The human cytomegalovirus UL100 gene encodes the gC-II glycoproteins recognized by Group 2 monoclonal antibodies

Bruce Kari,* Wei Li, Jeff Cooper, Ron Goertz and Bob Radeke

Biomedical Research Center, Children's Hospital, St Paul, Minnesota, U.S.A.

In human cytomegalovirus (HCMV) the envelope glycoprotein complexes designated gC-II contain two immunologically and biochemically distinct glycoproteins. Monoclonal antibodies (MAbs) recognizing the gC-II glycoproteins have been divided into two groups based on the Mr of the glycoproteins they recognize. We have now identified the HCMV UL100 gene as the gene encoding the gC-II glycoprotein recognized by the Group 2 MAbs. To do this, gC-II complexes were immunoaffinity purified and cleaved with cyanogen bromide (CNBr). CNBr peptides were separated by reverse phase high performance liquid chromatography (RPHPLC). Amino acid sequences which matched sequences found in the protein encoded by the HCMV UL100 gene were obtained from three purified peptides. To confirm the assignment we made synthetic peptides using amino acid sequence from the carboxyl terminus of the protein encoded by the UL100 gene. These peptides were used to make murine antibodies. The anti-UL100 antibodies immunoprecipitated gC-II complexes and were reactive with gC-II glycoproteins recognized by Group 2 MAbs in Western blotting. Several overlapping UL100 fusion proteins were expressed in E. coli. Only one of these fusion proteins was recognized by gC-II Group 2 MAbs. None of these UL100 fusion proteins were recognized by gC-II Group 1 MAbs. These data showed that the UL100 gene encoded the gC-II glycoprotein recognized by the Group 2 MAbs and that the epitope recognized by these antibodies was located between amino acids 315 to 372 at the carboxyl terminus.

Introduction

It has been established that herpesviruses use host cell plasma membrane heparin like proteoglycans as one attachment site (WuDunn & Spear, 1989; Mettenleiter et al., 1990; Lycke et al., 1991; Okazaki et al., 1991; Kari & Gehrz, 1992; Neyts et al., 1992; Shieh et al., 1992). Attachment to these proteoglycans is not absolutely necessary for infection, but it is required for efficient infection of the host cell (Grunheid et al., 1993). In human cytomegalovirus (HCMV) we have shown that a family of complexes designated gC-II is a heparin binding component of the envelope (Kari & Gehrz, 1992). Monoclonal antibodies (MAbs) recognizing gC-II glycoproteins can be divided into two groups. Group 1 MAbs recognize a glycoprotein which forms a broad band in SDS-PAGE covering Mr's of 47000 to 63000. Group 2 MAbs recognize glycoproteins with Mr's of 48000 to 200000 (Kari et al., 1990). The glycoproteins recognized by the Group 1 and Group 2 MAbs can also be differentiated by their glycan composition. Glycoproteins recognized by Group 1 MAbs contain a high content of O-linked glycans whereas those recognized by Group 2 MAbs contained N-linked glycans (Kari & Gehrz, 1993, 1988). Recently, we have shown that the glycoproteins recognized by Group 2 MAbs have affinity for heparin in the absence of gC-II glycoproteins recognized by Group 1 MAbs (Kari & Gehrz, 1993). These data suggest that the glycoproteins recognized by Group 2 MAbs contained the heparin binding site in the gC-II complexes. In order to better understand the structure and heparin binding properties of the gC-II complexes, we initiated experiments to identify the HCMV gene which encodes the gC-II glycoprotein recognized by the Group 2 MAbs. In this report we show that this glycoprotein is encoded by the HCMV UL100 gene.

Methods

Virus production, labelling of virion glycoproteins and extraction of envelope glycoproteins. Towne strain HCMV was used and was grown in skin fibroblast cultures as previously described (Kari et al., 1986). Virion glycoproteins were labelled by incubating infected cells with [3H]glucosamine (GlcN) (40 mCi/ml, 31 Ci/mmol; New England Nuclear) in Dulbecco's modified Eagle's medium. Virions were purified by differential centrifugation and the envelopes solubilized by mixing virions in a Tris-NaCl (TN) buffer (10 mM-Tris, 150 mM-NaCl, 1 mM-PMSF, pH 7.5) containing 10% NP40 for 1 h. Insoluble material was removed by centrifugation leaving the envelope glycoproteins in the supernatant.

