Arch Virol (1990) 113: 1–16

Archives of Virology

Cloning and characterization of major antigenic determinants of human cytomegalovirus Ad169 seen by the human immune system

W. Lindenmaier, Antje Necker*, Susanne Krause, Renate Bonewald, and **J. Collins**

GBF – Gesellschaft für Biotechnologische Forschung mbH, Braunschweig, Federal Republic of Germany

Accepted March 27, 1990

Summary. Starting from a cosmid library of HCMV Ad169-DNA random fragments of DNA were generated. Fragments about 200 to 600 bp in length were selected and cloned into open reading frame (ORF) expression vectors to create ORF-libraries that represent either the entire viral genome or defined subregions. About 120.000 clones were isolated and screened immunologically for the synthesis of fusion proteins consisting of an antigenic peptide encoded by the CMV sequence coupled to a truncated *E. coli* β-galactosidase molecule. Anti-CMV sera raised in animals as well as human hyperimmune globulin were used for colony screening. Distinct sets of antigenic fusion proteins were recognized by different antisera. Ten of the clones giving strong reactions with human immune sera were mapped on the CMV genome and the sequences of the CMV inserts determined. Antibodies against fusion proteins were raised in mice or rabbits to identify the corresponding CMV proteins. Antigenic fusion proteins described here were recognized by most individual human CMV immune sera tested. They allow determination of the humoral immune response to defined determinants and may therefore be particularly useful in diagnosis and vaccine development.

Introduction

Human CMV is a common pathogen of man. Primary CMV infections usually proceed without notice and the virus persists lifelong in a latent infection. However, in the immunocompromized host, major detrimental effects and life threatening disease are caused by CMV infection or reactivation (for review see [1]). Effective chemotherapy and protective immunization are not available.

* Present address: Centre d'Immunologie INSERM-CNRS de Marseille-Luminy, Case 906, F-13288 Marseille Cedex 9, France.

CMV is a member of the herpes virus group having a double stranded DNA genome of about 235 kb. About 30 structural proteins of the virion have been described [8, 36]. In addition a number of non-structural proteins are produced in infected cells. All of these polypeptides may be targets for the host's immune response during natural infection.

Using genetic engineering methods detailed analysis of relevant antigens is now feasible. Open reading frame cloning of c-DNA or genomic fragments, expression of fusion proteins in *E. coli* and immunological screening allow localization of antigenic determinants even in complex genomes. Immune screening using antisera or monoclonal antibodies to CMV-proteins allowed characterization and mapping of a number of CMV-genes from cDNA libraries of CMV infected cells [13, 14, 21, 23, 33]. These antibodies, however, often recognize distinct subsets of antigenic determinants which may not be representative for the human immune response against CMV. We therefore screened genomic ORF-libraries using human CMV-positive sera to identify antigenic determinants recognized by the human immune system.

Materials and methods

Plasmids and bacterial strains

CMV-cosmids pCM1050, 1058, 1065, 1106, 1007, 1029, 1049, 1017, 1015, 1035, and 1072 have been described [7]. pEX-vectors [35] were a gift from Dr. K. Stanley. pEXStuI [41] is a modification of pEX1. The synthetic oligonucleotide GATCTATCGAGGGT-AGGCCTG was integrated into the BamHI site of pEX1. Thereby the open reading frame of β-galactosidase is extended and blunt-ended DNA fragments can be inserted into the StuI-site immediately after the codons for amino acids Ile-Glu-Gly-Arg, the recognition sequence of protease factor Xa [29]. *E. coli* strains 5K or DH1 harboring pRK248cIts [3] and N4830 [9] were used for expression of ORFs cloned in pEX-vectors.

Viral antigens

HCMV Ad169 derived viral antigens were used. Stocks of HCMV Ad169 and primary human fibroblasts were provided by T. H. The, Groningen. Nuclear antigens and NP40-soluble antigens [20] were gifts of Dr. J. Middeldorp, Organon, Oss. Purified virions and dense bodies were kindly donated by Dr. B. Nowak, Erlangen.

Antisera and immunization

Goat-anti-CMV serum was obtained from Polyscience. Human CMV hyperimmune globulin was from Biotest. Mixed human sera and sheep-anti-CMV serum were from Dr. J. Middeldorp, Oss, rabbit-anti-CMV-virion serum was a gift of Dr. B. Fleckenstein, Erlangen. Human serum samples from renal transplant patients were provided by Dr. T. H. The, Groningen. pp28 specific monoclonal antibody P2G11 [31] was kindly provided by Dr. M. P. Landini, Bologna.

For the preparation of rabbit sera directed against fusion protein about 400 µg of purified CMV-β-galactosidase fusion protein was concentrated by aceton precipitation, resuspended and emulsified with an equal volume of adjuvant. Complete Freund's adjuvant was used for the first injection, incomplete adjuvant for two subsequent injections. Mice were immunized first with about 100 µg of fusion protein injected intraperitoneally and booster immunizations were done after 1, 4, and 7 weeks.

ORF-cloning

In general, standard cloning procedures described in Maniatis et al. [22] were used. To generate blunt-ended, random fragments from pCM-cosmids, DNA was either digested with DNAse I in the presence of Mn^{2+}-ions or fragmented by sonification. Fragments of about 200–600 bp in length were isolated from agarose gels. A library of cosmid fragmented in pEX1 was produced using the oligo (dG)-oligo (dC) tailing method [10]. For the second ORF-library and libraries of subgenomic regions containing the major late antigen [30] blunt ended fragments were cloned in pEXStuI. *E. coli* 5K [pRK248cIts] and *E. coli* DH1 [pRK248cIts] were used as host strains.

