1·24 by preparative IEF diminished precipitation during crystallization and eliminated proteolysis, as determined by the absence of the 49,000 and 37,000 M_r proteolytic fragments in fresh crystals analyzed by SDS/PAGE. Thus, it seems that preparative IEF reduced the effect of a protease contaminant. Our experiments also suggest that isoelectric focusing may have facilitated crystallization of the intact molecule of ILTat 1·24. Preparations that were not purified as described above did not yield crystals of intact ILTat 1·24. Furthermore, peak fractions of the Rotofor IEF profile (pH = 5·7 to 6·1) crystallized within two weeks in 50 mM-Mes (pH 6·5), containing 60 to 62% saturated $(NH_4)_2SO_4$, whereas outlying fractions isolated at lower pH values of 5·0 to 5·6 crystallized under similar conditions but only after three to six months.

It was previously reported that ILTat 1·24 underwent breakdown by endogenous protease(s) to

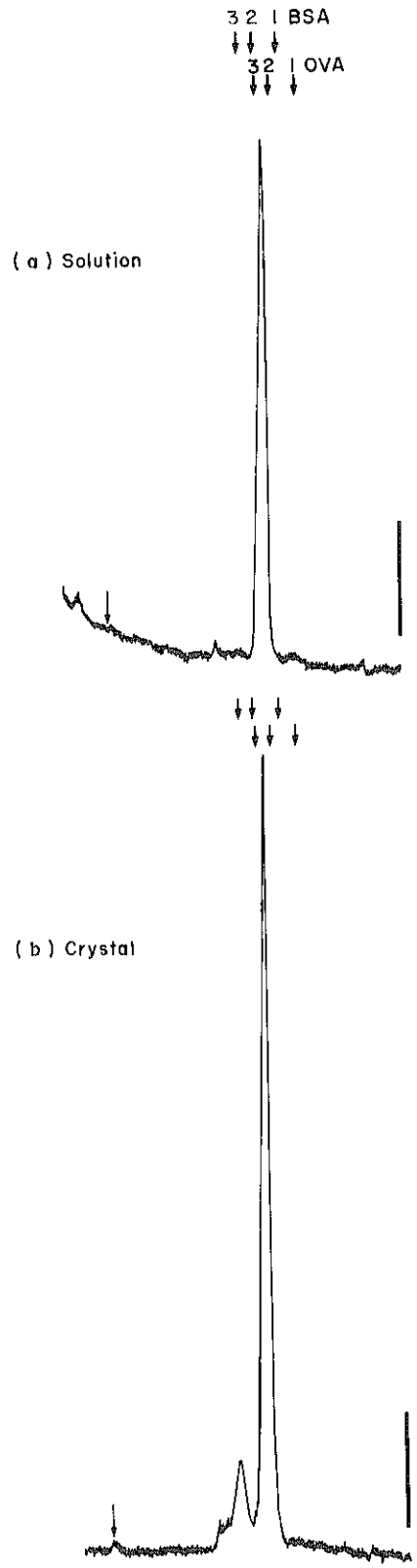

Figure 2. h.p.l.c. analysis of crystallized ILTat 1·24. Gel filtration h.p.l.c. was performed isocratically on a Waters h.p.l.c. system using a Protein Pak 300 column (Waters) pumped at 0·5 ml min^{-1}. The mobile phase consisted of MBS (0·15 M-NaCl, 0·025 M-Mes (pH 6·5), containing 0·1 % NaN_3). The column was calibrated with glutaraldehyde-crosslinked ovalbumin or bovine serum albumin. Lyophilized ILTat 1·24 was rehydrated directly in MBS buffer (a) and after X-ray data collection, a crystal of ILTat 1·24 was dissolved in MBS (b) and both samples were analyzed by gel filtration h.p.l.c. as described. Elution times were 13 : 00 and 13 : 02 (min : s), respectively. Migration of calibration standards are shown as arrows: bovine serum albumin (BSA) monomer (1), 65,000 M_r, 14 : 32; dimer (2) 130,000 M_r, 12 : 26; trimer (3) 195,000 M_r, 11 : 10; ovalbumin (OVA) monomer (1), 43,000 M_r, 15 : 45; dimer (2), 86,000 M_r, 13 : 48; and trimer (3), 129,000 M_r, 12 : 30. Sample injections are indicated by arrows on the baseline. The vertical bar indicates 0·1 o.d. units. The small peak at 10 : 35 in (b) is a high M_r oligomer of ILTat 1·24 and is possibly an artifact of X-ray irradiation.

