Expression and purification of a truncated recombinant streptococcal Protein G

Christopher R. GOWARD,* Jonathan P. MURPHY, Tony ATKINSON and David A. BARSTOW
Division of Biotechnology, PHLS Centre for Applied Microbiology and Research, Porton Down, Salisbury SP4 0JG, Wilts., U.K.

INTRODUCTION

Proteins which bind to the constant (Fc) region of IgG are located on the surface of a variety of staphylococci and streptococci (Langone, 1982). Protein A from Staphylococcus aureus is the best known of these Fc receptor molecules, and its ability to bind IgG has been exploited in many immunochemical methods (Goding, 1978; Langone, 1982). Protein G is a bacterial cell-surface-associated protein of group C and G streptococci and binds the IgG of different subclasses of most mammalian species (Åkerström et al., 1985; Guss et al., 1986; Reis et al., 1986).

In contrast with Protein A from Staph. aureus, Protein G from Streptococcus strain G148 binds to all four subclasses of human IgG; Protein A does not bind to the IgG subclass (Guss et al., 1986). However, Protein A has a higher overall affinity for human polyclonal IgG than does Protein G (Eliasson et al., 1989), and the two proteins have complementary binding patterns (Guss et al., 1986; Eliasson et al., 1988). Protein G has been shown to interact with the Fab regions of IgG, but with a 10-fold lower affinity than determined for the Fc region (Björck & Kronvall, 1984) and has been reported to bind F(ab')2 fragments of IgG (Erntell et al., 1988). There are independent and separate binding regions for Fab and Fc fragments of IgG on the Protein G molecule, and the elongated structure of the molecule may permit simultaneous binding of both Fab and Fc (Erntell et al., 1988).

The complete nucleotide sequence of the Protein G structural gene from Streptococcus G148 cloned by Olsson et al. (1987) indicates an Mₚ of 63294 and a pre-protein of 593 amino acids (including the N-terminal signal peptide). Protein and gene sequence analysis show that the Protein G gene has similar features to those of Protein A (Uhlén et al., 1984). Both proteins consist of repetitively arranged domains (Sjödahl, 1977; Fahnstock et al., 1986; Guss et al., 1986). Whereas Protein G from the streptococcal strains GX7809 and G148 consists of two and three IgG-binding domains respectively (Fahnstock et al., 1986; Olsson et al., 1987), Protein A from a range of different Staph. aureus strains has five IgG-binding domains (Uhlén et al., 1984; Guss et al., 1985). Protein G (from strain G148) also binds albumin at several sites which are structurally separate from the IgG-binding sites (Åkerström et al., 1987; Björck et al., 1987; Nygren et al., 1988; Sjöbring et al., 1988). Sjöbring et al. (1989) found Protein G from 31 strains of human group C and G streptococci had both IgG- and albumin-binding ability, which suggests both binding regions are essential to the mode of action of these bacteria. Sjöbring et al. (1988) isolated several IgG- and albumin-binding proteins, including a Protein G fragment of Mₚ 14000 which binds only albumin. A Protein G fragment of Mₚ 7500 binds IgG (Guss et al., 1986), but the exact residues involved in binding IgG have not yet been identified. Despite the similarities of function between Protein A and Protein G, with the exception of a short sequence at the C-terminus of the proteins probably associated with membrane anchorage, the genes and amino acid sequences show no homology.

Heterogeneous Protein G has previously been prepared by chromatography on IgG-Sepharose (Guss et al., 1986; Eliasson et al., 1988), by chromatography on DEAE-Sephadex, IgG-Sepharose and then Sephadex G-100 (Åkerström & Björck, 1986), and by chromatography on DEAE-cellulose, Sephadex G-100 and finally IgG-Sepharose (Björck & Kronvall, 1984). Falkenberg et al. (1988) prepared Protein G by h.p.l.c. (affinity), whereas Björck et al. (1987) isolated three major IgG-binding proteins from the Escherichia coli-cloned Streptococcus G148 Protein G gene by chromatography on IgG-Sepharose and then Sephadex G-200. The major protein of Mₚ 65000 was separated and described as 96% homogeneous.