Immunological screening

Recombinants expressing CMV-β-galactosidase fusion proteins were plated at a density of about 5,000 colonies per plate (85 mm ∅) and incubated at 30 °C until colonies became visible. Replicas on nitrocellulose filters were induced for production of fusion proteins by incubation at 42 °C for 3 h. Lysis of colonies on the nitrocellulose filters was carried out according to Verweij et al. [38]. The immunological reactions were performed according to the protocol of Biorad, Richmond. Unspecific bindig of antibodies was blocked by incubation of the colony-filters in T-TBS (20 mM Tris-HCl, 500 mM NaCl, 0.05% Tween 20, pH 7.5) containing 3% gelatine for 30 min. Sera were preincubated for 30 min at room temperature with protein lysates of corresponding strains without CMV insert to block binding of antibodies to *E. coli* determinants (200 µg *E. coli* protein/ml). Filters were washed twice in T-TBS and incubated with CMV-specific antisera for 16 h. After removal of unbound antibody by washing in T-TBS bound antibodies were detected by horseradish peroxidase labelled secondary antibodies (swine-anti-goat IgG (Tago) and goat-anti-human IgG (Biorad)Z for 4 h. Filters were washed twice in T-TBS and twice in TBS. Positive colonies were visualized by incubation with staining solution (3 mg/ml chloronaphtol in methanol diluted 1 : 6 in TBS, 1/2,000 vol. of 30% H_2O_2).

Fusion proteins

E. coli clones coding for CMV-β-galactosidase fusion protein were grown to $OD_{600} = 0.3$ in L-broth at 30 °C. Production of fusion protein was induced by shifting to 42 °C for 30 min and allowed to continue for further 2 h at 37 °C. Cells were collected by centrifugation, resuspended in 1/20 volume of 40 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.3 mg/ml lysozyme and incubated for 5 min at 0 °C. An equal volume of 20 mM $MgCl_2$, 20 mM Tris-HCl, pH 7.5, and 10 µg/ml DNAse I was added and incubation continued for another 5 min. For analytical purposes cells were lysed by addition of 1/200 vol of 10% sodium deoxycholate and inclusion bodies were solubilized by addition of 1/10 vol of urea buffer (9 M urea, 4% β-mercaptoethanol), 4% Triton X-100, 4% SDS, in 20 mM Tris-HCl, pH 7.5, 20% glycerol). For preparative purification of fusion proteins lysozyme-DNAse I treated cells were lysed by sonification. Inclusion bodies were pelleted by centrifugation, solubilized in 8 M urea, 50 mM Tris-HCl, pH 8.0, and dialyzed against 50 mM Tris-HCl, pH 8.3, 227 mM NaCl. After centrifugation (8,000 rpm, 5 min) the supernatant was applied to sepharose 6B-CL chromatography in 100 mM Tris-HCl, pH 7.0, 100 mM NaCl, 0.1% SDS. Fractions containing fusion protein were pooled and concentrated.

Blotting techniques

Immunoblots were performed as described [37]. Transferred proteins were stained with Ponceau S (Sigma) according to the instructions of the manufacturer. Mouse anti human IgM antibodies (BioYeda) were used to detect bound human IgM. Horseradish peroxidase labelled secondary antibodies (swine-anti-goat IgG (Tago) goat-anti-rabbit, goat-anti-

mouse and goat-anti-human IgG (Biorad) and chloronaphtol/H_2O_2 staining was used to visualize the antigens. Southern blots of CMV-cosmids were performed according to Maniatis et al. [22]. CMV insert fragments to be used as hybridization probes were isolated from agarose gels and ^{32}P-labelled by nicktranslation.

DNA sequencing

CMV insert fragments from clones of the oligo-(dG, dC) tailed library were sequenced according to Maxam and Gilbert [24]. Dideoxysequencing was used for pEXStuI clones. Direct plasmid sequencing was performed as described by Hattori and Sakaki [11] in conjunction with pEXStuI specific olignucleotide primers, GCCCGTCAGTATCGGCGG and CCGGTCAATCAATCAGC, kindly synthesized by the GBF DNA synthesis group (Dr. H. Blöcker).

Results

ORF-libraries

The aim of our studies was a more detailed characterization of antigenic determinants of HCMV which are recognized in the human humoral immune response. The system we chose was expression of open reading frames of CMV in *E. coli* and subsequent screening for fusion proteins recognized by human CMV immune sera. For the expression of ORFs as β-galactosidase fusion proteins in *E. coli* pEX1 [35] and pEXStuI which produce C-terminal fusions controlled by a temperature-sensitive bacteriophage λ repressor, proved to be superior to lac-repressor controlled N-terminal β-galactosidase fusions [15]. pEXStuI is a modification of pEX1 which contains the recognition site for protease factor Xa directly in front of the StuI cloning site [40]. CMV cosmid clones [7] used for ORF-cloning are indicated in Fig. 1. Random fragments of cosmid DNA were generated by DNAse I digestion in the presence of Mn^{2+} ions. Gel purified fragments of 200–600 bp were either tailed with oligo(dG) using terminal transferase and annealed to oligo(dC) tailed pEX1 or ligated in