produce amino-terminal domain fragments, which subsequently crystallized (Metcalf *et al.*, 1986, 1988). In contrast, data quality crystals of the intact protein have now been grown within two weeks from peak IEF fractions, starting as small (<100 μm) crystals that were observed within one or two days. The rate of ILTat 1·24 VSG breakdown is slowest compared to other variants that are currently under study in this laboratory (i.e., MITat 1·2, ILTat 1·23, and ILTat 1·25), which may be another factor contributing to our ability to crystallize the whole molecule.

Because of our previous inability to obtain crystals of intact VSGs it was necessary to establish unequivocally that the crystals described here contain intact ILTat 1·24. Three criteria suggest that this is so. The M_r of the ILTat 1·24 in crystals analyzed both by size exclusion h.p.l.c. (Fig. 2) and by SDS/PAGE (Fig. 3) was identical to that of the original intact protein. Crystallized and intact VSG had identical amino-terminal sequences (see above) as well as intact carboxy termini as shown by anti-CRD antibody reactivity on immunoblots (Fig. 3).

Still and precession photographs of ILTat 1·24 crystals were obtained with a Supper precession camera on Elliot GX-6 and GX-13 rotating anode X-ray generators operated at 19 mA and 39 kV, or 60 mA and 40 kV, respectively with Franks mirror focusing optics. X-ray data were collected on a Xentronics multiwire area detector (Durbin *et al.*, 1986) and analyzed with BUDDHA software (Blum *et al.*, 1987). Initially, the crystals of intact ILTat 1·24 diffracted to approximately 6·0 Å but improved growth conditions (see the legend to Fig. 1) have consistently produced crystals that diffracted to about 3·2 Å resolution. Their hexagonal shape suggested a trigonal or hexagonal space group. *hk*0, *h*0*l* and *hhl* precession photographs reveal a hexagonal lattice with 6/mmm symmetry. Systematic absences along the c^* axis are consistent with the presence of a 6_2 screw axis, establishing the space group as $P6_222$ (or $P6_422$). The unit cell dimensions are $a = b = 184$ Å and $c = 214$ Å. Cell constants have been observed to vary by about 0·5% on both a and c.

Based on a calculated mass of 50,200 M_r for the 468 amino acid residues, plus 2000 M_r for the phosphatidylinositol glycan, the molecular mass of an

