Biochem. J. (1991) 274, 503–507 (Printed in Great Britain)

The secondary structure of Protein G′, a robust molecule

Christopher R. GOWARD,*‡ Laurence I. IRONS,† Jonathan P. MURPHY* and Tony ATKINSON*
*Division of Biotechnology and †Division of Biologics, P.H.L.S. Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wilts. SP4 0JG, U.K.

The secondary structure of recombinant streptococcal Protein G′ was predicted and compared with spectropolarimetric data. The predicted secondary structure consisted of 37±4% α-helix and 30±5% β-sheet, whereas the values obtained from c.d. data were 29±2% α-helix and 41±3% β-sheet. An α-helix–β-sheet/turn–α-helix motif is conjectured to comprise the Fc-binding unit. The c.d. spectra in the near u.v. and far u.v. show that the Protein G′ molecule is stable to heating at 100 °C and to extremes of pH (pH 1.5 to 11.0). The protein retained biological activity at these extremes. The molecule uncoils above pH 11.5 in a time-dependent fashion. Unfolding of the molecule in guanidinium chloride was monitored by c.d. and fluorescence emission; 3 M-guanidinium chloride was required to unfold the protein by 50%. The protein was completely unfolded in 5.5 M-guanidinium chloride and fully refolded with restoration of activity after removal of guanidinium chloride.

INTRODUCTION

Protein G is located on the cell surface of group C and G streptococci and binds subclasses of IgG from many species (Åkerström *et al.*, 1985; Guss *et al.*, 1986; Reis *et al.*, 1986). Protein G has regions that bind albumin and the Fc portion of IgG. It binds to the Fab portion of IgG but with an affinity 10-fold lower than determined for the Fc regions (Björck & Kronvall, 1984), and also interacts with $F(ab')_2$ fragments (Erntell *et al.*, 1988). The complete nucleotide sequence of the Protein G structural gene from *Streptococcus* G148 has been reported and the protein was shown to consist of repetitively arranged domains by Guss *et al.* (1986) and Olsson *et al.* (1987). The amino acid residues involved in binding have not been identified, though a fragment of M_r 7500 binds IgG (Guss *et al.*, 1986).

Protein A from *Staphylococcus aureus* is another IgG-binding protein also composed of repetitively arranged domains (Sjödahl, 1977). The secondary structure was estimated by c.d. spectroscopy to consist of 31% α-helix and 13% β-sheet (Lindmark, 1982). The amino acid residues responsible for binding to IgG have been located by X-ray crystallography of an IgG-binding unit, which also showed the absence of β-sheet (Deisenhofer *et al.*, 1978; Deisenhofer, 1981). The stability of the molecule has been demonstrated by spectropolarimetric and spectrophotometric studies (Sjöholm, 1975). Protein G and Protein A have similar functions, but gene and amino acid sequence similarity is only found in the region probably involved in anchorage to the cell membrane.

We have cloned the Protein G gene and expressed a truncated protein, Protein G′ (Goward *et al.*, 1990). This recombinant protein was shown to bind only the Fc portion of IgG, and regions responsible for albumin binding and Fab binding, in addition to the cell-wall-spanning and membrane-anchoring portion, were removed. Homogeneous Protein G′, consisting of three putative IgG-binding domains, was prepared free from fragments of the major protein with a predicted M_r of about 20000 and a pI of 4.19 (Goward *et al.*, 1990).

The aim of the work was to initiate a study of the physical structure of Protein G′. In the present study we report secondary-structure predictions and compare these with c.d. data. We speculate on the structure of the IgG-binding unit and compare this with the known structure of a Protein A IgG-binding unit (Deisenhofer *et al.*, 1978; Deisenhofer, 1981). We also demonstrate the stability of Protein G′ in a number of harsh environments by c.d., fluorescence emission and e.l.i.s.a.