Isolation of gC-II complexes for CNBr digestion. gC-II complexes were immunoaffinity purified from NP40 extracts of virions. Immuno-
affinity columns were made using gC-II Group 2 MAb 27B4 and a Pierce Ag/Ab immobilization kit 1. Virion extracts were applied to the affinity column and the columns were washed with several column volumes of phosphate-buffered saline (PBS, pH 7.0) containing 0.1% NP40. The columns were then washed with PBS followed by water to remove as much NP40 as possible since residual NP40 would interfere with u.v. analysis of the CNBr peptides generated. The bound gC-II complexes were eluted with a 0.1 M-glycine–HCl buffer (pH 2.7). The gC-II complexes were completely soluble under acidic conditions. The gC-II complexes were concentrated in a speed vac (Savant) and then resuspended in 80% formic acid. Excess CNBr was added and the reaction vial purged with N2. The reaction was allowed to proceed overnight in the dark. The reaction was stopped by addition of water and the samples concentrated in a speed vac. Water was added to the samples a second time and the samples concentrated again with a speed vac.

Reduction and HPLC separation of gC-II peptides. Concentrated samples were resuspended in a 0.25 M-Tris buffer (pH 8.5) containing 6 M-guanidine. HCl and 1 nM-EDTA. Dithiothreitol (DTT; 5 mg) was added, the reaction vial purged with N2, and the reaction left at 37 °C for 1 h. The sample was then filtered with a 0.45 µm filter and applied to a Vydec C18 reverse phase HPLC (RP-HPLC) column (46 × 250 mm, 300 Å pore size). The solvent systems were: A, 0.06% trifluoroacetic acid (TFA) in water; B, 80% acetonitrile in 0.06% TFA. The gradient used was 20% B to 38% B during the first hour, 38% B to 75% B during the next 30 min, and 75% B to 98% B for the next 15 min. The flow rate was 0.5 ml/min. The column eluate was monitored at 225 nm and individual peaks were collected. Peptides in these peaks were sequenced using a Porton peptide sequencer.

Synthesis of synthetic peptides and production of monospecific antibodies and MAbs. Monospecific antibodies were made using synthetic peptides. Synthetic peptides were made using an Applied Biosystems 430 A peptide synthesizer using FMoc chemistry. The peptides were based on the carboxyl terminus amino acid sequences of the protein encoded by the HCMV UL100 gene. These sequences were HHDSLESRLREEED (amino acids 348 to 362, peptide 1) and REEEDDDDEDFEDA (amino acids 358 to 372, peptide 2). The peptides were conjugated to keyhole limpet haemocyanin and the conjugates used to immunize mice. Sera from immunized mice were tested against the peptides conjugated to BSA in an ELISA assay. Those mice which had a high titre to a peptide were given intraperitoneal injections of SP2 myeloma cells to induce the production of ascites. The ascites was collected and screened for antibody to the peptides. The gC-II Group 1 MAbs 8B4 and 9E10 and the gC-II Group 2 MAbs 12G9, 15G5, 15F9, 27B4, 25C8, 23B10 and 40B7 used in this study were described previously (Kari et al., 1990).

SDS-PAGE, immuno precipitations and Western blotting. SDS-PAGE was done using a Bio-Rad mini gel apparatus and the Laemmli (1970) buffer system. For SDS-PAGE, samples were solubilized in SDS by heating at 100 °C for 3 min. The relative amounts of proteins in Western blots were determined by densitometric scans of photographic negatives of the blots. Western blotting was done as described previously (Lussenhop et al., 1988).

UL100 fusion proteins. The E. coli expression vector pMAL-c (New England BioLabs), which generates a maltose binding protein (MBP) fusion protein, was used to clone three overlapping regions of the UL100 gene. The PCR products encoding UL100 amino acids (aa) 1 to 160, UL100 aa 135 to 315 and UL100 aa 263 to 372 were inserted into pMAL-c. These constructs were used to transform E. coli cells. After IPTG induction, bacterial cells were pelleted and directly lysed in 1 × Laemmli sample solubilization buffer for SDS-PAGE. The E. coli lysates containing MBP and MBP-UL100 fusion proteins were separated on 8% SDS-polyacrylamide gels and analysed by Western blotting as described previously (Lussenhop et al., 1988).