The aim of the present study was to clone the Protein G gene in order to be able to express it in a non-pathogenic host without having to use proteolytic enzymes for release of Protein G from the streptococcal cell wall, which appears to result in degradation products (Goward & Barstow, 1989). The map of various regions on the gene coding for different functions (Guss et al., 1986; Åkerström et al., 1987; Sjöbring et al., 1988) was used to design a Protein G molecule that would bind only the Fc portion of IgG (and neither the Fab region nor albumin) and which would have the cell-wall-spanning and membrane-anchoring regions removed to diminish both non-specific binding of IgG and potential problems with expression (Shuttleworth et al., 1987).

Abbreviations used: PBS, phosphate-buffered saline (composition and pH given in the text); HSA, human serum albumin; SPG, Streptococcus G148 Protein G.

* To whom correspondence and reprint requests should be addressed.
We further describe a simple method to prepare homogeneous Protein G' free from fragments of the major protein molecule.

**EXPERIMENTAL**

**Materials**

Radiochemicals were from Amersham International. X-Omat SX-ray film was from Kodak. Deoxy- and dideoxy-nucleotides, DNA ligase, restriction endonucleases and other DNA-modifying enzymes were from Boehringer. Agarose, acrylamide, bisacrylamide and phenol were from Bethesda Research Laboratories. Chromatography media were from Pharmacia—LKB (Uppsala, Sweden). Immunoglobulins were from Sigma or ICN Biomedicals Ltd., High Wycombe, Bucks., U.K. All other reagents were from Sigma or BDH. Nitrocellulose was purchased from Anderman and Co., Kingston-upon-Thames, Surrey, U.K.

**Media and culture conditions**

_E. coli_ was cultured in 2xYT broth [2% (w/v) tryptone/1% (w/v) yeast extract/1% (w/v) NaCl] at 37°C. Media were solidified with 2% (w/v) Bacto-agar (Difco). HT-agar for M13 overlays contained 1% (w/v) tryptone, 0.8% (w/v) NaCl and 0.8% (w/v) Bacto-agar (Difco). Ampicillin (25–50 μg/ml) was used where necessary for the selection and growth of transformants. Functional β-galactosidase was detected by addition of 5-bromo-4-chloroindolyl β-D-galactoside to a final concentration of 600 μg/ml and, where necessary, isopropyl β-D-thiogalactopyranoside to a final concentration of 200 μg/ml.

_E. coli_ containing the recombinant Protein G' gene were grown in a 400-litre fermentation on a medium containing yeast extract, Casamino acids and glycerol plus trace elements. The pH was maintained at 7.0 with H₃PO₄ or NaOH and the temperature at 37°C. After 8 h of growth the temperature in the vessel was rapidly reduced to below 10°C and the bacteria were harvested by centrifugation in a Westfalia KA25 centrifuge at a flow rate of 250 litres/h. The cell paste was quick-frozen and stored at −20°C.

**Isolation of DNA**

_Streptococcus_ G148 genomic DNA was isolated essentially as described by Guss et al. (1986). Plasmids were purified from _E. coli_ by Brij lysis (Clewell & Helinski, 1969) and CsCl/ethidium bromide density-gradient centrifugation (Radloff et al., 1967). A rapid small-scale plasmid-isolation technique (Holmes & Quigley, 1981) was used for screening procedures.

**Genetic manipulation procedures**

DNA-modifying enzymes were used in the buffers and under the conditions recommended by the supplier (Boehringer). Transformation of _E. coli_ was essentially as described by Cohen et al. (1972). Electrophoresis of DNA fragments was performed on vertical 1% (w/v)-agarose slabs gels in Tris/acetate buffer (40 mM-Tris/20 mM-sodium acetate/2 mM-EDTA, adjusted to pH 7.9 with acetic acid). DNA fragment sizes were estimated by comparison with fragments of λ-phage DNA digested with the restriction endonuclease HindIII. DNA fragments were purified by electrophoresion essentially as described by McDonnell et al. (1977). Southern transfers and hybridization conditions were performed by previously described procedures (Southern, 1975). Site-directed mutagenesis was performed by the methods of Carter et al. (1985).