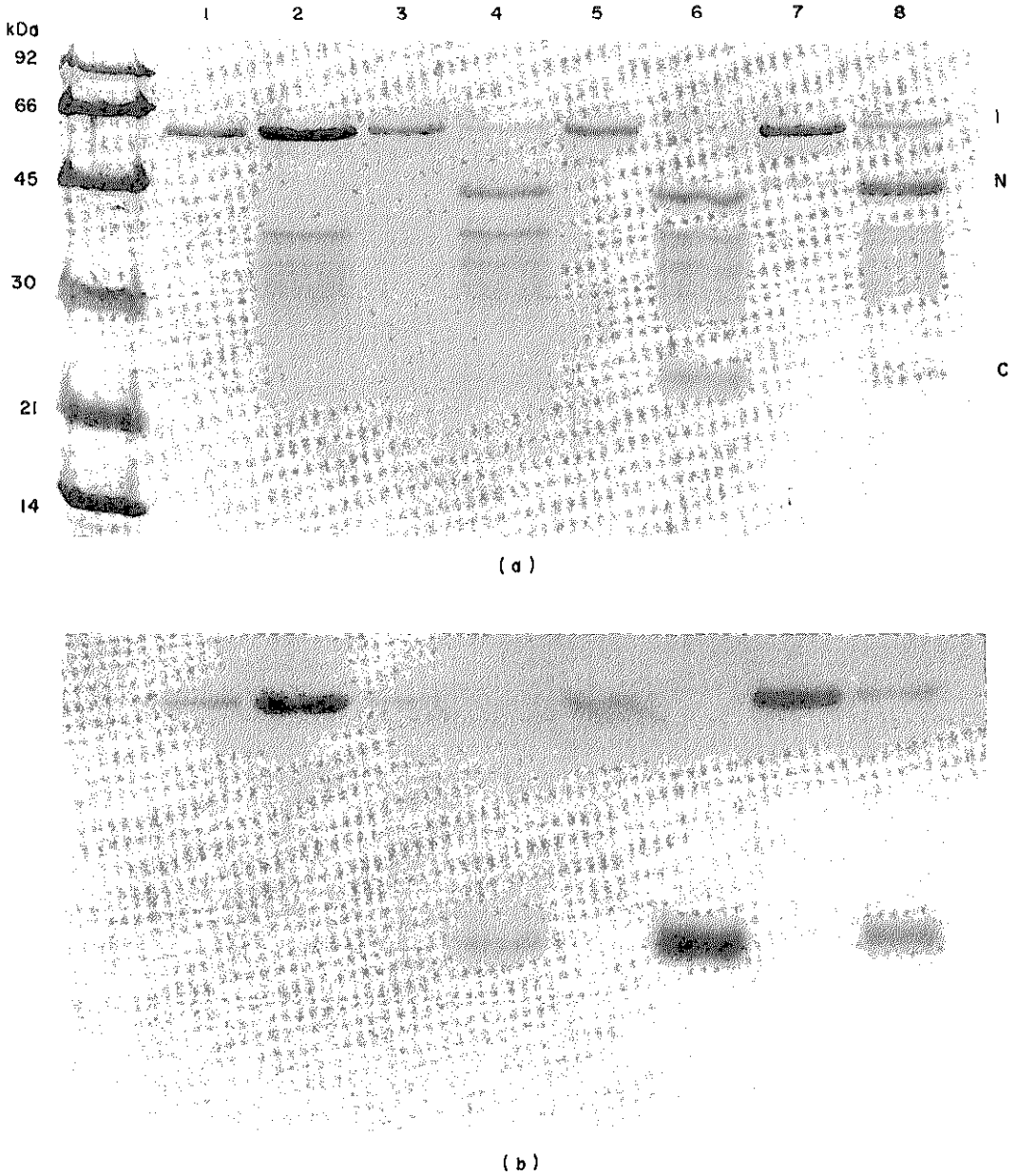

Figure 3. Comparison of reactivity toward anti-CRD antibody of ILTat 1·24 in crystals and solution and after treatment with *Staphylococcus aureus* V8 endoproteinase glu-c. Prior to SDS/PAGE analysis, crystals were washed by transfer through 10 × 2-ml harvest buffer consisting of 50 mM-Mes (pH 6·5) and $(NH_4)_2SO_4$ to 75% or 80% saturation. The crystals were transferred by a pulled glass capillary with a fresh capillary used for each transfer. Crystals were incubated in "crystal harvesting buffer" (80% $(NH_4)_2SO_4$, 50 mM-Mes (pH 6·5)) or "dissolving buffer" (50 mM-Tris (pH 8·1)) for 4 h at room temperature in the presence or absence of glu-c (0·02 units, Boehringer Mannheim), which cleaves ILTat 1·24 into amino and carboxy-terminal domains. The reactions were compared to identical control reactions performed with rehydrated, uncrystallized protein and analyzed by SDS/PAGE (Laemmli, 1970) and subsequent Western blotting (see below). Lanes 1, 2: VSG in intact crystals incubated in crystal harvesting buffer were not cleaved by glu-C (2) and retained the same M_r as protein in crystals that were not incubated with the protease (1). Lanes 3, 4: ILTat 1·24 crystal in dissolving buffer was proteolyzed by glu-c to produce N and C-terminal domains (4) while the control in the absence of enzyme remained intact (3). Lanes 5, 6: ILTat 1·24 rehydrated in crystal harvesting buffer in the absence (5) or presence (6) of glu-C. Lanes 7, 8: ILTat 1·24 rehydrated in dissolving buffer in the absence (7) or presence (8) of glu-C. (a) Shows the fixed, stained gel before re-incorporation of SDS and Western transfer. (b) Shows the autoradiogram of the Western blot after reaction with rabbit polyclonal anti-CRD antibodies followed by ^{125}I-labeled *S. aureus* protein A. I indicates the migration of the intact VSG, N indicates the position of the amino-terminal domain and C indicates the position of the carboxy-terminal domain. Migration of the molecular weight calibration standards are shown at the left. To visualize VSGs on both SDS/PAGE and Western blots, they were first run on SDS/PAGE in 12% (w/v) polyacrylamide gels, which were then fixed, stained and photographed. For Western transfer (Towbin *et al.*, 1979), SDS was re-incorporated into the gels by soaking in SDS/PAGE tank buffer, then VSGs were blotted onto PVDF membrane in an ice-cooled mini-blot apparatus (Biorad) containing SDS/PAGE tank buffer at 10 V, 60 mA for 1 h. Non-specific binding was blocked with 0·15 M-NaCl containing 20 mM-Tris (pH 7·4) and 5% (w/v) non-fat dry skim milk powder (Carnation) (Johnson *et al.