Human metapneumovirus (HMPV) is a ubiquitous pathogen of the Pneumovirinae subfamily of the Paramyxoviridae and causes serious respiratory illness in infants, young children, the elderly, and immunocompromised individuals (1–4). HMPV has a negative-sense, nonsegmented, single-stranded RNA genome of approximately 13 kb that encodes 9 proteins. Three of these are membrane-anchored glycoproteins, namely, the attachment (G), the small hydrophobic (SH), and the fusion (F) proteins. The F protein mediates fusion of the viral and cellular membranes during viral entry, induces syncytium formation in infected cells, and determines the cellular host range (5–8). While F is highly conserved, is immunogenic, and induces protective antibodies (9–11), the other surface glycoproteins, G and SH, have been shown to be only weakly immunogenic (11–14).
The G protein has been associated with binding to cellular glycosaminoglycans (15); however, this function was shown to be strain dependent (16). G-deleted recombinant HMPV is attenuated and induces high titers of HMPV-neutralizing serum antibodies, which confer protection against wild-type HMPV (14). Immunological studies have suggested that G inhibits host cell innate immune responses by targeting RIG-I-dependent gene transcription (17) and Toll-like receptor 4 (TLR-4)-dependent signaling (18, 19), although some of these findings have been mitigated (20, 21). Recently, deletion of the G and SH genes was shown to reduce HMPV internalization by monocyte-derived dendritic cells, leading to decreased activation of CD4 T cells (22).

In order to investigate the structural properties of HMPV G, we employed in silico predictions to compare conservation, disorder propensity, and localization of glycosylation sites in HMPV and the closely related human respiratory syncytial virus (HRSV) G (Fig. 1). Both glycoproteins possess a short conserved and structured N-terminal intracellular tail, followed by a transmembrane region (residues 31 to 53 in HMPV and residues 41 to 63 in HRSV) and an extracellular ectodomain (sG) of 182 and 236 amino acids in HMPV-A and in HRSV, respectively. The sG sequence is poorly conserved (23–26), with the exception of the cysteine-rich motif, present exclusively in HRSV (residues 160 to 190), which has been linked with immunomodulatory functions (27–29). Interestingly, the most variable regions are located near the C terminus and away from the transmembrane region (Fig. 1A and B), likely reflecting immune pressure. sG is predicted to be disordered, consistent with the large number of O-glycosylation sites (Fig. 1C and D) associated with a high content of serine, threonine, and proline residues (18.5, 21.7, and 10.9%, respectively).

Next, sG (residues 54 to 236) from strain NL-1-00 (A1) was cloned into pHLsec with an N-terminal secretion signal and a C-terminal His tag, transiently expressed in HEK293T cells, and purified from culture medium following standard procedures (30). Characterization using size exclusion chromatography combined with multiangle laser light scattering and refractometry (SEC-MALLS-RI) (31) indicates a molecular mass varying between 34 and 41 kDa (±1%). Comparison with the sequence-derived molecular mass (20.0 kDa) suggests that sG is a highly glycosylated monomer, with carbohydrates accounting for roughly 50% of the measured molecular mass, a property that may contribute to virion stability by preventing dehydration (32).

To gain insight into the solution structure of sG, we used small-angle X-ray scattering (SAXS) (Fig. 2B). Guinier plots were linear and unaffected by protein concentration (Fig. 2C), with a measured radius of gyration ($R_g$) of 5.5 nm. Kratky plots were linear at high scattering vector $Q$ ($Q = 4\pi \sin(\theta)/\lambda$) (Fig. 2D), indicating that sG behaves as a random polymer. The pair-distance distribution function $P(r)$ displayed a pronounced tail with a maximal intramolecular distance ($D_{max}$) of 20 nm (Fig. 2E), which is characteristic for elongated or disordered proteins (33).