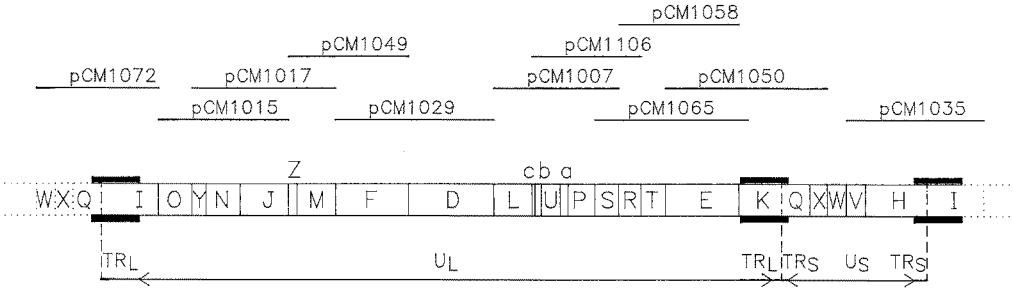

Fig. 1. CMV cosmid clones used for the construction of ORF-libraries. The HindIII restriction map of HCMV Ad169 is given [7]. Dashed lines delimit the small and large segments. Repeated elements are indicated by filled boxes. The location of repeated sesegments. Repeated elements are indicated by filled boxes. The location of repeated seare given below the restriction map. Cosmids used for ORF-cloning are indicated above the corresponding HindIII fragments. Fragment combinations derived from alternative conformations are indicated by dotted lines

blunt ended form to StuI digested pEXStuI. The pEX1-ORF-library and the pEXStuI-ORF-library consisted of about 1.5×10^4 and 1×10^5 independent clones with inserts >100 bp. One sixth of these clones should be in correct orientation and reading frame. In addition similar libraries of the subgenomic region encoding the so called "major late antigen" [30] were established. For these libraries CMV clones pCM1007 and pBW4 (Fig. 1, [33]) were used. The pCM1007-library and the pBW4-library represented 15,500 and 20,000 independent clones containing inserts of about 200–800 bp.

CMV-β-galactosidase fusion proteins

Clones which expressed CMV antigenic determinants were identified by immunological screening. Colonies grown on nitrocellulose filters were induced by a temperature shift to 42 °C and lysed in situ [38]. Goat-anti-CMV serum (Polyscience) and human CMV hyperimmune globulin (Biotest) were used for screening. To reduce background staining due to anti-*E. coli* reactivity of the sera, saturation with protein lysate of *E. coli* [pEXStuI] proved to be very important. Positively stained colonies were picked and retested in immunoblots. PAGE of protein lysates from induced cultures showed intense protein bands of a size equal or larger than β-galactosidase which was completely absent in uninduced cultures. Multiple bands are seen in lysates from "GC-tailed" library (e.g., CMV1, CMV2; Fig. 2A). The protein band of the size of β-galactosidase in protein lysates of clones from this library was not recognized by CMV immune

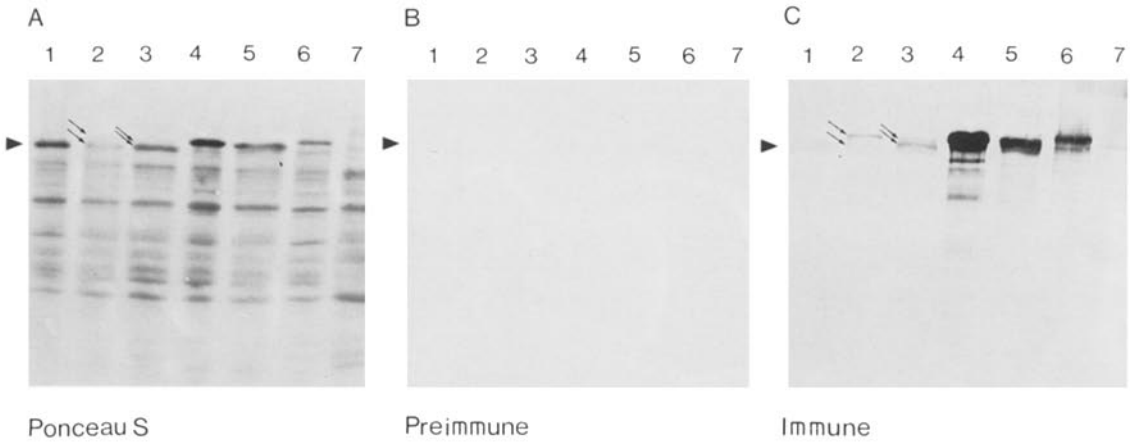

Fig. 2. Immunological detection of CMV antigenic determinants in CMV-β-galactosidase fusion proteins. *E. coli* cultures harboring pEX1 (*1*), pCMV1 (*2*), pCMV2 (*3*), pCMV13 (*4*), pCMV21 (*5*), pCMV23 (*6*), and pCMV24 (*7*), were induced for the production of fusion proteins and lysed as described in materials and methods. After separation by SDS-PAGE and blotting to nitrocellulose filters, proteins were stained with Ponceau S (**A**), destained and probed using goat anti CMV preimmune serum (**B**) and immune serum (**C**). Multiple induced protein bands in lysates of cells harboring pCMV1 and pCMV2 are indicated by arrows

sera. They were most probably products of premature termination due to complementary oligonucleotide tracts. On the other hand minor multiple immune reactive bands which were sometimes detected in protein lysates from the pEXStuI library were probably the result of partial degradation in *E. coli*. The fusion proteins reacted specifically with goat-anti-CMV immune serum but not with preimmune serum (Fig. 2B, C, and data not shown). They were produced as insoluble inclusion bodies in the *E. coli* cells [35]. This property allowed a considerable simplification of the purification procedure: pelleting of inclusion bodies, urea extraction and gel filtration yielded nearly homogenous fusion protein which was used to produce antisera monospecific for the CMV determinants expressed.