EXPERIMENTAL

Materials

Chromatography media were from Pharmacia–LKB (Uppsala, Sweden), and all reagents, including grade 1 guanidinium chloride (Gdn,HCl) were from Sigma Chemical Co. Milli-Q water (Millipore Corp.) was filtered through a 0.22 μm-pore-size filtration membrane.

Bacterial culture and protein purification

Escherichia coli containing the recombinant Protein G′ gene were cultured as previously described (Goward *et al.*, 1990). Homogeneous Protein G′ was prepared by affinity chromatography on IgG–Sepharose 4B, then anion-exchange chromatography by f.p.l.c. (Pharmacia–LKB) on Mono Q HR as previously described (Goward *et al.*, 1990), except that 170 mg of IgG–Sepharose 4B eluate was applied to a 20 ml Mono Q HR 16/10 column. The protein was dialysed exhaustively against Milli-Q water, then freeze-dried.

Protein concentration

Protein solutions were made by dilution of a known weight of salt-free freeze-dried Protein G′ in Milli-Q water. The molar concentration of Protein G′ was checked by analysis of amino acid composition in a Biotronik LC5000 Amino Acid Analyser after hydrolysis in 6 M-HCl at 160 °C. Protein concentrations were also determined by the Folin method of Lowry *et al.* (1951), with BSA as the standard. The concentration of functional Protein G′ was determined by an e.l.i.s.a. procedure as described previously (Goward *et al.*, 1990).

Secondary-structure predictions

Predictions of the secondary structure of Protein G′ were made by using the Robson–Garnier method (Garnier *et al.*, 1978) and the Chou–Fasman method (Chou & Fasman, 1978).

Abbreviations used: Gdn,HCl, guanidinium chloride.

‡ To whom correspondence and reprint requests should be addressed.

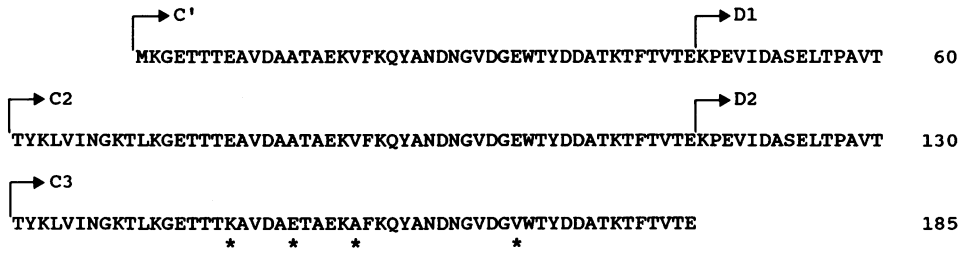

Fig. 1. Primary structure of Protein G′

The sequence of amino acid residues is grouped into the three putative IgG-binding domains, C′, C2 and C3, and may also include the D regions, which have been described as 'spacer regions' (Guss *et al.*, 1986; Olsson *et al.*, 1987). Translation is initiated part-way into the C′ domain from a TTG codon producing a methionine residue, whereas TTG codons in C2 and C3, in the same relative position, are translated to produce leucine residues. The remainder of C′/D1 and C2/D2 are identical. Several residues (*) are not conserved in the C3 domain.