Results

Preparation and sequence analysis of gC-II CNBr fragments

gC-II complexes were immunoaffinity purified and cleaved with CNBr. The resulting peptides were reduced with DTT, and separated by RPHPLC; individual peaks were collected and the amino acid sequences of the peptides in these peaks determined. Amino acid sequences were obtained from three peptides. Although other peptides were separated, only these peptides provided sequence data which could be interpreted. The sequences from these three peptides were compared to the predicted amino acid sequences for proteins of HCMV strain AD169 available through GenBank. Matches were made with sequences of the protein encoded by the HCMV UL100 gene. The amino acid sequence for this protein and the sequences of the gC-II CNBr peptides are shown in Fig. 1. Note that the amino terminus of each CNBr peptide starts after a methionine residue. This result was expected since CNBr was used to cleave gC-II glycoproteins. Furthermore, peptides 2 and 3 eluted from the reverse phase column in a single peak (data not shown). Even though this occurred sequences could still be obtained because these two peptides were not present in equal amounts. Thus, for each turn of the sequencer two peaks were obtained. One peak was always twice the area of the second peak. The larger peaks were taken to be amino acids from one peptide while the smaller peaks amino acids from a second peptide. If peptides 2 and 3 were not physically associated with each other, it would seem unlikely that they would co-elute from a reverse phase column. For example, fragment 2 contained 16 amino acids and fragment 3 contained 58 amino acids. Fragment 3 also contained many acidic amino acids and fragment 2 did not. It is interesting to note that fragments 2 and 3 contain cysteine residues (Fig. 1). It may be that disulphide bonds between these peptides were not reduced even though the gC-II complexes were subjected to denaturing conditions in the presence of DTT. It is also possible that these peptides remained associated by other types of bonds. Nonetheless, these data showed that the UL100 gene most likely encoded a gC-II glycoprotein.

Characterization of anti-UL100 peptide antibodies by immunoprecipitation

To support the assignment of the UL100 gene as a gene encoding a gC-II glycoprotein, we made synthetic peptides using amino acid sequences from the carboxyl...
MAPSHVD UNIXRT W SASIFMLTFVNSVHLVLSNF PHLGYPCV
YYHVDFLERL NMSAYVMLHOTPMLFLDSVQLCYAVFMQLVE
LAVTIYYLVICWIKISMRKD R GMSI N QS TRDL SYMDSLT AF LFLSM
1
DTFQLFTLM SPRL S MIAFMAAVHHFCLTFNVMQVRQSYK+R+S
LFFFSR+HL+PK+LK+GTQFR+TLVNLVEALGFNTTVVAMALC* YFGNN
2
FFVRTGRH MVLAFFYY AISIYLFLIEAVFFQYVKGQFGYHLGAF
FFVRTG WHMVLAFFYY AISIYLFLIEAVFFQYVKGQFGYHLGAF
3
GLGCIYIYDFTLSNYERTGISWSFGMLFFIWAM PTTC*RAVRY
FRGRSGSVKYQALATASGEVEAVLHSDLESRRLEEDDDDEFEDA

Fig. 1. Amino acid sequence of the protein encoded by the UL100 gene. The single letter designation is used. Hydrophobic sequences predicted by Hopp/Woods and Kyte/Doolittle algorithms are shown in bold type. The sequences obtained by amino acid sequence analysis of the CNBr peptides obtained from gC-II complexes are underlined and numbered. Methionine residues cleaved to generate the peptides are indicated in large type. Cysteine residues are indicated with an asterisk. This protein also contains an area which has a high proportion of basic amino acids. The positively charged amino acids in this area are indicated with a plus sign in the fourth and fifth lines down. Potential N-glycosylation sites are bold underlined and are in larger type.