**Nucleotide sequencing**

Nucleotide sequences were determined by the chain-termination procedure of Sanger et al. (1980) on M13 templates generated by the sonication procedure of Deininger (1983). The oligonucleotide sequencing primers were synthesized by using an Applied Biosystems 380B DNA synthesizer. All sequences of the coding and non-coding strands were confirmed with adequate overlap between contiguous sequences. Sequence data were assembled into contiguous sequence by using the computer programs of DNASTAR (Madison, WI, U.S.A.).

**Large-scale purification**

Disruption of cells. A 400-litre fermentation yielded 18 kg of cell paste. A 2 kg portion of cell paste was thawed at 4°C in 4 litres of 50 mM-Hepes/NaOH buffer, pH 8.0, containing 250 mM-NaCl and 0.5 mg of DNAase/litre. The thawed suspension was disrupted with a 15M-8BA Manton-Gaulin homogenizer at 550 kg/cm². The homogenate was centrifuged for 45 min at 13000 g at 4°C and passed through a 0.45 μm-pore-size filter to remove fine particulate matter.

Affinity chromatography on IgG-Sepharose 4B. IgG-Sepharose 4B was prepared by coupling porcine IgG to CNBr-activated Sepharose 4B (5 mg of IgG/ml of matrix). The cell-extract supernatant was applied to a 1-litre IgG—Sepharose 4B column ([17.5 cm long × 9 cm internal diameter (i.d.]) operated at a linear flow rate of 30 cm/h and equilibrated with 50 mM-Hepes/NaOH buffer, pH 8.0, containing 250 mM-NaCl. Unbound protein was removed with equilibration buffer and Protein G' was eluted in a single peak with 100 mM-glycine/HCl, pH 2.0. The eluate was immediately made 20 mm with respect to Tris and the pH was adjusted to 7.5 with NaOH. Protein G' was concentrated with an Amicon CH2A ultrafiltration unit fitted with an H10P10 hollow-fibre cartridge.

Anion-exchange chromatography on Q-Sepharose FF. The concentrated IgG—Sepharose 4B eluate was further purified in portions on a 32 ml column (16 cm × 1.6 cm i.d.) of Q-Sepharose FF operated at a linear flow rate of 60 cm/h and equilibrated with 20 mM-Tris/HCl, pH 7.5. The column was washed with equilibration buffer and the protein was eluted with a linear gradient of 0–500 mM-NaCl in 20 mM-Tris/HCl, pH 7.5.

Anion-exchange chromatography on Mono Q. The major fractions purified as above were separated by f.p.i.c. (Pharmacia LKB, Sweden) on an 8 ml Mono Q HR 10/10 column. The column was operated at a linear flow rate of 230 cm/h and equilibrated with 20 mM-Tris/HCl, pH 7.5. A 500 μl portion of concentrated IgG—Sepharose 4B eluate was applied. The column was washed with equilibration buffer and protein was eluted with a linear gradient of 0–250 mM-NaCl in 20 mM-Tris/HCl, pH 7.5.

**Protein assay**

Protein concentrations were determined by the Folin method of Lowry et al. (1951), with bovine serum albumin as the standard. The protein content of column eluates was also monitored by absorbance at 280 nm.

**Protein G assay**

Protein G concentration was determined functionally by an e.l.i.s.a. procedure. It was assumed that uni-, bi- or poly-valent Protein G fragments are detected, since a fragment containing one IgG-binding domain can be separated by chromatography on IgG—Sepharose (Guss et al., 1986). The method was adapted from that described for Protein A (Warnes et al., 1986). Microtitre plates were coated with capture antibody (human polyclonal IgG) at 2 μg/ml in 15 mM-NaCO₃/35 mM-NaHCO₃, pH 9.6 (100 μl/well) at room temperature overnight. The plates were washed six times after each stage of the assay with phosphate-buffered saline (PBS): 8 mM-NaHPO₄/1.5 mM-KH₂PO₄/137 mM-
Truncated recombinant streptococcal Protein G