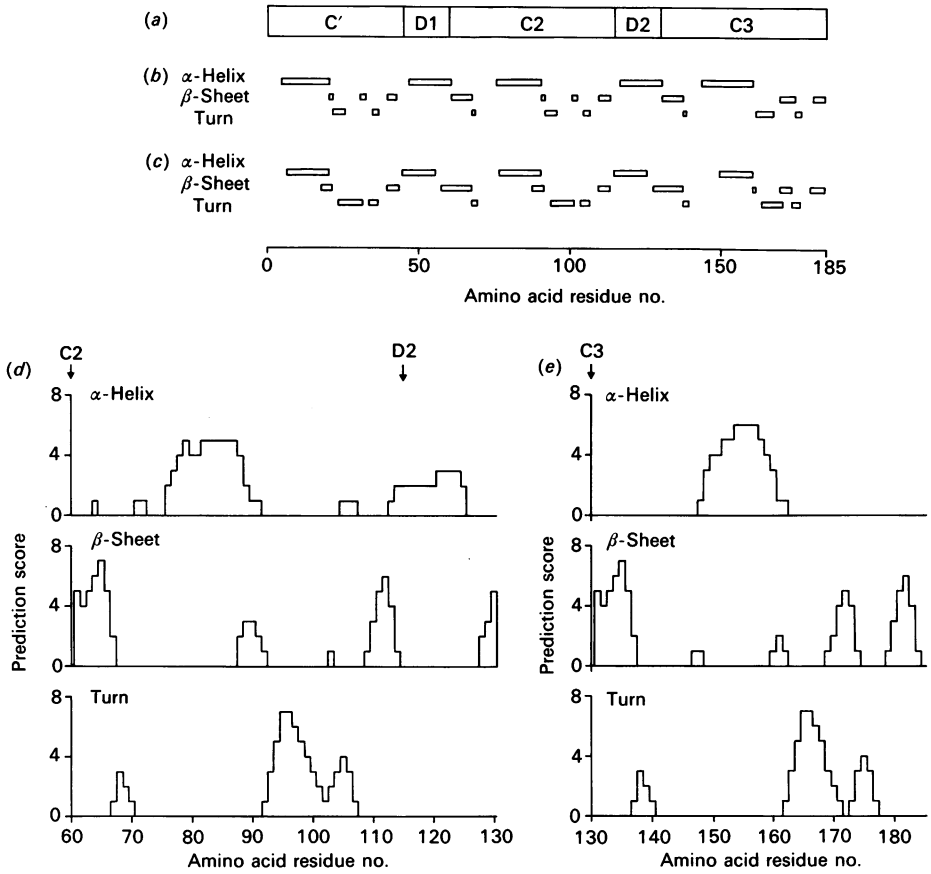

Fig. 2. Secondary-structure predictions of Protein G′

(*a*) Schematic representation of the Protein G′ molecule. (*b*) Robson–Garnier computer predictions of secondary structure (Garnier *et al.*, 1978). (*c*) Chou–Fasman computer predictions of secondary structure (Chou & Fasman, 1978). Chou–Fasman computer prediction scores are shown for the secondary-structure components of (*d*) C2 and D2 and (*e*) C3.

C.d. spectroscopy

C.d. spectra were obtained at room temperature (20–23 °C) with a JASCO J-600 spectropolarimeter calibrated with 0.6% (w/v) ammonium (+)-camphor-10-sulphonate (Takakuwa *et al.*, 1985). The spectra were recorded in the near u.v. (320 to 250 nm) and the far u.v. (250 to 190 nm) in cylindrical cells of pathlength 0.2 cm and 0.02 cm respectively. Each spectrum was the mean of four accumulated scans obtained with a time constant of 4 s and a scan speed of 10 nm/min. Unspecified solvent dichroic absorbances were subtracted from the spectra by computer manipulation. The final spectra were converted into mean residue ellipticities by using a mean residue weight of 108.7 for Protein G′. The amount of secondary structure was estimated from the far-u.v. c.d. spectra by using the computer program CONTIN version 2DP of Provencher & Glöckner (1981).

Fluorescence emission

Protein G′ was dissolved in 10 mM-potassium phosphate buffer, pH 7.0, containing various concentrations of Gdn,HCl. The final concentration of Protein G′ was 2.6 μg/ml. The fluorescence emission was measured at 345 nm after excitation at

Table 1. Percentages of α-helix and β-structures in Protein G′

Predicted values were estimated from data in Fig. 2. Calculated values were calculated from far-u.v. c.d. data.