(a) (b) (c)

Fig. 2. Virions were labelled with [3H]GlcN and extracted with NP40 and the extracts used for immunoprecipitation. (a) Extracts were immunoprecipitated with MAb 27B4 or with anti-UL100 peptide 2 antibody. The immunoprecipitates were solubilized in SDS at 100 °C and reduced, prior to SDS-PAGE and fluorography. Lane 1, glycoproteins immunoprecipitated by MAb 27B4. Lane 2, glycoproteins immunoprecipitated with anti-UL100 peptide 2 antibody. (b) Serial immunoprecipitations were done by using anti-UL100 peptide 2 antibodies first and then Group 2 MAb 27B4. Immunoprecipitated gC-II complexes were examined by SDS-PAGE and fluorography under non-reducing conditions. Lane 1, complexes immunoprecipitated by anti-UL100 peptide 2 antibodies. Lane 2, complexes immunoprecipitated by Group 2 MAb 27B4 after the UL100 antibody. (c) Western blot analysis was done with whole virions which had been solubilized at 100 °C in SDS and reduced prior to SDS-PAGE. Lane 1, SP2 negative control. Lane 2, glycoprotein detected with Group 1 MAb 9E10. Lanes 3 and 4, glycoprotein detected with Group 2 MAb 27B4 and 15F9 respectively. Lanes 5 and 6, glycoproteins detected by polyclonal antibodies to UL100 peptides 1 and 2. Numbers to the left indicate Mr markers.

terminus of the protein predicted by the UL100 gene. These were designated UL100 peptide 1 (aa 348 to 362) and UL100 peptide 2 (aa 358 to 372) respectively. These peptides were used to make polyclonal antibodies in mice. We first examined the antibodies for their ability to immunoprecipitate gC-II complexes. To do this, gC-II complexes were labelled with [3H]GlcN and extracted from purified virions in NP40. These extracts were immunoprecipitated with Group 2 MAb 27B4 or with anti-UL100 peptide 2 antibodies. These immunoprecipitates were reduced with 2-mercaptoethanol prior to SDS–PAGE. The glycoproteins immunoprecipitated by MAb 27B4 and the anti-UL100 peptide 2 antibody were typical of gC-II. Moreover, the relative amounts and Mr's of the glycoproteins were identical when comparing the two immunoprecipitates (Fig. 2a, lanes 1 and 2). Serial immunoprecipitations were also done by using anti-UL100 peptide 2 antibodies first followed by Group 2 MAb 27B4. These immunoprecipitates were not reduced prior to SDS–PAGE. When anti-UL100 peptide 2 antibody was used complexes typical of gC-II were immunoprecipitated (Fig. 2b, lane 1). When gC-II MAb 27B4 was used after anti-UL100 antibody only a small amount of additional gC-II was immunoprecipitated.
Fig. 3. Western blot analysis using extracts from E. coli transformed with constructs containing the p-MAL vector only or the pMAL vector and one of the three overlapping regions of the ULI00 gene. These three regions contained aa 1 to 160, aa 135 to 315 or aa 263 to 372. MBP, extracts from E. coli cells expressing maltose binding protein only. MBP UL100 aa 1 to 160, extracts from E. coli cells expressing the fusion protein MBP with ULI00 aa 1 to 160. MBP UL100 aa 135 to 315, extracts from E. coli cells expressing fusion protein MBP UL100 aa 263 to 315. MBP UL100 aa 263 to 372, extracts from E. coli cells expressing fusion protein MBP UL100 aa 263 to 372. In all cases, lane 1 is the SP2 negative control, lane 2 gC-II Group 1 MAb 8B4, lanes 3 and 4 gC-II Group 2 MAbs 40B7 and 15F9 respectively and lane 5 anti-UL100 peptide 1 antibody, fp, Fusion protein; Im, low Mr proteins detected non-specifically.

(Fig. 2b, lane 2). These data established that the anti-UL100 antibodies recognized the same gC-II complexes as our gC-II MAbs and support the designation of the ULI00 gene as a gene encoding a gC-II glycoprotein.

To determine if the ULI00 gene encoded glycoproteins recognized by gC-II Group 1 or Group 2 MAbs, Western blot analysis was done with virions as the antigen. Samples were solubilized in SDS by heating at 100 °C. As a negative control, SP2 ascites was used and no reactivity was detected (Fig. 2c, lane 1). The gC-II glycoprotein detected with Group 1 MAb 9El0 is shown in Fig. 2(c), lane 2. This glycoprotein formed a broad band covering Mr 47 000 to 63 000. Glycoproteins recognized by Group 2 MAbs 27B4 and 15F9 are shown in Fig. 2(c), lanes 3 and 4. These MAbs reacted with a relatively well focused band with an Mr of 48 000. However, they also reacted with proteins having Mr's of 90 000 and greater (Fig. 2c, lanes 3 and 4). The ascites containing antibodies to ULI00 peptides reacted with the same glycoproteins as the gC-II Group 2 MAbs, including the proteins which had Mr's of 90 000 and greater (Fig. 2c, lanes 5 and 6). The Mr of the protein predicted from the ULI00 gene would be approximately 45 000 without post-translational processing. Thus, the reactivity with the high Mr proteins would not be expected. Since the ULI00 protein is very hydrophobic it would seem likely that these high Mr bands were aggregates of the Mr 48 000 protein. Nonetheless, these data showed that the Group 2 MAbs recognized the ULI00 gene product and that the Group 1 MAbs did not.