We then used ensemble optimization to quantify sG flexibility. sG was modeled using an ensemble of bead models, as implemented in the program RANCH, and the data were fitted using GAJOE (34). Although these models are coarse and cannot reproduce the branched nature of the glycoprotein, they remain useful in extracting the distributions of $R_g$ and $D_{max}$ (Fig. 3A and B),
indicating that sG populates a broad ensemble of conformations with $R_g$ of 3 to 8 nm and $D_{\text{max}}$ ranging from 10 to >25 nm. SAXS profiles were well fitted (Fig. 3C), with the goodness of fit ($\chi^2_{\text{exp}}$) decreasing smoothly from 3.3 to 2.0 when the optimized ensemble size was varied between 1 and 50 models (Fig. 3D), consistent with high levels of intrinsic disorder (35).

The large dimensions of sG and the reported association of F and G at the surface of viral particles (36, 37) suggested that G may have a shielding function and prompted us to compare the size of the two proteins (Fig. 4). The trimeric fusion protein F, which can exist in both pre- and postfusion conformations, possesses a large extracellular region for which extensive structural information is available (38–42). The ability of sG to extend up to >25 nm from the membrane would allow the protein to tower above the smaller F trimers, a phenomenon that might be amplified through oligomerization mediated by the transmembrane region (43). Steric hindrance by G may additionally decrease F binding to neutralizing antibodies, such as DS7 (a Fab fragment has been solved in complex with a fragment of the HMPV F protein [42]), or host factors, such as the innate immune sensor TLR-4–myeloid differentiation factor 2 (MD-2) complex, which is activated through binding of HRSV F to MD-2 (44).

This “steric masking” hypothesis is supported by the hyper-variability of the C-terminal region of sG, the increased capture radius and binding rates associated with intrinsic disorder (45, 46), and the decreased binding affinity of soluble proteins to membrane-anchored substrates in the presence of crowding factors (47). In addition, because of the nonprotective and cross-protective nature of antibodies directed against G and F, respectively (10–14, 48), transient immunity leading to reinfection could be explained if G can reduce the immunological footprint associated with F. This is consistent with the increased CD4$^+$ T cell activation observed with HMPV lacking the G and SH proteins (22). Interestingly, an avian metapneumovirus isolate bearing a long (585-residue) G protein was found to replicate efficiently without signs of disease in domestic turkeys, suggesting decreased activation of the innate immune response (49–51).

The properties of pneumovirus G proteins, such as intrinsic
disorder, sequence hypervariability, and heavy O-glycosylation, contrast with the structured attachment glycoproteins in other paramyxoviruses (52, 53). Interestingly, the surface glycoprotein GP from Ebola virus is dominated by a mucin-like domain of 150 amino acids, which was shown to shield viral epitopes and impair immune recognition (54, 55). This suggests similarities between immune evasion strategies employed by pneumoviruses and filoviruses, whose evolutionary relationship was recently highlighted by structural comparison of the matrix and M2-1 transcriptional antiterminator proteins (56–58).

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under SILVER grant agreement no. 260644. The work was also supported by the Wellcome Trust Core award (090532/Z/09/Z). M.R. was supported by a SILVER grant agreement no. 260644. The work was also supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 282445–4512. J.V. was supported by a Wellcome Trust fellowship (099667/Z/12/Z).

We thank the European Synchrotron Radiation Facility for beamtime and the staff of beamline BM29 for assistance with data collection.

