Vaccine xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Structural, antigenic and immunogenic features of respiratory syncytial virus glycoproteins relevant for vaccine development

José A. Melero [a,*], Vicente Mas [a], Jason S. McLellan [b,*]

[a] Unidad de Biología Viral, Centro Nacional de Microbiología and CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain
[b] Department of Biochemistry and Cell Biology, Geisel School of Medicine at Dartmouth, Hanover, NH, United States

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
RSV glycoproteins and vaccine development

ABSTRACT

Extraordinary progress in the structure and immunobiology of the human respiratory syncytial virus glycoproteins has been accomplished during the last few years. Determination of the fusion (F) glycoprotein structure folded in either the prefusion or the postfusion conformation was an inspiring breakthrough not only to understand the structural changes associated with the membrane fusion process but additionally to appreciate the antigenic intricacies of the F molecule. Furthermore, these developments have opened new avenues for structure-based designs of promising hRSV vaccine candidates. Finally, recent advances in our knowledge of the attachment (G) glycoprotein and its interaction with cell-surface receptors have revitalized interest in this molecule as a vaccine, as well as its role in hRSV immunobiology.

Human respiratory syncytial virus (hRSV) was recently classified in the genus *Orthopneumovirus* of the newly created *Pneumoviridae* family within the order *Mononegavirales*, detached from the original *Paramyxoviridae* family [1]. hRSV is an enveloped virus with a genome made of a single-stranded RNA molecule of negative polarity and about 15.2 kb in length. Molecules of nucleoprotein (N) wrap around the entire length of this RNA to form a stable ribonucleoprotein (RNP) complex, which encodes 11 proteins, three of which are membrane-bound glycoproteins (G, F and SH) (for a review, [2]). The G glycoprotein was originally described as the receptor-binding or attachment protein [3]. F was identified by Walsh and Hruska [4] as the fusion protein that fuses the viral and cell membranes enabling the virus RNP to reach the cell cytoplasm. Finally, SH was initially described as a viroporin—a class of small viral proteins that modify membrane permeability [5]—and was later found to form pentameric pore-like structures in the membrane that confer cation-selective channel-like activity, compatible with its initial designation as a viroporin [6].

It is widely accepted that protection against hRSV is conferred mainly by neutralizing antibodies. For instance, high levels of neutralizing antibodies correlate with protection of human adult volunteers to hRSV challenge [7], as well as a lower risk of hRSV infection in children [8] and in the elderly [9]. Therefore, the surface glycoproteins, particularly F, have recently received much attention as targets of neutralizing and protective antibodies and as potential antigens to be included in a hRSV vaccine [10]. These aspects of hRSV vaccinology are the topic of this review.

Structure and function of the hRSV G glycoprotein: The G protein is synthesized as a polypeptide precursor of about 300 amino acids (depending on the viral strain) with a single hydrophobic domain (residues 38–63) near the N-terminus that acts as a combined signal and membrane anchor domain [11] (Fig. 1). This hydrophobic region targets the nascent chain, as it emerges from the ribosome, to the endoplasmic reticulum and ensures translocation of the polypeptide chain across the membrane while anchoring the G protein to the lipid bilayer. G has neither sequence nor structural homology with the attachment protein of viruses in the *Paramyxoviridae* family [11].

The G polypeptide precursor is extensively modified by the addition of both *N*- and *O*-linked oligosaccharides and is also palmitoylated at a single cysteine residue in its N-terminal cytoplasmic tail [12]. High-mannose *N*-linked glycans are co-translationally added to the G protein precursor, followed by the conversion of these sugars to the complex type and addition of *O*-linked glycans in the Golgi compartment. These modifications convert the 32 kDa precursor into a mature protein of 80–90 kDa (estimated by SDS-PAGE) in most immortalized cell lines [13], whereas a 180 kDa form has been described in human airway epithelial (HAE) cultures [14] that is postulated to represent either a dimer of the 90 kDa G protein or the 90 kDa form with additional or more extensive *O*-linked carbohydrate chains.