	α-Helix (%)	β-Structure (%)
Predicted		
Robson–Garnier	41	35
Chou–Fasman	33	25
Calculated after treatment of sample		
20 °C at pH 5	30	44
100 °C at pH 5	26	40
pH 1.5	27	45
pH 5	30	44
pH 7	27	38
pH 10	28	39
pH 12	30	38
pH 12 for 6 days, then pH adjusted to 7	15	44
After removal of 5.5 M-Gdn,HCl	32	36

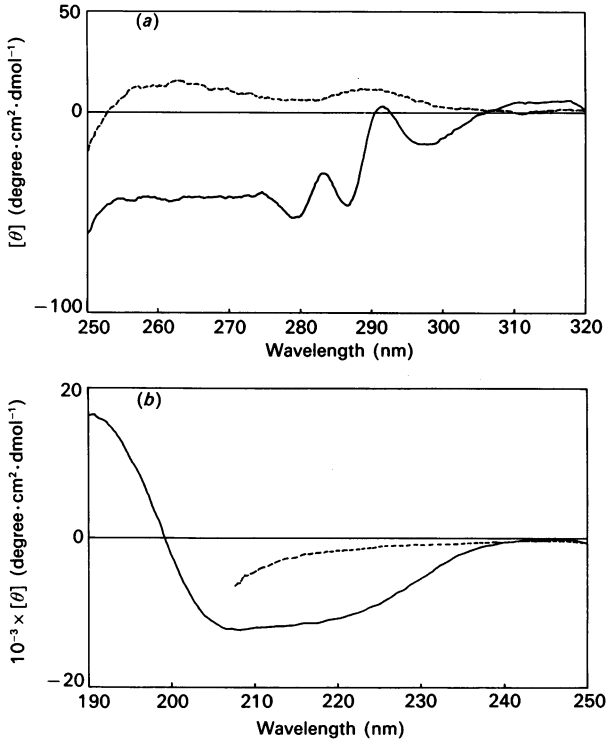

Fig. 3. C.d. spectra of native and unfolded Protein G′

(*a*) Near-u.v. and (*b*) far-u.v. c.d. spectra of Protein G′ in 10 mM-potassium phosphate buffer, pH 7.0, in the presence (-----) and absence (——) of 5.5 M-Gdn,HCl. The protein concentrations were 2.675 mg/ml for the near-u.v. spectra and 0.355 mg/ml for the far-u.v. spectra. Samples were incubated at 20 °C for 24 h before scanning. The background noise due to Gdn,HCl is excessive below 208 nm.

295 nm and at 20 °C in a Perkin–Elmer LS-5 Luminescence Spectrometer. Spectra were recorded 24 h after addition of Gdn,HCl to the protein. The concentrations of Gdn,HCl solutions were confirmed by measurement of refractive index and using the data of Nozaki (1972).

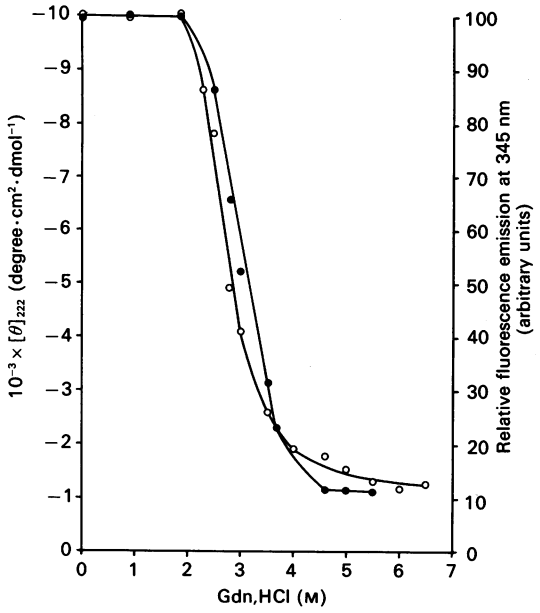

Fig. 4. Unfolding of Protein G′ in Gdn,HCl

Protein G′ (0.355 mg/ml) was incubated for 24 h at 20 °C in 10 mM-potassium phosphate buffer, pH 7.0, containing various concentrations of Gdn,HCl. C.d. spectra were recorded in the far u.v. and the mean ellipticity at 222 nm (●) was calculated. Fluorescence emission intensity (○) of Protein G′ (2.6 μg/ml) was monitored at 345 nm after excitation at 295 nm.