REFERENCES

1. Boivin G, Abed Y, Pelletier G, Ruel L, Moisan D, Cote S, Peret TC, Erdman DD, Anderson LJ. 2002. Virological features and clinical manifestations associated with human metapneumovirus: a new paramyxovirus responsible for acute respiratory-tract infections in all age groups. J. Infect. Dis. 186:1330–1334. http://dx.doi.org/10.1086/344319.
2. Boivin G, De Serres G, Cote S, Gilca R, Abed Y, Rochette L, Bergeron MG, Dery P. 2003. Human metapneumovirus infections in hospitalized children. Emerg. Infect. Dis. 9:634–640. http://dx.doi.org/10.3201 /eid0906.030017.
3. Osterhaus A, Fouchier R. 2003. Human metapneumovirus in the community. Lancet 361:890–891. http://dx.doi.org/10.1016/S0140-6736(03) 12785-7.
4. van den Hoogen BG, de Jong JC, Groen J, Kuiken T, de Groot R, Fouchier RA, Osterhaus AD. 2001. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. Nat. Med. 7:719–724. http://dx.doi.org/10.1038/89098.
5. Herfst S, Mas V, Ver LS, Wierda RJ, Osterhaus AD, Fouchier RA, Melero JA. 2008. Low-pH-induced membrane fusion mediated by human metapneumovirus F protein is a rare, strain-dependent phenomenon. J. Virol. 82:8891–8895. http://dx.doi.org/10.1128/JVI.00472-08.
6. Schowalter RM, Smith SE, Dutch RE. 2006. Characterization of human metapneumovirus F protein-promoted membrane fusion: critical roles for proteolytic processing and low pH. J. Virol. 80:10931–10941. http://dx.doi.org/10.1128/JVI.01287-06.
7. de Graaf M, Schrauwen EJ, Herfst S, van Amerongen G, Osterhaus AD, Fouchier RA. 2009. Fusion protein is the main determinant of metapneumovirus host tropism. J. Gen. Virol. 90:1408–1416. http://dx.doi.org/10.1099/vir.0.009688-0.
8. Schlender J, Zimmer G, Herrler G, Conzelmann KK. 2003. Respiratory syncytial virus (RSV) fusion protein subunit F2, not attachment protein G, determines the specificity of RSV infection. J. Virol. 77:4609–4616. http://dx.doi.org/10.1128/JVI.77.8.4609-4616.2003.
9. Tang RS, Mahmood K, Macphail M, Guzzetta JM, Haller AA, Liu H, Kaur J, Lawlor HA, Stillman EA, Schickli JH, Fouchier RA, Osterhaus AD, Spaete RR. 2005. A host-range restricted parainfluenza virus type 3 (PIV3) expressing the human metapneumovirus (hMPV) fusion protein elicits protective immunity in African green monkeys. Vaccine 23:1657–1667. http://dx.doi.org/10.1016/j.vaccine.2004.10.009.
10. Skiaopoulos MH, Biacchesi S, Buchholz UJ, Riggs JM, Surman SR, Amaro-Carbomot E, McAluliffe JM, Ellkins WP, St Claire M, Collins PL, Murphy BR. 2004. The two major human metapneumovirus genetic lineages are highly related antigenically, and the fusion (F) protein is a major contributor to this antigenic relatedness. J. Virol. 78:6927–6937. http://dx.doi.org/10.1128/JVI.78.9.6927-6937.2004.
11. Skiaopoulos MH, Biacchesi S, Buchholz UJ, Amaro-Carbomot E, Surman SR, Collins PL, Murphy BR. 2006. Individual contributions of the human metapneumovirus F, G, and SH surface glycoproteins to the induction of neutralizing antibodies and protective immunity. Virology 345:492–501. http://dx.doi.org/10.1016/j.virol.2005.10.016.
12. Tedcastle AB, Fenwick F, Robinson MJ, Toms GL. 2014. Immunogenicity in mice of human metapneumovirus with a truncated SH glycoprotein. J. Med. Virol. 86:547–557. http://dx.doi.org/10.1002/jmv.23731.
13. Ryder AB, Tollefson SJ, Podsaid AB, Johnson JE, Williams JV. 2010. Soluble recombinant human metapneumovirus G protein is immunogenic but not protective. Vaccine 28:4145–4152. http://dx.doi.org/10.1016/j.vaccine.2010.04.007.
14. Biacchesi S, Skiaopoulos MH, Yang L, Lamirande EW, Tran KC, Murphy BR, Collins PL, Buchholz UJ. 2004. Recombinant human metapneumovirus lacking the small hydrophobic SH and/or attachment G glycoprotein: deletion of G yields a promising vaccine candidate. J. Virol. 78:12877–12887. http://dx.doi.org/10.1128/JVI.78.23.12877-12887.2004.