* Corresponding authors.
E-mail addresses: jmelero@isciii.es (J.A. Melero), Jason.S.McLellan@Dartmouth.edu (J.S. McLellan).

http://dx.doi.org/10.1016/j.vaccine.2016.09.045
0264-410X/

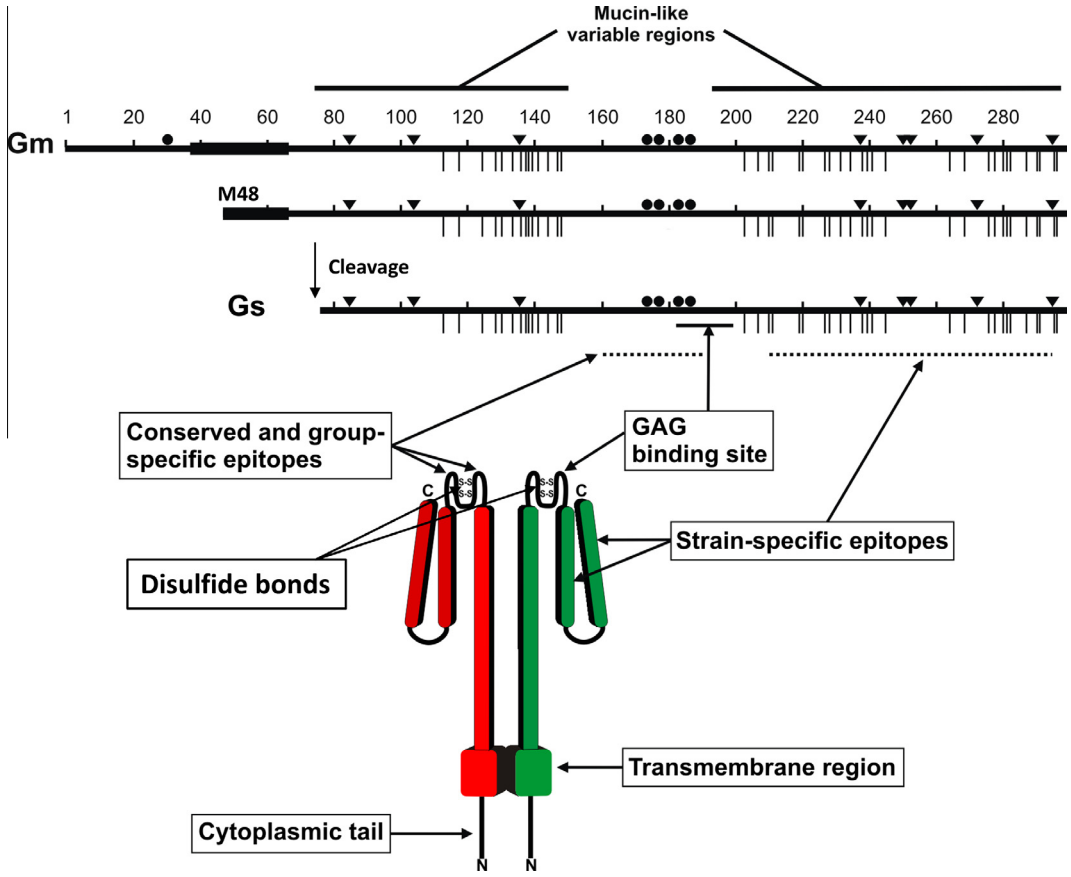

Fig. 1. Human respiratory syncytial virus G glycoprotein. The full length, 298 amino acid membrane-anchored G protein (Gm) and the 233 amino acid soluble G protein (Gs) are shown (Long strain). Hydrophobic regions are denoted by thick lines. Gs is formed by alternative translation initiation at M48, followed by cleavage after residue 65. Inverted triangles represent *N*-linked glycosylation sites and vertical lines indicate *O*-linked glycosylation sites. Cysteine residues overlapping the central conserved domain are represented by solid circles. The lower part of the figure depicts a model of the 3-dimensional structure of Gm. While Gm is probably tetrameric [38], a dimer is shown for simplicity. The mucin-like regions are depicted as extended rod-like structures due to the presence of multiple *O*-linked sugars that have a tendency to stretch the polypeptide backbone [93]. The second hypervariable region is externally located in the model to denote that it harbors multiple epitopes and it is shown as two halves joined by a protease susceptible site [94]. Antibody epitopes and the glycosaminoglycan (GAG) binding site are indicated by arrows. Figure provided by Alfonsina Trento.

The G protein ectodomain consists of two large heavily glycosylated "mucin-like" domains, rich in serine, threonine and proline residues (characteristic of mucins), connected by a short central region devoid of carbohydrates (Fig. 1) [15]. The sequence of the two mucin-like domains is extremely variable among viral strains [16] but they all have several potential sites for *N*-glycosylation and multiple serines and threonines that are predicted to be *O*-glycosylated by the NetOGlyc software [17]. This sequence variability has been used in numerous studies of molecular epidemiology and evolution of hRSV [18]. Thus, hRSV strains have been classified into two genetic groups, A and B, that correlate with the antigenic groups, initially identified by reactivity with certain monoclonal antibodies (mAbs) [19,20]. Within each group numerous clades or genotypes have been identified. Viruses of different genotypes and even different antigenic groups frequently co-circulate in each yearly outbreak. The dominance of these genotypes changes in successive epidemics and replacement of certain genotypes by others has been noticed at the global level [21].