gC-II Group 2 MAbs and the anti-peptide antibodies recognize a ULI00 gene product expressed in prokaryotic cells

To further establish the ULI00 gene as the gene encoding the gC-II glycoproteins recognized by the Group 2 MAb, three overlapping regions of the ULI00 gene were cloned into a pMAL vector and these constructs expressed in E. coli. Cells were extracted with SDS-PAGE sample solubilization buffer. These extracts were analysed by SDS-PAGE and Western blotting using gC-II MAbs and anti-UL100 antibodies. Cell extracts from E. coli expressing only the pMAL protein or the ULI00 fusion proteins, ULI00 aa 1 to 160 or ULI00 aa 135 to 315, did not contain any proteins recognized by any of the gC-II MAbs in Western blotting. Cell extracts from E. coli expressing the fusion protein containing the ULI00 carboxyl terminus (UL100 aa 263 to 372) all the gC-II Group 2 MAbs and the anti-UL100 antibodies recognized a protein (representative data for MAbs 40B7, 15F9 and ULI00 #1 antibody are shown in Fig. 3, MBP UL100 aa 263 to 372, lanes 3, 4 and 5). In contrast, the gC-II Group 1 MAb 8B4 or the SP2 negative control did not react with this protein (Fig.
Discussion

The HCMV UL100 gene has been identified as the gene encoding the gC-II glycoproteins recognized by the Group 2 MAbs. This conclusion is based on three separate observations. First, CNBr peptides were obtained from gC-II glycoproteins which had sequences found in the protein encoded by the UL100 gene. Second, anti-UL100 antibodies recognized the same gC-II glycoproteins as the Group 2 MAbs in Western blot analysis. Third, the pMAL fusion protein containing aa 263 to 372 of the UL100 glycoprotein was recognized by all gC-II Group 2 MAbs. An M, 45,000 envelope protein encoded by the UL100 gene has also been described by Lehner et al. (1989). The UL100 gene has homology to predicted translation products from EBV (BRRF3), VZV (ORF50) and HSV (UL10) (Lehner et al., 1989). Like UL100 all of these proteins have multiple hydrophobic domains and would most likely span the membrane several times. However, these proteins do not contain acidic carboxyl termini like the UL100 protein. Another protein which is similar to these proteins is the Epstein–Barr virus (EBV) latent membrane protein (LMP) encoded by the BNLF1 gene. It has been established that the LMP protein of EBV is expressed in the plasma membrane of the infected cell, is found in the virion (Mann et al., 1985) and is a target for cytotoxic T cells (Thorley-Lawson & Isaelsen, 1987). The LMP protein and the UL100 glycoprotein have several characteristics in common. First, both the UL100 glycoprotein and the LMP protein have several hydrophobic domains which would allow them to span a lipid bilayer several times (Liebowitz et al., 1986). Second, the amino terminus of both proteins is hydrophilic and thus not characteristic of a signal sequence. This would suggest that the amino terminus of both proteins is internal. Third, both proteins have acidic carboxyl termini. While these similarities exist, the LMP protein or the homologous proteins in other herpesviruses are not heparin binding proteins. It is interesting to note that in HSV glycoprotein C (gC) is the principal heparin binding envelope glycoprotein (Herold et al., 1991). Whereas HSV contains a homologue to the UL100 glycoprotein, HCMV does not contain a homologue to the HSV gC glycoprotein. It may be that the HCMV UL100 glycoprotein evolved to provide the heparin binding function found on the HSV gC glycoprotein. However, the exact role or roles played by the UL100 protein in virion replication will require further studies.

This work was supported by NIH grant HD19937-02 and by a grant from Children's Biomedical Research Center, St Paul, Minnesota.

References


(Received 24 January 1994; Accepted 27 June 1994)