NaCl/2.7 mM-KCl, pH 7.4, containing 0.1% (v/v) Tween 20 ('polyoxyethylene sorbitan monolaurate'). This solution was also used as a diluent in all stages of the procedure. Protein G standards, prepared by chromatography on IgG-Sepharose 4B followed by anion-exchange FPLC on Mono Q HR, and other samples, were diluted as appropriate. Samples (100 µl) of 2-fold dilution series of each sample were transferred to the microtitre plate and incubated for 90 min at room temperature. Bound Protein G was detected by incubation with goat anti-rabbit IgG IgG–horseradish peroxidase conjugate for 90 min at 20 °C. The bound conjugate was incubated for 30 min at room temperature with 0.1% (w/v) 5-aminosalicylic acid/0.006% (v/v) H2O2 in 50 mM-Na2HPO4/NaH2PO4, pH 6.0, and the absorbance of the coloured product was determined at 450 nm by using a Titertek Multiscan MCC automatic plate reader. The Protein G concentration was determined by comparison of the samples with Protein G standard. The human serum albumin (HSA) binding capacity of the Protein G was determined by substitution of HSA for the capture antibody.

PAGE

Acrylamide (12.5%, w/v) slab gels were run in an LKB vertical electrophoresis unit (Laemmli, 1970). Proteins were stained with Coomassie Brilliant Blue R-350, and protein bands were scanned with a Chromoscan-3 laser optical densitometer (Joyce–Loebl, Gateshead, Tyne and Wear, U.K.), to estimate the apparent Mr.

Western blotting

Proteins were applied to nitrocellulose membranes by electrophoretic transfer from SDS/polyacrylamide gels as described by Towbin et al. (1979). The nitrocellulose membranes were incubated with 125I-labelled human polyclonal IgG or HSA to detect the Protein G. Autoradiography was performed at −70 °C with Kodak X-Omat S X-ray film.

Determination of pl

The pl of Protein G was determined with a Pharmacia PhastGel apparatus and broad-pH-range gels (pH 3.5–9.5), followed by narrow-pH-range gels (pH 4.0–6.5) as described by the manufacturer (Pharmacia–LKB). The appropriate Pharmacia–LKB calibration protein kits were used, and the pl was estimated from densitometer scans of the protein bands.

N-Terminal sequencing

N-Terminal sequence analysis was performed to locate the proteins on the gene sequence. Sequences were determined on an Applied Biosystems 470A protein sequencer by automated Edman phenylthiohydantoin degradation (Hunkapiller et al., 1983). Protein G samples were dialysed against 50 mM-NaCl, and about 500 pmol was applied to the gel-phase sequencer. The equipment was operated essentially according to the manufacturer’s instructions. Repetitive Edman degradations provided sequential removal of amino acids from the peptide, which were identified by using reversed-phase h.p.l.c. (Hunkapiller & Hood, 1983).

RESULTS AND DISCUSSION

Cloning and characterization of the Protein G gene from Streptococcus G148

The strategy to clone the Streptococcus G148 Protein G (SPG) gene was to use the previously determined incomplete DNA sequence (Guss et al., 1986) to design specific oligonucleotide probes for detection of the gene. Thus two synthetic oligonucleotides were constructed; SPG1 (sequence 5'-GGTAAAA-CATTGAAAGGCGAA) specific for the C repeat regions (Guss et al., 1986) and SPG2 (sequence 5'-AAATATGGAGT-AAGTAGCTAT) specific for the A repeat regions, each being represented three times in the Streptococcus G148 sequence (Fig. 1a).

Streptococcus GX7805 was shown to have a 2.3 kb HindIII fragment containing the SPG gene (Fahnstock et al., 1986; Filopula et al., 1987). As the nucleotide sequence of GX7805 was shown to be identical with that of the SPG gene (Guss et al., 1986; Filopula et al., 1987; Olsson et al., 1987) we attempted to clone the entire SPG gene as a single HindIII fragment. HindIII-digested fragments of genomic DNA isolated from Streptococcus G148 were subjected to Southern-blot analysis (Southern, 1975) and probed independently with 32P-labelled oligonucleotides SPG1 and SPG2. However, both probes hybridized to a 4.2 kb fragment. Streptococcus G148 genomic DNA was digested with HindIII, separated by electrophoresis on a 1% (w/v)-agarose gel, and DNA fragments of 4.0–4.4 kb were excised and purified by electroelution. The purified genomic DNA fragments were ligated to HindIII dephosphorylated plasmid vector pUC8 and transformed into E. coli TG2.

