Expression of the G glycoprotein gene of human respiratory syncytial virus in *Salmonella typhimurium*

Antonia Martin-Gailardo, Edward Fleischer, Shawn A. Doyle, Rasappa Arumugham, Peter L. Collins, Stephen W. Hildreth and Peter R. Paradiso

1 Department of Virology, Praxis Biologics Inc., 300 East River Road, Rochester, New York 14623 and 2 Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892, U.S.A.

The attachment protein, G, of human respiratory syncytial virus (RSV) is an M_r 84K to 90K species which has a high content of N-linked and O-linked carbohydrates. The unglycosylated form of this protein was expressed by inserting a full-length cDNA copy of the mRNA from the A2 strain of RSV into a prokaryotic expression vector under the control of the lambda P_L promoter. *Salmonella typhimurium* cells transformed with the G-containing plasmid synthesized a protein of M_r 40000 that specifically reacted with polyclonal and two neutralizing monoclonal antibodies raised against the native RSV G glycoprotein. Recombinant G protein was purified by immunoaffinity chromatography using a neutralizing monoclonal antibody. Cotton rats immunized with the recombinant G protein produced serum antibodies to the G glycoprotein that neutralized RSV *in vitro*. The study demonstrates that the G protein of RSV can be expressed in bacteria and that at least one neutralizing epitope is not structurally dependent on carbohydrates.

Introduction

Human respiratory syncytial virus (RSV), a member of the genus Pneumovirus within the Paramyxoviridae family (McIntosh & Chanock, 1990), is the major cause of lower respiratory disease in infants (Chanock et al., 1982). It has a negative-sense ssRNA genome that encodes 10 proteins including two envelope glycoproteins, G and F, that mediate attachment and penetration respectively, and have been identified as the major viral antigens involved in inducing RSV-neutralizing antibodies and protective immunity (Huang et al., 1985; Levine et al., 1987; Olmsted et al., 1986).

The RSV G protein has structural features that make it unique among the known viral surface glycoproteins (Satake et al., 1985; Wertz et al., 1985). It has a predicted sequence of 298 amino acids (for strain A2) with a calculated M_c of 32587, but the mature protein migrates on PAGE with an M_r of 84000 to 92000 (Levine et al., 1987; Wertz et al., 1985). This is due to extensive glycosylation, most of which consists of O-linked oligosaccharide chains (Satake et al., 1985; Wertz et al., 1985). Serine and threonine residues, which provide the sites for O-glycosylation, constitute 30-6% of the protein. In addition, four potential sites for attachment of N-linked oligosaccharides have been identified for the A2 strain and appear to be glycosylated (Satake et al., 1985; Wertz et al., 1985).

The apparent high carbohydrate content of the G glycoprotein raises questions concerning the role of oligosaccharides in the immune response to this protein. Studies of serum samples taken from children who have had several RSV infections showed that the IgG subclass immune response to the heavily glycosylated G glycoprotein differed from that shown to a polysaccharide antigen, being more like the immune response to a protein antigen (Wagner et al., 1989). However, the contribution the carbohydrate moiety makes to the immunogenicity and antigenic variation of the G protein remains unknown. Expression of the G protein in a bacterial system would offer one approach for studying this aspect, since mammalian proteins synthesized in bacteria are unglycosylated (Bialy, 1987; Ellis & Gerety, 1990). In this study, we have expressed the gene encoding the G protein from the A2 strain of RSV (subgroup A) in *Salmonella typhimurium* to analyse the antigenicity of the unglycosylated G protein and to assess the ability of the unglycosylated recombinant G protein to induce a neutralizing antibody response in cotton rats.