RESULTS AND DISCUSSION

Predictions of Protein G′ secondary structure

The primary structure of the Protein G′ molecule shows the repetitive nature of its domains (Fig. 1). Structural analyses of the molecule by using the Robson–Garnier prediction algorithms (Garnier *et al.*, 1978) and the Chou–Fasman prediction algorithms (Chou & Fasman, 1978) indicate an α-helix content of 37±4% and a β-structure content of 30±5% (Fig. 2). It should be noted that the accuracy of secondary-structure predictions is only about 50–55% (Kabsch & Sander, 1983; Nishikawa, 1983). However, c.d. analysis of Protein G′ suggests an α-helix content of 29±2% and a β-sheet content of 41±3% (Table 1). Estimation of α-helix and β-sheet contents from c.d. data of a number of proteins was demonstrated to be quantitatively accurate (Provencher & Glöckner, 1981). C.d. analysis of Protein G′ underestimates α-helix content and overestimates β-sheet content compared with secondary-structure predictions. C.d. has, however, also been noted to underestimate α-helix content and overestimate β-sheet content compared with X-ray crystal structures (Bayley *et al.*, 1989). Preliminary n.m.r. data confirm the presence of a significant amount of α-helical secondary structure in Protein G′ (L. Y. Lian, J. Derrick & G. C. K. Roberts, personal communication). Secondary analysis of the areas between the strongly predicted β-pleated structures by using Chou–Fasman or Robson–Garnier parameters with respect to each individual amino acid residue indicates that, if the β-sheet content is to be higher than that predicted, the most likely occurrence is a continuation of the β-pleated sheet between residues 20 and 44, residues 90 and 114 and residues 160 and 184 (Fig. 2). If this continuation of sheet were to occur the β-pleat and α-helix contents would be relatively closely in line with c.d. data. A turn function is present in the predicted β-pleat.

The Fc portion of IgG binds to Protein A on two antiparallel α-helices (Deisenhofer, 1981) connected by several residues of

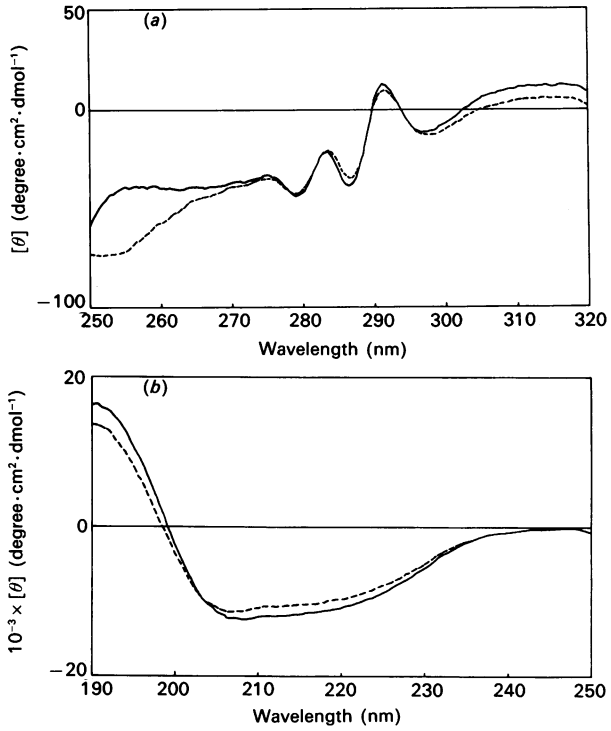

Fig. 5. C.d. spectra of Protein G′ after heating

(*a*) Near-u.v. and (*b*) far-u.v. c.d. spectra of Protein G′ after incubation in Milli-Q water, pH 5, at 20 °C (——) and 100 °C (-----) for 20 min. The protein concentrations were 2.675 mg/ml for the near-u.v. spectra and 0.322 mg/ml for the far-u.v. spectra. Protein G′ solutions were sealed in Pyrex tubes and heated for 20 min in a water bath. After heating to 100 °C, the sample was rapidly cooled in ice-cold water.