In all 1000 colonies were probed by colony hybridization in situ (Grunstein & Hognes, 1975) using the SPG1 and SPG2 oligonucleotides, and nine positives were detected. Plasmid DNA was isolated from four by CsCl density-gradient centrifugation (plasmids pSPG2, 3, 5 and 6) and characterized by restriction-endoenzyme analysis. Clones pSPG3, 5 and 6 showed identical restriction-endoenzyme cleavage patterns after digestion with PstI, EcoRI, DraI and HindIII. Single and double restriction enzyme digests were performed on pSPG3 and a 'crude' restriction endonuclease cleavage map was constructed (Fig. 1b), which correlated well with the restriction-endoenzyme cleavage maps of previously cloned Protein G genes (Guss et al., 1986; Fahnstock, 1987). The SPG gene in the plasmid construct pSPG3 was found to be in the opposite orientation to the lac promoter of the vector pUC8; thus expression of Protein G from this clone was from its own promoter. The clone produced a protein which bound human polyclonal IgG and HSA, as confirmed by Western-blot analysis and e.l.i.s.a.

Sequencing of clone pSPG3

The 4.2 kb HindIII fragment from pSPG3 (Fig. 1c, i) was purified by gel electroelution; M13 templates isolated from sonicated fragments of this circularized HindIII fragment were DNA-sequenced. Three changes in the sequence were found from that previously published (Guss et al., 1986; Olsson et al., 1987), as shown in Fig. 2. All were located in the 5' non-coding region of the gene; two of these were direct changes, whereas the third was an insertion.

Site-directed mutagenesis of the Protein G gene

Attempts to clone the entire 4.2 kb HindIII fragment from pSPG3 directly into M13 repeatedly failed, as did attempts to subclone the large blunt-ended 2.1 kb DraI–HindIII fragment directly into M13. However, it proved possible to clone the blunt-ended DraI–HindIII fragment into SmaI-cleaved pM13L22. The Protein G gene was then excised on a BamHI–BglI fragment, ligated into BamHI dephosphorylated M13mp8 and transformed into E. coli TG2. Templates were made from 12 plaques, and the DNA was sequenced by using oligonucleotides spanning the entire gene as primers. Two templates (9 and 13) were found to contain the Protein G gene, but in different orientations from the lac promoter of M13.

Deletion of the cell-wall and membrane-spanning regions from the translated protein was achieved by synthesis of an oligo-
nucleotide (5'-CCTCTGTAACCTTATTCAGTT) in which, by site-directed mutagenesis, the ATG codon between nucleotide positions 1259 and 1261 was replaced with the stop codon TAA (Fig. 2). To prevent read-through, a guanine residue was inserted at nucleotide position 1262 after the TAA stop codon to alter the reading frame of the gene and introduce further stop codons downstream (Fig. 2). Mutants were confirmed by DNA sequencing. The mutated gene was designated pMSPG594-12 (Fig. 1c, ii). The albumin-binding domains were deleted from pMSPG594-12 by looping out the relevant regions with an oligonucleotide (5'-TAAAATTTCACTATGAAGAAATTCTTAAAG) that hybridized to the sequence 15 bp before the A1 domain and 15 bp after the A3 domain. Removal of the albumin-binding domains was confirmed by DNA sequencing of the final mutant, pMSPG631 (Protein G'), using primers from the mutation (Fig. 1c, iii).

**Construction of a production strain expressing the recombinant Protein G’**

The recombinant Protein G' gene was removed from pMSPG631 on a KpnI–HindIII fragment and ligated into KpnI–HindIIICleaved dephosphorylated pMTL22. After transformation into *E. coli* TG2, 250 recombinant clones were screened for the Protein G’ gene by colony hybridization *in situ*, and a positive colony (pSPG29) was cultured to purify and quantify any translated product.

Cell-free extracts from 100 ml of culture were passed through a 1 ml column (1.6 cm x 0.9 cm i.d.) of IgG-Sepharose 4B equilibrated and washed with 50 mM-Hepes/NaOH, containing 250 mM-NaCl. Protein G’ was eluted with 100 mM-glycine/HCl, pH 2.0; it showed no binding to HSA and gave only one predominant protein, of apparent *M*ₐ 35,000, which bound to human polyclonal IgG.