Methods

Cells and viruses. RSV strains were plaque-purified and propagated in HEp-2 cell monolayers; strains A2, Long, 137 and G-2 are subgroup A viruses and strains 18537, 9320 and 43 are subgroup B viruses.
Bacterial strains and plasmids. Escherichia coli strains used were the lambda lysogen N96cT and N4830 (Pharmacia). N96cT carries the wild-type cT repressor of the lambda Pp promoter and was used for the construction and growth of Pp-containing plasmids. N4830 carries the temperature-sensitive cB37 lambda repressor and was used for thermoinducible protein expression (Gottesman et al., 1981). S. typhimurium strain LB5010 is a galE derivative of strain LB5000, provided by L. B. Bullas, via B. A. D. Stocker (Stanford University, Stanford, Ca., U.S.A.) (Bullas & Ryu, 1983). The full-length cDNA of the G glycoprotein mRNA of the A2 strain of RSV was present as a BamHI fragment in the replicative form of M13mp18 phage (Elango et al., 1986). The G cDNA was cloned into pBR322 and the resulting plasmid was termed pPX1003. Plasmid pPL-lambda (Pharmacia) carried the Pp promoter and the N gene on a 1215 base pair (bp) segment of the bacteriophage lambda genome (Drahos & Szybalski, 1981; Reed, 1981), and contained a unique Hpal site that is located within the N gene 321 bp downstream from the start of Pp transcription. The expression vector pPX1600 was provided by W. Majarian (Praxis Biologics Inc., Rochester, N.Y., U.S.A.) and contained a synthetic 60 bp oligonucleotide inserted into the Hpal site of pPL. This oligonucleotide included a translation termination codon in frame with the N coding sequences, followed by translation initiation signals and restriction sites (NcoI, StuI, EcoRV) for the insertion of heterologous DNA sequences (see Fig. 1).

Protein analysis. E. coli cells (N4830 strain) harbouring the expression plasmid with and without the G cDNA insert (pPX1044 and pPX1600, respectively) were grown to mid-log phase at 30 °C in L broth containing 100 mg/ml ampicillin, and then shifted at 42 °C for either 30 min or 1 h. S. typhimurium cells harbouring the same plasmids were grown at 37 °C. Cells were harvested by centrifugation, resuspended in 1 mM-PMSF, 1 mM-EDTA, 25 mM-HEPES pH 7.4, and lysed using a French press. Proteins from the bacterial lysates were separated by SDS-PAGE (Laemmli, 1970) under both reducing and non-reducing conditions and electrotransferred to nitrocellulose papers (Towbin et al., 1979) which were incubated with polyclonal and monoclonal antibodies (MAbs) raised against the RSV G glycoprotein. For Western blots of cotton rat sera, blots received an additional incubation with anti-cotton rat IgG rabbit sera. The antigen-antibody complexes were detected using horseradish peroxidase (HRP)-labelled anti-rabbit or anti-mouse IgG and 4-chloro-1-naphthol. Anti-G MAbs (L7 and L9) were provided by E. E. Walsh (University of Rochester, Rochester, N.Y., U.S.A.), and have been previously described (Walsh & Hruska, 1983).

Protein purification. The G glycoprotein was purified from RSV (strain A2)-infected HEp-2 cells by immunoaffinity chromatography using MAb L9, as described by Walsh et al. (1984). The recombinant G protein was purified from pPX1044-containing S. typhimurium cells that were grown to late-log phase at 37 °C. A lysate of the bacteria was centrifuged at 100000 g for 1 h, and the pellet from this centrifugation, which was enriched with expressed G (Gsα), was resuspended in 25 mM-HEPES pH 7.4, containing 1% Triton X-100, 2% deoxycholate and 2 mM-PMSF. The soluble material was isolated by ultracentrifugation and passed through a Sepharose-L9 immunoaffinity column, then eluted using buffer (0.1 m-glycine pH 2.5, with 0.1% Triton X-100) as described for native G glycoprotein (Walsh et al., 1984).

Serological response to immunization. Cotton rats were immunized intramuscularly with 1 μg of Gsα, G native (glycoprotein purified from HEp-2 cells) or with PBS (control) on weeks 0, 2, 4 and 6. The immunogens were adsorbed to aluminium hydroxide at 0.5 mg/ml and rats received a final volume of 100 μl/dose. Antibody responses were measured by ELISA using purified native G and F glycoproteins and Gsα as antigens, and by 60% plaque reduction neutralization assay (Walsh & Hruska, 1983).