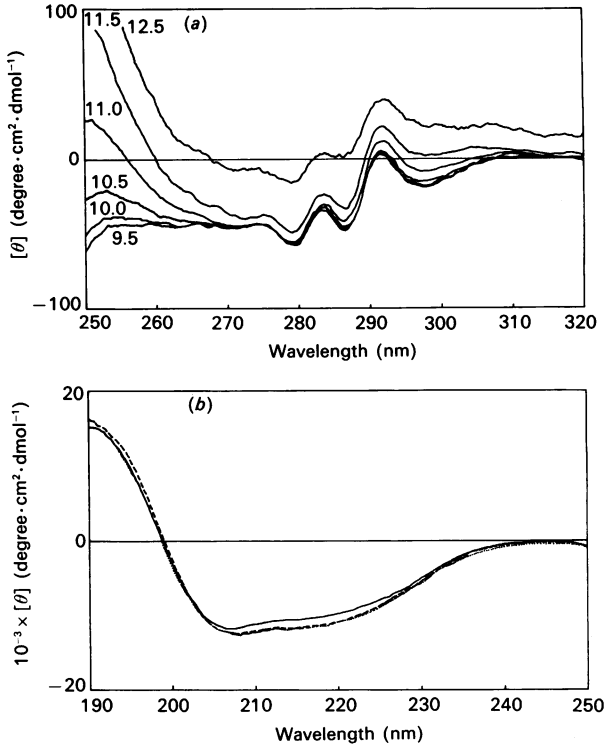

Fig. 6. C.d. spectra of Protein G′ at various pH values

(*a*) Near-u.v. and (*b*) far-u.v. c.d. spectra of Protein G′ in Milli-Q water at various pH values. The protein concentration was 2.675 mg/ml for the near-u.v. spectra and 0.332 mg/ml for the far-u.v. spectra. The pH of samples was adjusted with HCl or NaOH. The pH values at which spectra were measured in the near u.v. are shown on the Figure. In the far u.v. the spectra shown were measured at pH 1.5 (——), 7 (-----) and 12.5 (·····).

irregular conformation (Deisenhofer *et al.*, 1978) containing a turn function. Our prediction for Protein G′, however, would be an α-helix–β-sheet/turn–α-helix motif covering residues 6–58 in one domain (C′ and D), and similarly in the second domain covering residues 76–128 (C2 and D). The third domain (C3) appears to be incomplete. This is in contrast with the report of Guss *et al.* (1986), who conjecture that IgG binding is limited to the C regions and does not involve the D domain. It is noteworthy that Protein A and Protein G compete for the same binding region of the Fc portion of IgG (Eliasson *et al.*, 1989), indeed they both bind the $C_\gamma2$–$C_\gamma3$ domain interface region of IgG (Stone *et al.*, 1989). The lack of amino acid sequence similarity, however, does not preclude three-dimensional structural similarities; binding to IgG may require a similar shape even if contacts are made by different residues. It is relevant to note that interaction of the second α-helix of the B fragment of Protein A with Fc may be a crystal contact only (Deisenhofer, 1981) and so the binding could be just between the first α-helix and Fc.

C.d.

The chiral nature of the environments near aromatic amino acid residues of proteins often produces Cotton effects in near-u.v. c.d. spectra, which are sensitive indicators of structural change. Protein G′ contains six phenylalanine, eight tyrosine and three tryptophan residues (Fig. 1) and its near-u.v. c.d. spectrum at pH 7 showed three distinct negative Cotton effects, with minima at 279, 286 and 298 nm (Fig. 3*a*), which probably arise from highly immobilized tryptophan residues (Strickland, 1974). These Cotton effects appeared to be superimposed on a broad negative band with a minimum between 260 and 270 nm. No fine structure from phenylalanine residues at 250–270 nm, which are prominent in the c.d. spectra of Protein A (Sjöholm, 1975), could be observed in the spectrum of Protein G′ (Fig. 3*a*).