Genetic localization of antigenic determinants

The map position of clones derived from the genomic libraries were determined by hybridization of ^{32}P-labelled insert sequences to EcoRI/HindIII digested cosmid DNAs representing the complete CMV genome. Two clones (CMV1, CMV13) were mapped within HindIII fragment J (Fig. 3A, B), three (CMV2, CMV21, CMV23) were located in HindIII R (Fig. 3A, C). CMV26 was derived from HindIII M. Clone CMV24 insert hybridized to pHC79 vector DNA in addition to CMV specific fragments from HindIII F (Fig. 3A). DNA sequencing (see below) identified pBR322-derived sequences cloned in frame to the CMV insert. More accurate mapping was done by hybridization to multiple digests of positive cosmids. Identical subfragments were recognized by CMV1 and CMV13 within pCM1015 and by CMV2, CMV21, and CMV23 within pCM1058, respectively (Fig. 3B, C, and data not shown). Clones from the subgenomic libraries were mapped by colony hybridization with subfragments of pCM1007 cloned in pAN26 [19], a R1-derived vector without homologies to pEX1. 35 clones of these ORF-libraries were analyzed in detail. All clones which were strongly reactive with human sera were clustered in the HindIII b, c, L region which had been shown to encode the 65 k "major late antigen" and a 71 k phosphoprotein [30, 33]. No strongly immune reactive clone was derived from HindIII S, P, a, and U. In summary, clones strongly reactive with human sera were isolated from five regions of the HCMV genome (Fig. 3D).

Immunological reactivity

Taking into account how positive clones had been isolated, one would assume that preferentially strong immunodominant antigenic epitopes had been picked out. To test whether the CMV-β-galactosidase fusion proteins actually represented antigenic determinants recognized by a wide variety of CMV positive sera, reactivity of individual patient sera and additional animal sera was determined. Clones CMV1, CMV13, CMV23, CMV24, CMV26, B1B, C35, and