Results

Construction of bacterial expression plasmids

The entire coding sequence of the RSV G protein was inserted into the expression vector pPX1600 (Fig. 1), which carries the Pp promoter and the N gene from the bacteriophage lambda (Drahos & Szybalski, 1981). Additional features of pPX1600 include an in-frame

![Fig. 1. Construction of recombinant plasmids. Plasmid pPL-lambda (Pharmacia), which contained the Pp promoter (■) and the N gene (□) from bacteriophage lambda, was modified by inserting a synthetic oligonucleotide (ON; stippled box) into the Hpal site. This oligonucleotide included a translation stop codon for the N coding sequences, followed by a ribosomal binding site, translation initiation codon (underlined sequences) and three unique restriction sites for the insertion of a foreign gene. The new recombinant vector was termed pPX1600. Plasmid pPX1003 contained a full-length cDNA copy of the RSV G gene (■) cloned into the BamHI site of pBR322. The RSV G cDNA was excised from pPX1003 by BamHI digestion, filled-in with Klenow enzyme, and ligated to Stul-cleaved pPX1600. The resulting expression plasmid, pPX1044, should direct the expression of the RSV G gene by reading from the synthetic ATG. The direction of transcription is indicated by arrows. * indicates the DNA sequence at the junction between pPX1600 and the 5' end of the G cDNA; amino acids are designated by the single-letter code; ● represents the initiation codon for the authentic G protein.](image-url)
RSV G gene expression in S. typhimurium

Translation stop codon placed in the unique HpaI site of the N coding sequences followed by a ribosomal binding site, a translational start site, and three unique sites (NcoI, StuI, EcoRV) for insertion of a foreign gene; translation termination signals in all reading frames are located downstream of these cloning sites. The G cDNA was excised from pPX1003, filled in, and blunt end-ligated to the StuI site of pPX1600 to make pPX1044 (Fig. 1). The DNA sequences at the junctions between the plasmid vector and both 5’ and 3’ ends of the G cDNA were confirmed by the dideoxynucleotide chain-termination method of Sanger et al. (1977). In this construction the initiation codon for the G protein is in frame with the ATG included in the oligonucleotide sequences, which lies 9 bp downstream from the synthetic ribosomal binding site. Therefore, the recombinant G protein is predicted to have six additional amino acids at its NH₂ terminus with respect to the authentic G protein.

Expression of recombinant RSV G protein in bacteria

To analyse P₄-directed expression of the RSV G protein, the plasmid pPX1044 and the vector pPX1600 without the G cDNA insert were introduced into E. coli cells (strain NB830) carrying the temperature-sensitive cI857 repressor of the P₄ promoter. Heat induction failed to cause production of detectable amounts of recombinant G protein, as determined by Western blots probed with rabbit anti-RSV G serum and developed with HRP-labelled anti-rabbit IgG (not shown). However, expression was demonstrated in S. typhimurium cells (strain LB5010), which were tested in parallel. Differences between the levels of gene expression in E. coli and S. typhimurium have been previously observed for the RSV F protein (A. Martin-Gallardo, unpublished results). Specifically, a protein band of about 40K was detected in extracts of S. typhimurium cells carrying pPX1044 (Fig. 2, lane 2). This protein was not detected in extracts of induced cells harbouring pPX1600 (Fig. 2, lane 1), or when the blots were probed with rabbit preimmune serum (data not shown). In addition, the 40K protein reacted with anti-G MAb L7 and L9 (Fig. 2, lanes 4 and 6, respectively). Both of these MAbs neutralize RSV and have been shown to bind to a common subtype neutralizing epitope (Walsh et al., 1989). Other smaller species (Mᵣ < 30000) reacted with the anti-G MAb L9 and might represent degradation products of the 40K protein.