The far-u.v. c.d. spectrum of native Protein G′ was typical for a protein containing only a moderate amount of α-helix (Fig. 3*b* and Table 1). The change in structure of the protein with increasing concentration of Gdn,HCl was monitored by measuring changes in both mean residue ellipticity at 222 nm and fluorescence emission intensity contributed by tryptophan residues. These results (Fig. 4) showed that there was no change in structure up to about 2 M-Gdn,HCl but that at higher concentrations there was a progressive loss of structure. The unfolding of the secondary structure as monitored by change in mean residue ellipticity is paralleled by a change in tertiary structure as shown by fluorescence emission intensity. However, changes in fluorescence emission intensity occur before changes in mean residue ellipticity indicating a change in the tryptophan environment before unfolding of the secondary structure. A concentration of 3 M-Gdn,HCl was required to unfold the protein by 50%. Complete loss of structure appeared to occur in 5.5 M-Gdn,HCl from near-u.v. (Fig. 3*a*) and far-u.v. (Fig. 3*b*) c.d. spectra. When refolding was initiated by replacing 5.5 M-Gdn,HCl with 10 mM-potassium phosphate buffer by rapid buffer-exchange chromatography on Sephadex G-25, the c.d. spectra in the near u.v. and far u.v. showed the refolding to be complete (results not shown) and the values of α-helix and β-sheet were similar to those of the control (Table 1). E.l.i.s.a. results showed that full IgG-binding capability was restored after removal of Gdn,HCl.

When Protein G′ was heated at temperatures up to 90 °C and cooled to room temperature, no changes were found in either near-u.v. or far-u.v. c.d. spectra (results not shown). After heating at 100 °C, there was a small change in the near-u.v. c.d. spectrum between 250 and 270 nm (Fig. 5*a*), the far-u.v. c.d. spectrum showed a possible minor loss of α-helix (Fig. 5*b* and Table 1) and e.l.i.s.a. showed the protein retained its full IgG-binding capacity. Cooling the sample slowly offered no advantage over rapid cooling. The stability of Protein G′ to heat is achieved without disulphide bonds to aid stabilization; indeed the molecule contains no cysteine residues, which also eliminates any possibility of interchain disulphide linkages.

The near-u.v. c.d. spectrum of Protein G′ varied little between pH 1.5 and 10, but at more alkaline pH values a positive peak with a maximum at about 248 nm appeared and the Cotton effects associated with tryptophan residues became slightly more positive (Fig. 6*a*). The peak at 248 nm corresponds to ionized tyrosine residues (Chen & Woody, 1971), and above pH 12 the changes in c.d. spectra associated with tyrosinization were found to be time-dependent. The near-u.v. c.d. spectrum at pH 11.5 remained unchanged after 24 h, and so the changes in c.d. spectra found at pH 11 and pH 10.5 (Fig. 6*a*) probably represent immediate changes in structure. When Protein G′ was left at pH 12 for 6 days and the pH was then adjusted to 7, the near-u.v. c.d. spectrum was similar to that of the control sample at pH 7 but with less-well-defined Cotton effects associated with the tryptophan residues (results not shown). There was, however, a marked loss of α-helix (Table 1), and e.l.i.s.a. showed that the IgG-binding capacity was not restored. The far-u.v. c.d. spectra were virtually unchanged as the pH was varied (Fig. 6*b*), and the secondary structure remained intact (Table 1). E.l.i.s.a. showed that the protein retained 100% of its IgG-binding capability after exposure to pH values between 1.5 and 11; above this value there was a progressive decrease in IgG binding.

Sjöholm (1975) showed that Protein A is also very stable. Tyrosinization was observed above pH 8, but below this value the near-u.v. spectra were similar. Above pH 9 there was considerable loss of negative ellipticity. The tertiary structure of Protein A appears less stable at high pH than that of Protein G′. Protein A was also less stable than Protein G′ at high temperature; after heating at 80 °C for 3.5 h considerable ellipticity is lost in the near-u.v. spectrum.