The central conserved region of hRSV G (aa 163–189) has four cysteines (residues 173, 176, 182 and 186) that are conserved in all viral strains. Within this region there is a stretch of 13 amino acids (164–176) that is strictly maintained in all strains while the remaining sequence of the central region is somewhat group-specific. Disulfide bridges are formed between Cys173 and Cys186, and between Cys176 and Cys182, resulting in a cystine noose motif which resembles the structure found in the 55 kDa tumour necrosis factor receptor [22,23]. The Cys-rich motif is missing in the highly related G protein of human metapneumovirus (hMPV), which otherwise shares the overall amino acid composition and sequence variability of hRSV G [24,25]. The ectodomains of both hRSV and hMPV G are predicted to be disordered (except for the hRSV cystine noose), consistent with the high content of serine, threonine, and proline residues and extensive *O*-glycosylation [26].

The central region of hRSV G contains the CX3C motif (aa 182–186) that can bind to CX3CR1—the specific receptor of the fractalkine chemokine—and hence induce leukocyte chemotaxis [27]. Several authors have reported hRSV binding to differentiated HAE cells by the interaction of the G protein with CX3CR1 in the apical surface of ciliated cells [28–30]. Inhibition of CX3CR1 binding reduces but does not entirely suppress infection of HAE cultures, indicating that CX3CR1is an important but not the only hRSV receptor in these cells. It has been reported that hRSV uses cell surface proteoglycans for attachment to established cell lines [31–33] mainly by interactions of the G protein with glycosaminoglycans (GAGs) [34,35]. Whether proteoglycans may also act as an hRSV G receptor in HAE cells is still controversial [28,30].

In addition to the membrane bound form of hRSV G, infected cells also produce a soluble form of G (sG) [36] by internal initiation of translation at a second AUG codon (Met48) located in the middle of the transmembrane region and subsequent cleavage after residue Asn66 (Fig. 1) [37]. While sG is monomeric,

Please cite this article in press as: Melero JA et al. Structural, antigenic and immunogenic features of respiratory syncytial virus glycoproteins relevant for vaccine development. Vaccine (2016), http://dx.doi.org/10.1016/j.vaccine.2016.09.045