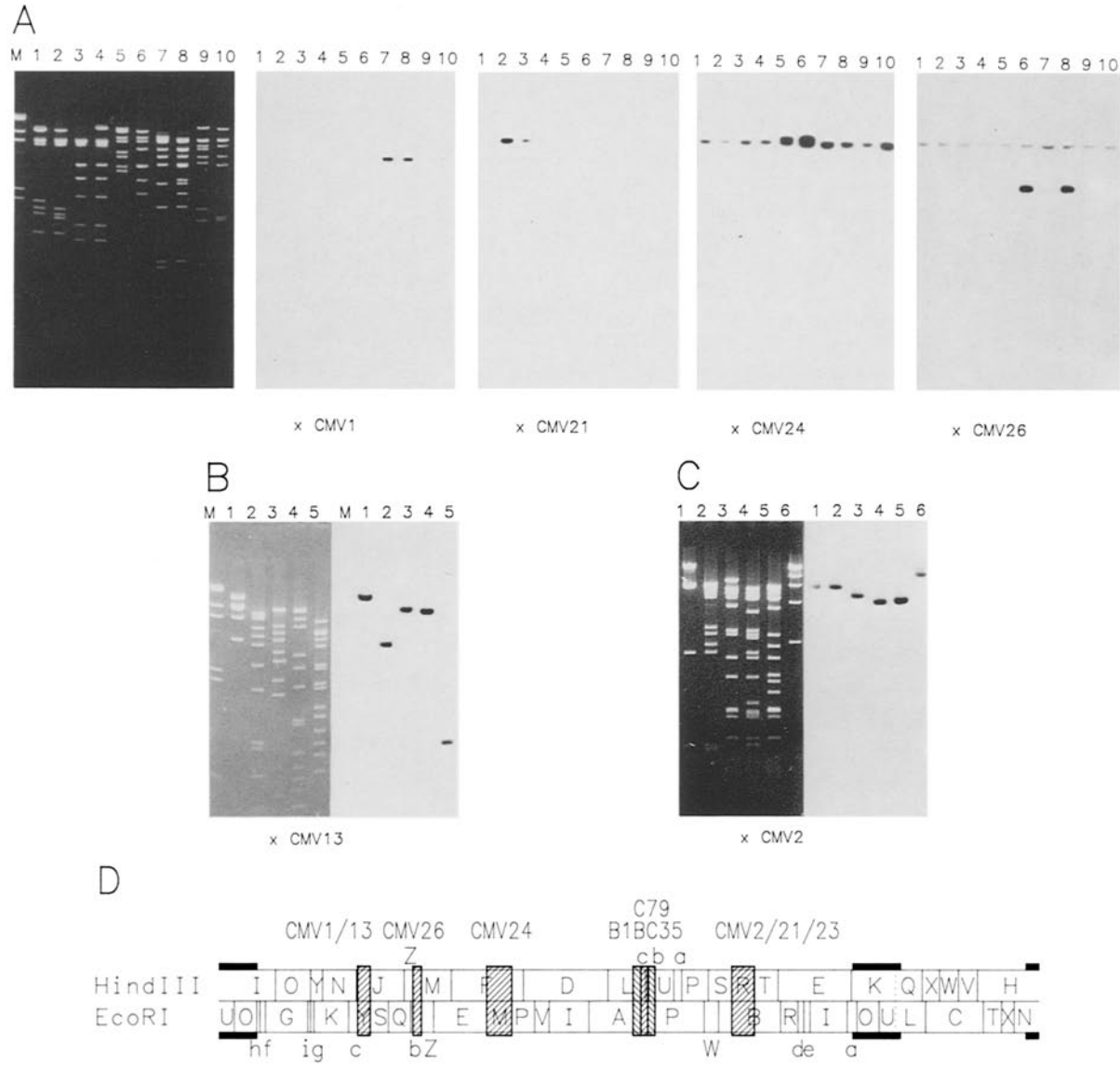

Fig. 3. Mapping of CMV antigenic determinants by Southern blotting. **A** CMV cosmids pCM1050 (*1*), pCM1058 (*2*), pCM1106 (*3*), pCM1007 (*4*), pCM1029 (*5*), pCM1049 (*6*), pCM1015 (*7*), pCM1017 (*8*), pCM1035 (*9*), and pCM1072 (*10*) digested with EcoRI and HindIII were hybridized with inserts of pCMV1, pCMV21, pCMV24, and pCMV26 as indicated below the autoradiograms. **B** CMV13 was mapped within cosmid pCM1015. pCM1015 was digested with HindIII (*1*), HindIII-EcoRI (*2*), HindIII-BglII (*3*), HindIII-BamHI (*4*), and HindIII-PstI (*5*). **C** CMV23 was hybridized to a Southern blot of pCM1058 cut with HindIII (*1*), HindIII-BamHI (*2*), PstI (*3*), HindIII-PstI (*4*), PstI-XbaI (*5*), and XbaI (*6*). λ-DNA digested with HindIII was used as DNA size marker (*M*). CMV inserts were ^{32}P-labelled by nick translation. **D** Genomic localisation of antigenic determinants CMV1 and CMV13, CMV26, CMV24, B1B, C35, C79, and CMV2/21/23 is indicated within the EcoRI and HindIII restriction maps of HCMV Ad169 by hatched boxes

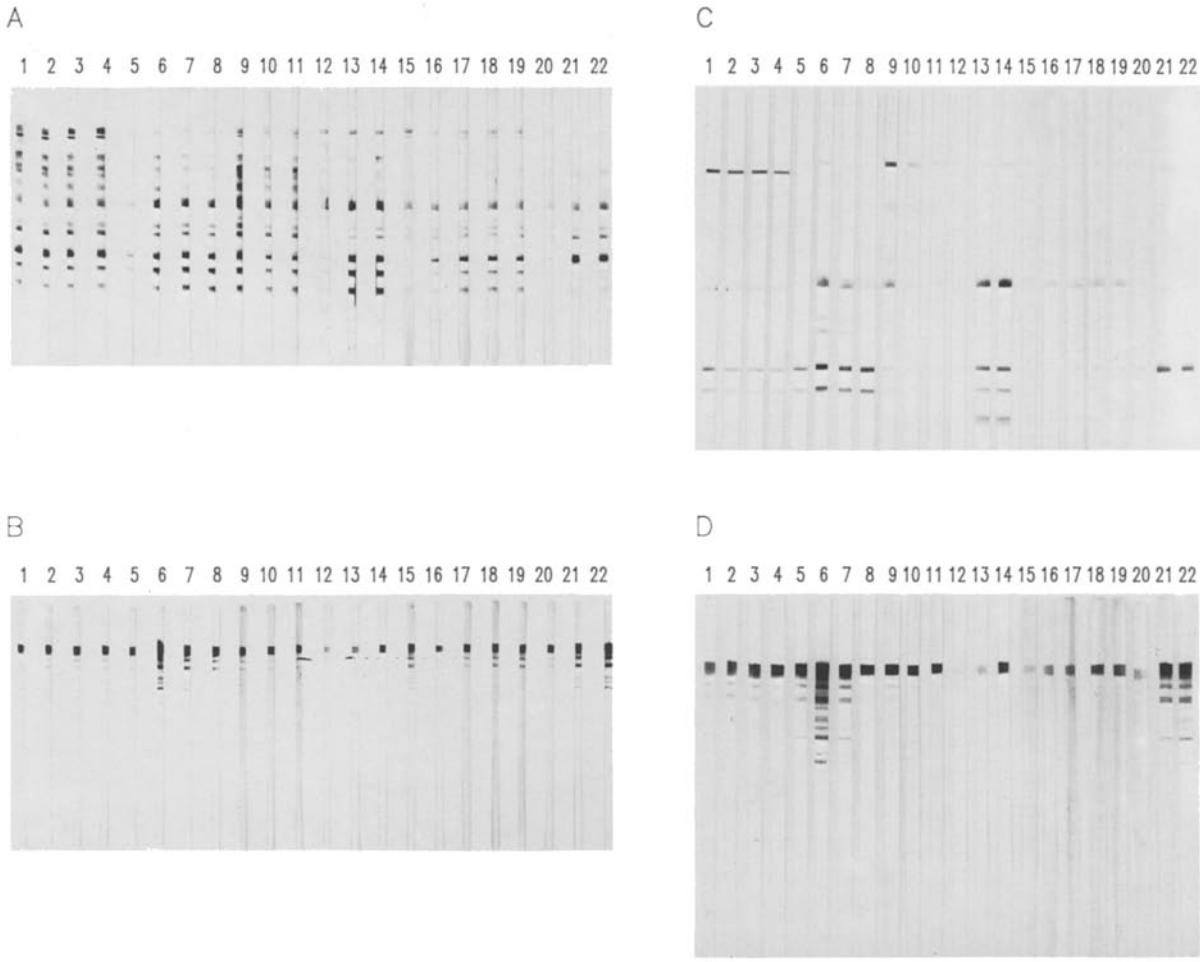

Fig. 4. CMV nuclear antigens (**A, C**) and proteins of the *E. coli* clone expressing CMV13 fusion protein (**C, D**) were separated by 10% SDS-PAGE and blotted to nitrocellulose filter paper. Filters were cut and individual strips were incubated with human patient sera (Table 1). Dilution of human sera for detection of IgG was 1 : 500. To establish IgM binding dilutions of 1 : 100 were used. Bound antibodies were detected by incubation with horseradish peroxidase labelled rabbit anti human IgG antiserum (**A, B**) or a monoclonal antibody specific for human IgM and a tertiary HRPO-labelled rabbit anti mouse Ig antiserum (BioYeda). Visualization of bound antibodies was done by H_2O_2/chloronaphtol staining

C79 which corresponded to different antigenic regions were tested. Immunoblots of fusion proteins and CMV antigens were incubated with individual sera at a dilution of 1 : 500 for IgG or 1 : 100 for IgM-reactions (Fig. 4, and data not shown). The results are summarized in Table 1. All fusion proteins were specifically recognized by CMV positive sera, but universal and most intense reaction was seen with CMV13 (Fig. 4). Reactivity to anti-CMV sera from sheep, goat or rabbit was quite distinct for different clones (Table 1). Thus using the antigenic determinants expressed as fusion proteins in *E. coli* sensitive and specific identification of anti-CMV antibodies in human patient sera is possible.