In summary, Protein G′ has predicted secondary structure values of 37 ± 4% α-helix and 30 ± 5% β-sheet with calculated secondary-structure values of 29 ± 2% α-helix and 41 ± 3% β-sheet. Binding to IgG is conjectured to be through an α-helix–β-sheet/turn–α-helix motif. The molecule is very stable over a wide range of pH values, is stable to boiling and requires a high concentration (> 4 M) of Gdn,HCl to unfold the molecule, the effect of which is reversible. This stability is of great importance in the purification and maintenance of this useful immunochemical and potential therapeutic reagent. A marked structural stability of native Protein G on the outer surface of the streptococcal cell (and similarly for Protein A on the staphylococcal cell) may be critically important for the molecule to continue to function in its natural environment.

We thank Professor Gordon Roberts for providing us with data before publication and for critical comments, Roy Hartwell for amino acid analyses and Richard Seabrook for measurements of fluorescence emission intensity.

REFERENCES

Åkerström, B., Brodin, T., Reis, K. & Björck, L. (1985) J. Immunol. **135**, 2589–2592
Bayley, P. M., Janot, J.-M. & Martin, S. R. (1989) FEBS Lett. **250**, 389–394
Björck, L. & Kronvall, G. (1984) J. Immunol. **133**, 969–974
Chen, A. K. & Woody, R. W. (1971) J. Am. Chem. Soc. **93**, 29–37
Chou, P. Y. & Fasman, G. D. (1978) Adv. Enzymol. Relat. Areas Mol. Biol. **47**, 45–148
Deisenhofer, J. (1981) Biochemistry **20**, 2361–2370
Deisenhofer, J., Jones, T. A., Huber, R., Sjödahl, J. & Sjöquist, J. (1978) Hoppe-Seyler's Z. Physiol. Chem. **359**, 975–985
Eliasson, M., Andersson, R., Olsson, A., Wigzell, H. & Uhlén, M. (1989) J. Immunol. **142**, 575–581
Erntell, M., Myhre, E. B., Sjöbring, U. & Björck, L. (1988) Mol. Immunol. **25**, 121–126
Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) J. Mol. Biol. **120**, 97–120
Goward, C. R., Murphy, J. P., Atkinson, T. & Barstow, D. A. (1990) Biochem. J. **267**, 171–177
Guss, B., Eliasson, M., Olsson, A., Uhlén, M., Frej, A.-K., Jornvall, H., Flock, J.-I. & Lindberg, M. (1986) EMBO J. **5**, 1567–1575
Kabsch, W. & Sander, C. (1983) FEBS Lett. **155**, 179–182
Lindmark, R. (1982) Mol. Immunol. **19**, 957–959
Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. **193**, 265–275
Nishikawa, K. (1983) Biochim. Biophys. Acta **748**, 285–299
Nozaki, Y. (1972) Methods Enzymol. **26**, 43–50
Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M. & Uhlén, M. (1987) Eur. J. Biochem. **168**, 319–324
Provencher, S. W. & Glöckner, J. (1981) Biochemistry **20**, 33–37
Reis, K. J., Hansson, H. F. & Björck, L. (1986) Mol. Immunol. **23**, 425–431
Sjödahl, J. (1977) Eur. J. Biochem. **73**, 343–351
Sjöholm, I. (1975) Eur. J. Biochem. **51**, 55–61
Stone, G. C., Sjöbring, U., Björck, L., Sjöquist, J., Barber, C. V. & Nardella, F. A. (1989) J. Immunol. **143**, 565–570
Strickland, E. H. (1974) CRC Crit. Rev. Biochem. **2**, 113–175
Takakuwa, T., Konno, T. & Meguro, H. (1985) Anal. Sci. **1**, 215–218

Received 2 July 1990/28 July 1990; accepted 7 September 1990