*, 1984). The blots were then reacted for 1 h with a 1 : 100 dilution of polyclonal VSG-absorbed rabbit anti-VSG CRD antibody in TBS/milk, then washed with TBS/milk. Specifically bound antibodies were detected by ^{125}I-labeled *S. aureus* protein A and subsequent autoradiography.

ILTat 1·24 monomer is estimated at 52,200 M_r. The amino-terminal domain of ILTat 1·24 crystallized as dimers (Metcalf *et al.*, 1988) and protein from dissolved crystals of intact ILTat 1·24 migrate on size exclusion h.p.l.c. as dimers (Fig. 2). If the solvent content were about 55% (as found in the N-terminal domain crystals), the crystals of intact ILTat 1·24 would contain two dimers per asymmetric unit.

The present crystals have a relatively short lifetime (8 h) in the X-ray beam. Preliminary experiments indicate that after introduction of ethylene glycol as a cryo-protectant (Petsko, 1975; Douzou *et al.*, 1975) the crystals continue to diffract when cooled to 100°K in a stream of cryo-cooled N_2 (Hope, 1988; Hope *et al.*, 1989).

We thank our colleagues at Harvard University and Merck, Sharp and Dohme, especially Edward Hayes, Paula Dulski, Stefan Galuska and Richard Crouse for helpful support, William S. Lane of the Harvard Microchemistry Facility for sequence analyses and Denise Boldt of Becton Dickinson Information Services for literature search. We also thank Terry Pearson, Michael Clarke, John Donelson and Mark Carrington for helpful discussions. J.A.D. was supported by a fellowship from Merck, Sharp and Dohme. This work was supported in part by the National Institutes of Health, grant AI-21324 and the Howard Hughes Medical Institute (to D.C.W.) and was initiated by the United Nations Developmental Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Disease (Project nos. 810278 to D.C.W. and 800328 to M.J.T.).

References

Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. (1987). A system for collection and on-line integration of X-ray diffraction data from a multiwire area detector. *J. Appl. Crystallogr.* **20**, 235–242.

Cross, G. A. M. (1975). Identification, purification and properties of clone specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology*, **71**, 393–417.

Cross, G. A. M. (1990). Cellular and genetic aspects of antigenic variation in trypanosomes. *Annu. Rev. Immunol.* **8**, 83–110.