FIG 4 Comparison of the molecular dimensions of the extracellular regions of the HMPV F and G proteins, illustrating the G protein induced steric hindrance potentially hampering host factor-F protein interactions. The extracellular region of the Toll-like receptor 4–myeloid differentiation factor 2 (TLR-4/MD-2) complex (PDB identifier 3FXI) is represented as an example of such host factor. Homology models of the F protein trimers in the pre- and postfusion states form protrusions of 13 and 17 nm, respectively. Models were generated in HOMER (63) and are represented as red and blue surfaces. The homodimeric TLR-4 is shown in cyan and MD-2 in magenta, and three superimposed low-resolution models of sG are shown in orange, wheat color, and yellow. The anti-HMPV DS7 Fab bound to the prefusion model of F is shown in green based on PDB identifier 4DAG.
28. van den Hoogen BG, van Boheemen S, de Rijck J, van Nieuwkoop S, Smith DJ, Laksono B, Gultyaev A, Osterhaus AD, Fouchier RA. 2014. Excessive production and extreme editing of human metapneumovirus defective interfering RNA is associated with type I interferon induction. J. Virol. 89:10155–10163. 

29. Le Nouen C, Hillyer P, Brock LG, Winter CC, Rabin RL, Collins PL, Polack FP, Irusta PM, Hoffman SJ, Schiatti MP, Melendi GA, Delgado PM, Skiazi E, Stocks AM, Shaw PD, Loo VG, Muchondo G, Sadlon T, Gordon D. 2011. Antigenic and genetic variability of human metapneumoviruses. Emerg. Infect. Dis. 17:90–97. http://dx.doi.org/10.3201/eid1701.101196.

30. Swanson KA, Settembre EC, Shaw CA, Dey AK, Rappuoli R, Mandl CW, Dormitzer PR, Carfi A. 2011. Structural basis for immunization with postfusion respiratory syncytial virus fusion glycoprotein (RSV F) to elicit high neutralizing antibody titers. Proc. Natl. Acad. Sci. U. S. A. 108:9619–9624. http://dx.doi.org/10.1073/pnas.1016310108.

31. Valencia AM, Yu Q, Zsak L. 2009. Structure and flexibility within human metapneumovirus glycoproteins interact to form a complex on the surface of infected cells. J. Mol. Biol. 393:1093–1104. http://dx.doi.org/10.1016/j.jmb.2009.08.018.

32. Shoemaker BA, Portman JJ, Wolynes PG. 2000. Speeding molecular recognition by using the folding funnel: the flying-casting mechanism. J. Mol. Biol. 301:1143–1159. http://dx.doi.org/10.1006/jmbi.2000.4132.

33. Bennett RS, LaRue R, Shaw D, Yu Q, Nagaraja KV, Halvorson DA, Njenga MK. 2005. A wild goose metapneumovirus containing a large attachment glycoprotein is avirulent but immunoprotective in domestic turkeys. J. Virol. 79:14834–14842. http://dx.doi.org/10.1128/JVI.79.23.14834–14842.2005.

34. Yu Q, Estève C, Song M, Kapczynski D, Zsak L. 2009. Generation and biological assessment of recombinant avian metapneumovirus subgroups C and D viruses containing different lengths of the G gene. Virus Res. 147:182–188. http://dx.doi.org/10.1016/j.virusres.2009.10.021.

35. Pelikan M, Hura GL, Hammel M. 2009. Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen. Physiol. Biophys. 28:174–189. http://dx.doi.org/10.1128/GB.02_09097.

36. Wyatt PJ. 1998. Submicrometer particle sizing by multilange light scattering following flocculation. J. Colloid Interface Sci. 197:9–20.