Table 1. Reactivity of antisera to cloned antigenic determinants expressed in *E. coli*

Serum no.	Code	Antigen											
		CMV1	CMV13	CMV23	CMV24	CMV26	CMV	pEXStuI	B1B	C35	C79	CMV13 IgM	CMV IgM
1	GSW 0804	++	++	++	±	+	+++	−	+	+	+	+++	++
2	1804	++	++	++	±	+	+++	−	+	+	+	+++	++
3	2904	+++	++	++	+	+	+++	−	++	+	+	+++	++
4	2705	+++	++	++	±	+	+++	±	++	+	+	+++	++
5	GTW 1002	++	+	−	−	±	±	−	±	−	−	+++	+
6	1103	+++	+++	++	+	++	++	±	+	−	−	++++	++
7	1504	+++	+++	++	+	++	++	−	+	−	−	++++	++
8	2705	++	++	++	±	++	++	−	+	−	−	+++	++
9	GAL 1706	+	++	++	±	+	+++	−	+++	++	++	+++	++
10	0209	±	++	++	±	±	++	−	++	+	+	++	+
11	1410	±	++	++	−	+	++	−	++	+	+	++	±
12	GAS 344	±	±	+	+	+	±	±	+	−	−	±	−
13	366	±	+	±	+	++	++	±	+	−	−	+	++
14	392	±	+	±	+	++	++	±	+	−	−	++	++
15	GCD 348	±	++	−	−	++	++	−	±	−	−	+	−
16	398	±	+	±	±	++	+	±	+	±	±	++	++
17	436	±	++	+	+	+	++	±	++	+	+	++	+
18	450	+	++	+	+	+	++	±	++	+	+	++	+
19	479	+	++	+	+	+	++	±	++	+	+	++	+
20	GCA 251	+	+	+	±	+	±	−	+	−	−	+	±
21	311	++	++	++	−	+	++	−	+	±	±	+++	+
22	323	++	+++	++	±	+	++	−	+	+	+	+++	+
23 human	pos. mix	++	+++	++	+	+	+++	−	++	+	+	n.d.	n.d.
24 human	neg.	−	±	−	−	−	−	−	−	−	−	±	−
25 goat	anti CMV	+	++	+	±	−	+	−	+	−	+	n.d.	n.d.
26 goat	preimm.	−	−	−	−	−	−	−	−	−	−	n.d.	n.d.
27 rabbit	anti CMV	+++	++	−	−	−	+++	−	n.d.	n.d.	n.d.	n.d.	n.d.
28 sheep	anti CMV	−	−	−	−	+	++	−	++	++	+	n.d.	n.d.

Protein lysates of *E. coli* expressing CMV-β-galactosidase fusion protein as well as control antigens [pEXStuI and CMV nuclear antigens (CMV)] were separated by SDS-PAGE, blotted and probed with individual sera (diluted 1 : 500 for Ig and 1 : 100 for IgM) as in Fig. 4

Correlation of antigenic determinants to viral proteins

Cloned ORFs of CMV were correlated with their corresponding viral proteins in two ways: (1) Preparation of anti-fusion protein sera and reaction of determinant specific sera with viral antigens, and (2) DNA sequencing and comparison to known sequences of HCMV or related herpes viruses (for review see [26]). The first method is generally applicable and leads to direct identification of viral proteins. The second approach is very useful when gene sequences have already been identified and will be more and more important as sequence information accumulates [2, 4, 5, 6, 16, 17, 25, 39]. However correlation of an open reading frame to a distinct protein species is not possible from DNA sequence data alone, since many CMV gene products are post-transcriptionally and posttranslationally processed. We therefore prepared antisera to fusion proteins to identify corresponding HCMV encoded proteins.