**Nucleotide-sequence analysis of pSPG29**

The original aim of this work was to isolate a derivative of the Protein G gene and to express a protein that lacked the portion containing albumin-binding domains A1, B1, A2, B2 and A3 and terminated after the C3 domain (Fig. 1). However, N-terminal protein sequence analysis (Fig. 2) showed translation was initiated several residues into one of the C domains. To clarify this anomaly, the nucleotide sequence of clone pSPG29 was determined. The BamHI–BglII fragment from pSPG29 containing the Protein G’ was purified by gel electroelution; M13 templates isolated from sonicated fragments of this circularized fragment were sequenced (Fig. 2). Sequencing confirmed the two mutations introduced by site-directed mutagenesis and the three differences already discussed that existed between our sequence and the sequences previously published (Guss et al., 1986; Olsson et al., 1987). However, a major additional and unsought change was the deletion of a single guanine residue from the original sequence between nucleotide positions 564 and 565 (Fig. 2), which changed the reading frame of the gene. The deletion was found to have probably occurred during construction of pMSPG594-12. As a consequence, the open reading frame of the gene was not preserved, and stop codons were introduced immediately downstream of this deletion (Fig. 2). This accounts for the absence of full-length Protein G. The TTG codon at nucleotide positions 705–707 was found (by N-terminal sequence analysis of pure protein) to have acted as an alternative initiation codon within the gene, giving rise to a functional IgG-binding protein (Protein G').
The sequence was compared with that of Olsson et al. (1987). The C’ region corresponds to Cl, but translation of the protein was initiated part of the way into this domain. The boxed regions show the stop codons. Differences between this sequence and that described by Olsson et al. (1987) are: 1, substitution of G for C; 2, insertion of C; 3, substitution of T for A; 4, deletion of G and a corresponding shift in the reading frame; 5, stop codon introduced due to the shift in reading frame; 6, HSA-binding domains deleted; 7, translation initiated again; 8, stop codon introduced; 9, insertion of G to prevent any possible continuation of translation by introduction of putative stop codons, 10 and 11. Differences 1, 2 and 3 are also found in the native Protein G gene. The N-terminal amino acid sequence of the purified protein is indicated by a continuous line.
Guss et al. (1986) reported translation of two proteins from an EcoRI–HindIII fragment, both of which had the same N-terminal sequence as the Protein G' we have prepared; translation was thought to be initiated in the system using the TTG codon at nucleotide positions 705–707 and 915–917 (Fig. 2). Guss et al. (1986) also postulated a poor ribosomal binding sequence of GGT complementary to the 16S rRNA of prokaryotes upstream from the TTG start codon. However, the −10 and −35 promoter sequences recognized by E. coli that Guss et al. (1986) suggested, must be ruled out in our gene, since they lie within the A and B repeat regions which have been removed by site-directed mutagenesis.

Purification

The protein was heterogeneous after affinity chromatography, possibly due to proteolytic ‘nicking’ of the molecule. Anion-exchange chromatography on Q-Sepharose FF removed some of the smaller contaminating Protein G' fragments, but anion-exchange chromatography on Mono Q HR (Fig. 3) was used to prepare Protein G' with an overall recovery of 45%, which gave a single protein band after SDS/PAGE and isoelectric focusing (Figs. 4 and 5). Results in Table 1 show Protein G' was expressed at 0.4% total soluble cell protein and in different cultures results have ranged between 0.3 and 0.6%.

SDS/PAGE

SDS/PAGE showed the IgG-Sepharose 4B eluate contained a heterogeneous mixture of protein (Fig. 4), but Western-blot analysis with 125I-human polyclonal IgG showed the minor bands to bind IgG and therefore they may be post-translational products. The apparent Mr of 35000 is in conflict with the predicted Mr of about 20000. Other authors observed that recombinant IgG-binding proteins had lower mobilities on SDS/PAGE than the full amino acid sequence would predict (Guss et al., 1986; Nygren et al., 1988). This anomalous behaviour may be attributable to an excessively elongated structure of the molecule in SDS, low SDS binding to the protein or a C-terminal post-translational modification (limited proteolytic cleavage) of the protein removing disproportionate capacity for SDS binding.