Purification and immunogenicity of recombinant G

In order to study the immunogenic properties of the unglycosylated G protein, the recombinant G product synthesized in S. typhimurium cells (Gₛₐ₁) was partially purified by immunoaffinity chromatography. The eluted fraction contained both anti-G MAb L9-reactive and non-reactive proteins (Fig. 3). Using both silver-stained gels and Western blots, the eluted protein fraction was estimated to contain about 20% MAb L9-reactive protein with the major band at 40K.
Fig. 3. Purification of the protein of RSV expressed in S. typhimurium. Cell extracts (lanes 1 and 5) were applied to a G protein-specific immunoaffinity column and the non-binding fraction (lanes 2 and 6) was examined. After extensive washing, proteins were eluted (lanes 3 and 7) with elution buffer. The SDS-polyacrylamide gels were run under non-reducing conditions; lanes 1 to 4 were developed by silver staining and lanes 5 to 8 by Western blotting using anti-G MAbs (L9). Lanes 4 and 8 contain Mr markers.

Cotton rats immunized with $G_{\text{sal}}$ developed moderately high levels of serum antibodies that bound to the authentic A2 G glycoprotein purified from RSV-infected cells (Fig. 4a, lane 3). This pattern of reactivity to G protein was similar to that seen using antibody to the $G_{\text{native}}$ protein (Fig. 4b, lane 3). The antibody to $G_{\text{sal}}$ did not react with the F glycoprotein (Fig. 4a, lane 2) whereas the antibody to $G_{\text{native}}$ did, due to co-purification of a small amount of F protein (5%) in the $G_{\text{native}}$ preparation (unpublished data). This pattern of reactivity was confirmed by ELISA using purified F and G proteins (Table 1). The anti-G protein response was around eight-fold higher in cotton rats immunized with the $G_{\text{native}}$. The relative level of neutralizing antibody response was similar but the data are complicated by the potential contribution of anti-F antibody in the animals vaccinated with $G_{\text{native}}$. Sera from the animals inoculated with $G_{\text{sal}}$ displayed a strong subgroup-specific neutralizing response, neutralizing all of four subgroup A viruses but only one of three subgroup B viruses.

**Discussion**

In this study we demonstrated the synthesis of the RSV G protein by S. typhimurium cells containing a recombinant expression plasmid. This plasmid was designed to

Table 1. Immunogenicity of $G_{\text{sal}}$ recombinant protein in cotton rats

<table>
<thead>
<tr>
<th>Immunization*</th>
<th>ELISA† (× 100)</th>
<th>Neutralization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G_{\text{native}}$</td>
<td>$F_{\text{native}}$</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>Long</td>
</tr>
<tr>
<td>PBS</td>
<td>&lt; 1.5</td>
<td>&lt; 1.5</td>
</tr>
<tr>
<td>$G_{\text{sal}}$</td>
<td>6.5</td>
<td>&lt; 1.5</td>
</tr>
<tr>
<td>$G_{\text{native}}$</td>
<td>52.5</td>
<td>55.0</td>
</tr>
</tbody>
</table>

* Rats were immunized intramuscularly with 1 µg of G/dose on weeks 0, 2, 4, 6 and bled for serological examination on week 8. The total protein dose for $G_{\text{sal}}$ was 5 µg with 20% estimated to be L9 MAb-reactive G protein. Purified $G_{\text{native}}$ from HEp-2 infected cells contains copurified F protein (approximately 5% of total protein).
† Mean serum antibody titre (reciprocal, × 100) using purified F and G glycoproteins from RSV A2 or $G_{\text{sal}}$ recombinant protein.
‡ Mean serum antibody titre measured by 60% plaque reduction neutralization of RSV.
§ ND, Not determined.
direct the expression of the 298 amino acid open reading frame of the G protein fused to a recombinant DNA sequence that coded for only six non-G amino acids. The immunologically reactive G product (G~sal) migrated with an \( M_r \) of about 40,000, which is somewhat higher than that deduced from the amino acid sequence (33,400). However, unglycosylated G protein synthesized in vitro using cell-free translation of mRNA also showed a retarded electrophoretic mobility with respect to its calculated \( M_r \) (Satake et al., 1985). Therefore, the gel mobility of G~sal would be consistent with that of the complete unglycosylated G protein.