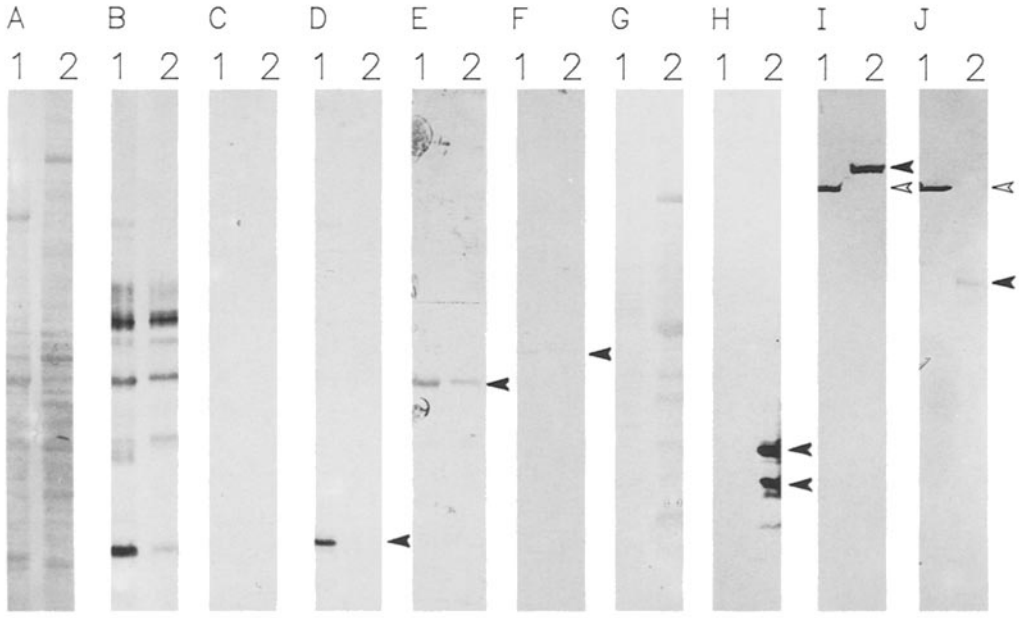

Fig. 5. Identification of CMV proteins corresponding to cloned antigenic determinants. Reactivity of rabbit antisera raised against B1B, C35, and C79 to nuclear (*1*) and NP40-soluble (*2*) CMV antigens is demonstrated in **D**, **E**, and **F**, respectively. Parallel immunoblots with human CMV-positive serum and rabbit anti-B1B preimmune serum are shown in **B** and **C**. For comparison parallel lanes of the same filter were stained with Ponceau S (**A**). **H** Rabbit anti-CMV26 serum was analyzed on immunoblots of NP40-soluble (*1*) and nuclear (*2*) CMV antigens from a different gel. PonceauS stained proteins are shown in **G**. **I**, **J** Immunoblots of marker proteins (myosin, β-galactosidase, phoshorylase B, bovine serum albumin, ovalbumin, and carboanhydrase; (*1*) and virion proteins (*2*) were probed with mouse anti-CMV13 serum (**I**) and mouse anti-CMV24 serum (**J**). Specific binding of antibodies was detected using HRPO-labelled secondary antibodies and H_2O_2/chloronaphtol staining. ▶ Position of p38 (D), p65 (E), p71 (F), p36 and p40 (H), p150 (I), and p58 (J). ▷ Staining of β-galactosidase marker protein by anti-fusion protein sera

```
CMV1      IRRPAGGGGG GGGGQQQRHA AFSLVSPQVT KASPGRVRRD SAWDVRPLTE TRGDLFSGDE DSDSSDGYPP NRQDPRFTDT LVDPPP..

CMV2      IRRPAGGGGG GGGGATPGEP LKDALGRQVS LRSYDNIPPT SSSDEGEDDD DGEDDDNEEP PPP..

CMV13     ARGSIEGRLT PTPVNPSTAP APAPTPTFAG TQTPVNGNSP WAPTAPLPGD MNPANWPRER AWALKNPHLA YNPFRMPTPG SVDLQPSLLID*

CMV21/23  DEEDTSIYLS PPPVPPVQVV AKRLPRPDTP RTPRQKKISQ RPPTPGD

CMV24     ARGSIEGRLA NGFTTPRIGA NQFLRRTVNA QTNPWQNGTD SLDGQTGTQD KGQKPNLLDR LRHRKNGYRH LKDSDEEENV*

CMV26     ARGSIEGRGK SRGGGGGGGS LSSLANAGGL HDDGPGLDND LMNEPMGLGG LGGGGGGGGK KHLDPSTCSQ AC*

C35       ARGSIEGRGL LCPKSIPGLS ISGNLLMNGQ QIFLEVQAIR ETVELRQYDP VAALFFFDID LLLQRGPQYS EHPTFTSQYR IQGKLEYRHT WDRHDEGAAQG
          DDDVWTSGSD SDEELVTTER KTPRVTGGGA MAGASTSAGR KRKSASSATA CTSGVMTRGR LKAESTVAPE EDTDEDSDNE ILDPSTCSQA C*

C79       ARGSIEGRIF FPENIPGVSI EAGPLPDRVR ITLRVTLTGD QAVHLEHRQP LGRIHFFRRG FWTLTPGKPD KIKRPQVQLR AGLFPRSNVM RGAVSEFLPQS
          PGLPPTEEEE EEEEEDDEDD LSSTPTPTPL SEAMFAGFEE ASGDEDSDTQ AGLSPALILT GQRRRSGNNG ALTLVIPSPG SVDLQPSLLI D*

B1B       ARGSIEGRDP HYPGWGRRYE PAPSLHPSYP VPPPPSPAYY RRRDSPGGMD EPPSGWERYD GGHRGQSQKQ HRHGGSGGHN KRRKETAAAS SSSSDEDLSFP
          GEAEHGRARK RLKSHVNSDG GSGGHAGSNQ QQQQRYDELR DAIHELKRDL FAARQSSTLL SAALPSAASS SPTTTTVCTP TGELTSGGGE TPTALLSGGAK
          VAERAQAGVV NASCRLATAS GSEAATAGPS TAGSSSCPAS VVLAAAAAQA AAASQSPPKD MVDLNRRIFV AALNKLE*
```