Donelson, J. E. (1987). Antigenic variation in African trypanosomes. *Contrib. Microbiol. Immunol.* **8**, 138–175.

Douzou, P., Hoa, G. H. & Petsko, G. A. (1975). Protein crystallography at sub-zero temperatures: lysozyme–substrate complexes in cooled mixed solvents. *J. Mol. Biol.* **96**, 367–380.

Durbin, R. M., Burns, R., Moulai, J., Metcalf, P., Freymann, D., Blum, M., Anderson, J. E., Harrison, S. C. & Wiley, D. C. (1986). Protein, DNA and virus crystallography with a focused imaging proportional counter. *Science*, **232**, 1127–1132.

Freymann, D., Metcalf, P., Turner, M. & Wiley, D. (1984). 6 Å X-ray structure of a variable surface glycoprotein from *Trypanosoma brucei*. *Nature (London)*, **311**, 167–169.

Freymann, D., Down, J., Carrington, M., Roditi, I., Turner, M. & Wiley, D. (1990). The 2·9 Å structure of the N-terminal domain of a VSG from *Trypanosoma brucei*. *J. Mol. Biol.* **216**, 141–160.

Hope, H. (1988). Cryocrystallography of biological macromolecules: a generally applicable method. *Acta Crystallogr. sect. B*, **44**, 22–26.

Hope, H., Frolow, F., von Bohlen, K., Makowski, I., Kratky, C., Halfon, Y., Danz, H., Webster, P., Bartels, K. S., Wittmann, H. G. & Yonath, A. (1989). Cryocrystallography of ribosomal particles. *Acta Crystallogr. sect. B*, **45**, 190–199.

Johnson, D. A., Gautsch, J. W., Sportsman, J. R. & Elder, J. H. (1984). Improved technique utilizing non-fat dry milk for analysis of proteins and nucleic acids transferred onto nitrocellulose. *Gene Anal. Tech.* **1**, 3–8.

Laemmli, U. K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature (London)*, **227**, 680–685.

Lanham, S. M. & Godfrey (1970). Isolation of Salivarian trypanosomes from man and other mammals using DEAE-cellulose *Expt. Parisitol.* **28**, 521–534.

Matsudira, P. (1987). Sequence from picomole quantities of protein electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**, 10035–10038.

McPherson, A. (1985). Crystallization of macromolecules: general principles. *Methods Enzymol.* **114**, 112–125.

Metcalf, P., Blum, M., Freymann, D., Turner, M. J. & Wiley, D. C. (1986). Two variant surface glycoproteins of *Trypanosoma brucei* of different sequence classes have similar 6 Å resolution X-ray structures. *Nature (London)*, **325**, 84–86.

Metcalf, P., Down, J. A., Turner, M. J. & Wiley, D. C. (1988). Crystallization of the amino-terminal domains and domain fragments of variant surface glycoproteins from *Trypanosoma brucei brucei*. *J. Biol. Chem.* **263**, 17030–17033.

Pays, E. & Steinert, M. (1988). Control of antigen expression in African Trypanosomes. *Annu. Rev. Genet.* **22**, 107–126.

Petsko, G. A. (1975). Protein Crystallography at sub-zero temperatures: cryoprotective mother liquors for protein crystals. *J. Mol. Biol.* **96**, 381–392.

Towbin, H., Staehlin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. *Proc. Nat. Acad. Sci., U.S.A.* **76**, 4350–4354.

Turner, M. J. (1984). Antigenic variation in its biological context. *Phil. Trans. Roy. Soc. ser. B*, **307**, 27–40.

Turner, M. J. (1985). Structure and variation within variant surface glycoproteins of *Trypanosoma brucei*. *Ann. Inst. Pasteur Immunol.* **136C**, 41–49.

Vickerman, K. & Luckins, A. G. (1969). Localization of variable antigens in the surface coat of *Trypanosoma brucei* using ferritin-conjugated antibody. *Nature (London)*, **224**, 1125–1126.

Edited by A. Fersht