The recombinant G~sal protein reacted with two MAbs (L7 and L9) raised to the native G glycoprotein (Walsh & Hruska, 1983). These MAbs recognize a single epitope which is cross-reactive between the two subgroups of RSV, as demonstrated by competitive-binding studies (Walsh et al., 1989). Furthermore, the L7 and L9 MAbs neutralized both the Long (subgroup A) and 18537 (subgroup B) strains of RSV, and provided significant lung protection against live RSV challenge upon passive transfer to cotton rats (Walsh et al., 1989). Palomo et al. (1991) have observed variability in binding of mouse MAbs to unglycosylated and partially glycosylated forms of mammalian cell-derived G proteins. They observed that the majority of the MAbs would bind only to mature or partially glycosylated forms, whereas only a few antibodies would bind to both the unglycosylated and the glycosylated forms of G protein. Of the latter group, all have been reported to have some ability to neutralize RSV (Garcia-Barreno et al., 1989). Both MAbs used to identify the G~sal protein have been observed to bind to the 33K to 38K G from HEp-2-grown Long and 18537 strains of RSV (Walsh et al., 1989 and personal communication), and therefore were considered good candidates for detecting neutralizing epitopes in a G protein expressed by bacteria. The fact that the antigenic site defined by these MAbs is conserved in the unglycosylated G~sal product suggest that this epitope is determined primarily by the amino acid sequence of the G glycoprotein and is not dependent on the attachment of carbohydrates. It also suggested that the bacterially expressed protein contained an authentic neutralization epitope.

Consistent with this latter point, cotton rats immunized with purified G~sal protein developed an antibody response to the G glycoprotein that neutralized RSV infectivity in vitro. The titres of antibodies, measured both by ELISA and neutralization, were approximately eight-fold lower than those produced in response to the \( G_{native} \) from RSV-infected HEp-2 cells. However, the observed greater potency of the latter for inducing higher and broader neutralizing titres may be biased because of the presence of a minor amount of F in the immuno-

affinity-purified native polypeptide. Therefore, we can conclude that at least one neutralization epitope is defined by the polypeptide moiety of the G protein, but it is difficult to quantify the difference in the relative immunogenicities of the glycosylated and unglycosylated forms. It remains to be determined whether the lack of carbohydrates and possible differences in conformation are important to G~sal antigenicity. Future studies will determine whether larger doses of G~sal would boost the neutralization titres to levels considered to be protective (Prince et al., 1985).

The antigenic differences between the G glycoproteins of the RSV subgroups (Garcia-Barreno et al., 1989; Johnson et al., 1987; Walsh et al., 1987) was reflected in the serum neutralizing antibody responses to the recombinant G~sal. The levels of these antibodies against three of the four homologous subgroup A viruses tested were significantly greater than those observed for the heterologous subgroup B viruses. The subgroup specificity of the neutralizing response in animals immunized with G~sal agrees with that described in cotton rats immunized with a vaccinia virus A2-G recombinant (Johnson et al., 1987). Antigenic variation of the G protein has been recently found within viruses of the same subgroup with only limited divergence of amino acid sequence between strains (Garcia-Barreno et al., 1989). Our results would be consistent with this observation and suggest that the polypeptide moiety alone is important in determining at least one neutralization epitope of the heavily glycosylated G protein of RSV.

The challenge to develop an RSV vaccine is considerable. The bacterial recombinant G protein described here could be a potential candidate for an RSV vaccine subunit. Furthermore, secretory IgA antibodies play a leading role in the control of the infectious process at its onset (Leschinsky et al., 1988). The synthesis of recombinant G in S. typhimurium cells makes it possible to examine the ability of a live salmonella oral vaccine to elicit both secretory and cellular immunity against RSV.

We thank Dr William Matijarjan for the plasmid pPX1600, Dr Bruce Stocker for S. typhimurium strains, Dr Caroline Hall for RSV strains, Dr Edward Walsh for L7 and L9 MAbs, and Dan Speelman and Pamela Yeh for technical support.

References


(Received 5 November 1991; Accepted 3 November 1992)