Specificity of binding

The Protein G' was shown to bind Fc fragments, but not Fab fragments, of human polyclonal IgG by e.l.i.s.a. using Fc or Fab as the capture molecule. The autoradiograph of a Western blot probed with 125I-HSA showed no evidence of HSA binding; this was further confirmed by e.l.i.s.a. using HSA as the capture molecule when again no HSA binding was observed.

\[
\text{Pl}
\]

The pl of Protein G' was shown to be 4.19 by using narrow-range (pH 4.0–6.5) isoelectric-focusing gels (Fig. 5), and the theoretical pl was calculated as 4.20 by the computer program of DNASTAR Inc. The pl of Protein G cleaved from the cell walls of Streptococcus G148 with papain was determined to be less than 3.5 (Åkerström & Björcr, 1986).

N-Terminal sequence

The first 35 N-terminal amino acid residues of the Protein G' molecule were sequenced and are indicated by the continuous line in Fig. 2. The sequence starts with an N-terminal protein-sequence-identified methionine residue from the TTG codon. TTG is a common start codon in Gram-positive bacteria (Uhlen et al., 1983), but is uncommon in E. coli. However, Protein A is well expressed from a TTG initiation codon in E. coli (Shuttleworth et al., 1987). Translation was initiated part-way in to one
Truncated recombinant streptococcal Protein G

Table 1. Purification of Protein G

<table>
<thead>
<tr>
<th>Step</th>
<th>Volume (ml)</th>
<th>Total protein (mg)</th>
<th>Protein G’ (μg)</th>
<th>Specific amount (μg/mg of total protein)</th>
<th>Recovery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgG-Sepharose 4B</td>
<td>4500</td>
<td>163 000</td>
<td>655 600</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>Cell-free extract</td>
<td>1100</td>
<td>560</td>
<td>471 000</td>
<td>841</td>
<td>72</td>
</tr>
<tr>
<td>Q-Sepharose FF</td>
<td>45</td>
<td>189</td>
<td>159 000</td>
<td>841</td>
<td>100</td>
</tr>
<tr>
<td>IgG–Sepharose 4B eluate*</td>
<td>41</td>
<td>102</td>
<td>92 000</td>
<td>902</td>
<td>58</td>
</tr>
<tr>
<td>Mono Q HR</td>
<td>0.5</td>
<td>6.72</td>
<td>5650</td>
<td>841</td>
<td>100</td>
</tr>
<tr>
<td>IgG–Sepharose 4B eluate*</td>
<td>8.0</td>
<td>3.52</td>
<td>3560</td>
<td>1011</td>
<td>63</td>
</tr>
</tbody>
</table>

* The IgG-Sepharose 4B eluate was concentrated before application to Q-Sepharose FF and Mono Q HR.

of the IgG-binding domains; however, Guss et al. (1986) have shown that such translated molecules still bind IgG efficiently. Initiation of translation is likely to have occurred at the first TTG codon, in view of the Mr of the protein expressed by E. coli, so the Protein G’ probably has three functional IgG-binding domains (Fig. 1c, iii). Although the most likely initiation position of Protein G’ is indicated in Fig. 2, giving a protein with three IgG-binding domains and an Mr of 20 000 (apparent Mr 35 000 by SDS/PAGE), it is conceivable that a molecule with the same N-terminal sequence, with two IgG-binding domains, the same theoretical pi and an Mr of 12 500 could be produced by initiation of translation from a TTG codon starting at nucleotide position 915 instead of 705. The N-terminal protein sequence and present data are not adequate to distinguish between these possibilities other than to note that 2-fold differences between real and apparent Mr on SDS/polyacrylamide gels are known, whereas 3-fold differences are not. Protein G with three IgG-binding domains has been demonstrated to have greater affinity for human polyclonal IgG than Protein G with two IgG-binding regions (Eliasson et al., 1989). Guss et al. (1986) reported that N-terminal sequence analyses of the two gene products of their Protein G gene suggest two different TTG codons may be recognized in E. coli for initiation of translation to yield two proteins. However, the same TTG codon may have been recognized for initiation of translation, and the smaller molecule may have been generated by C-terminal proteolytic enzyme cleavage. In contrast, we found only one major product, even though there are three TTG codons available for initiation of translation.

We thank Dr. S. P. Chamber for culture of bacteria and Mr. D. R. Hartwell for oligonucleotide synthesis and N-terminal sequence analysis.

REFERENCES


Received 30 June 1989/23 October 1989; accepted 1 November 1989

Vol. 267