37. Verdugo P. 1991. Mucin exocytosis. Am. Rev. Respir. Dis. 144:S33–S37.

38. Receiver-Brechot V, Durand D. 2007. How random are intrinsically disordered proteins? A small angle scattering perspective. Curr. Prot. Pept. Sci. 13:55–57. http://dx.doi.org/10.1002/0471148688.101688.88.

39. Bennett RS, LaRue R, Shaw D, Yu Q, Nagaraja KV, Halvorson DA, Njenga MK. 2005. A wild goose metapneumovirus containing a large attachment glycoprotein is avirulent but immunoprotective in domestic turkeys. J. Virol. 79:14834–14842. http://dx.doi.org/10.1128/JVI.79.23.14834–14842.2005.
virus subtype C wild bird isolates in domestic turkeys. Virol. J. 10:38. http://dx.doi.org/10.1186/1743-422X-10-38.

52. Jardetzky TS, Lamb RA. 2014. Activation of paramyxovirus membrane fusion and virus entry. Curr. Opin. Virol. 5:24–33. http://dx.doi.org/10.1016/j.coiviro.2014.01.005.

53. Plattet P, Plemper RK. 2013. Envelope protein dynamics in paramyxovirus entry. mBio 4(4):e00413-13. http://dx.doi.org/10.1128/mBio.00413-13.

54. Francica JR, Varela-Rohena A, Medvec A, Plesa G, Riley JL, Bates P. 2010. Steric shielding of surface epitopes and impaired immune recognition induced by the Ebola virus glycoprotein. PLoS Pathog. 6:e1001098. http://dx.doi.org/10.1371/journal.ppat.1001098.

55. Reynard O, Borowiak M, Volchkova VA, Delpeut S, Mateo M, Volchkov VE. 2009. Ebola virus glycoprotein GP masks both its own epitopes and the presence of cellular surface proteins. J. Virol. 83:9596–9601. http://dx.doi.org/10.1128/JVI.00784-09.

56. Blondot ML, Dubosclard V, Fix J, Lassoued S, Aumont-Nicaise M, Bontems F, Eleouet JF, Sizun C. 2012. Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2–1 protein. PLoS Pathog. 8:e1002734. http://dx.doi.org/10.1371/journal.ppat.1002734.

57. Leyrat C, Renner M, Harlos K, Huiskonen JT, Grimes JM. 2014. Structure and self-assembly of the calcium binding matrix protein of human metapneumovirus. Structure 22:136–148. http://dx.doi.org/10.1016/j.str.2013.10.013.

58. Leyrat C, Renner M, Harlos K, Huiskonen JT, Grimes JM. 2014. Drastic changes in conformational dynamics of the antiterminator M2-1 regulate transcription efficiency in Pneumovirinae. eLife:02674. http://dx.doi.org/10.7554/eLife.02674.

59. Gerard FC, Ribeiro Ede A Jr., Leyrat C, Ivanov I, Blondel D, Longhi S, Ruigrok RW, Jamin M. 2009. Modular organization of rabies virus phosphoprotein. J. Mol. Biol. 388:978–996. http://dx.doi.org/10.1016/j.jmb.2009.03.061.

60. Pei J, Grishin NV. 2001. AL2CO: calculation of positional conservation in a protein sequence alignment. Bioinformatics 17:700–712. http://dx.doi.org/10.1093/bioinformatics/17.8.700.

61. Hamby SE, Hirst JD. 2008. Prediction of glycosylation sites using random forests. BMC Bioinformatics 9:500. http://dx.doi.org/10.1186/1471-2105-9-500.

62. Svergun D. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. 25:495–503. http://dx.doi.org/10.1107/S0021889892001663.

63. Tosatto SC. 2005. The victor/FRST function for model quality estimation. J. Comput. Biol. 12:1316–1327. http://dx.doi.org/10.1089/cmb.2005.12.1316.