Fig. 6. Amino acid sequences of CMV antigenic determinants. Amino acid sequences were translated from nucleic acid sequence data. Sequences derived from pEX-vectors are printed in bold, CMV24 sequences derived from pBR322 are underlined. CMV1 and CMV13 correspond to amino acids 486 to 554 and 594 to 663 of pp150 [13]. CMV2 is derived from pp28 (amino acids 16 to 59 [27]). CMV24 contains sequences derived from pBR322 (nucleotides 1294 to 1380) fused in frame to amino acids 862 to 906 of gB [4]. C35 and C79 correspond to amino acids 303 to 476 and 317 to 487 of pp65 and pp71, respectively [33]. B1B is derived from p38 and corresponds to putative amino acids 103 to 373 of this protein (A. Necker and W. Lindenmaier, unpublished data). One letter code is used for amino acids (A-Ala, C-Cys, D-Asp, E-Glu, F-Phe, G-Gly, H-His, I-Ile, K-Lys, L-Leu, M-Met, N-Asn, P-Pro, Q-Gln, R-Arg, S-Ser, T-Thr, V-Val, W-Trp, Y-Tyr)

Fusion proteins were partially purified as described above and used to immunize mice or rabbits. Protein blots of nuclear and cytoplasmic viral antigens, or virion proteins were tested (Fig. 5). Antisera against CMV1 and CMV13 recognized a protein of 150 k. A protein of this size is considered the most antigenic one according to reactivity with human patient sera [13, 18, 34]. Antisera to CMV24 recognized a viral protein of about 58 k, corresponding to glycoprotein B [4]. Antisera to CMV26 recognized multiple protein bands of variable intensity depending on the antigen preparation used. In nuclear extracts of CMV infected fibroblasts proteins with a molecular mass of 36.000 and 40.000 were stained very intensely (Fig. 5H). Antiserum against C35 reacted with a 65 k protein, anti-C79 serum with a 71 k protein. These two proteins were as expected from mapping data of Nowak et al. [30] and represent the 65 k "major late antigen" and the 71 k phosphoproteins [33]. However, antisera to B1B, a third clone derived from the same genomic region, recognized a 38 k protein (Fig. 5; A. Necker and W. Lindenmaier, unpublished data). No specific reaction to the blotted viral antigens was detected using antisera raised against fusion proteins CMV2, 21, or 23, all of which were derived from HindIII fragment R. However, a monoclonal antibody to pp28 [31] recognized the antigenic determinant of CMV2 (data not shown).

DNA sequencing was employed to determine sequence, reading frame, and direction of transcription and translation of the cloned antigenic determinants.

Direct plasmid sequencing [11] was used for pEXStuI clones. Using pEX-specific primers direct, unambiguous identification of the reading frame was possible. For oligo (dG/dC) tailed clones the method of Maxam and Gilbert [24] was used because dideoxy-sequencing from vector specific primers was hindered by the oligo(dG) and olig(dC) tracts. For C35 and C79 which were derived from the pp65 and pp71 genes only start and end sequences of the cloned fragments were determined and compared to the published sequence [33]. Amino acid sequences corresponding to CMV inserts of clones CMV1, 2, 13, 21/23, 24, 26, C35, C79, and B1B are shown in Fig. 6. Sequence comparison to CMV sequences published while this work was in progress revealed homologies to six of these clones. CMV1 and CMV13 are derived from the socalled 150 k "basic phosphoprotein" [13, 34] and correspond to amino acids 486–554 and 594–663 of the sequence of Jahn et al. [13]. They represent at least two independent antigenic epitopes, the first of which contains many charged amino acids and is predicted by an antigenicity prediction program [12, 13]. The second one (CMV13), however, which is the most reactive of our determinants is missed by these programs. The most striking property of this antigenic determinant is its high proline content (Fig. 6). This is also a feature of the common sequence of CMV21 and CMV23, although in this case the sequence is rather hydrophilic. CMV2 represents a very hydrophilic negatively charged part of pp28 [27]. CMV24 contains an insert of in frame pBR322- and CMV-sequences, the latter representing the carboxy-terminal region of glycoprotein gB, amino acids 862–906 [4]. The cloned segment contains a peptide region belonging to the cytoplasmic domain of the protein. It is only weakly recognized by human sera (Table 1). Clones C35, C79, and B1B were isolated from subgenomic libraries of pCM1007 and pBW4. All antigenic determinants of pp65 and pp71 proteins recognized by the human immune system which we had cloned were localized in the C-terminal regions of the corresponding proteins (A. Necker and W. Lindenmaier, unpublished results).

Discussion

Our approach of ORF cloning and expression in *E. coli* allowed isolation and mapping of HCMV antigenic determinants from five regions of the genome corresponding to at least seven different proteins (pp150, pp71, pp65, gB, p36/40, p38, and pp28). In addition to the identification of antigenic regions recognized by the human immune system within already mapped CMV proteins two genes for CMV proteins of 38 k (B1B) and ca. 36/40 k (CMV26) could be localized. Antisera to B1B fusion protein stained a 38 k CMV protein. 3′ sequences of the B1B insert overlapped the 3′ flanking sequences of pp71, therefore direction of transcription/translation is opposite to that found for the 65 k/71 k-proteins. p38 is encoded by the 3′half of a reading frame which has coding capacity for a 78 k protein (A. Necker and W. Lindenmaier, unpublished data) and represents the HCMV homologue of the recently described "assembly

protein" of Colburn CMV [8, 32]. Antibodies to CMV26 recognized predominantly proteins of 36 k and 40 k in immunoblots of nuclear antigens. The gene was mapped within HindIII-fragment M. A gene for a protein family which includes the 52 k DNA binding protein has been mapped to the same genomic fragment [28], but the protein pattern recognized by our anti-CMV26 serum is different. Whether p36/40 could be related to minor coprecipitated components of this family is not clear.

The statistical genomic cloning approach is not biased to specific genes expressed in large quantities during defined times in the infectious cycle as is the cDNA cloning approach. In addition, interference with residual or induced cellular RNAs is avoided. Screening with human sera enables direct isolation of antigenic determinants relevant to human humoral immune response avoiding problems due to the unreliability of antigenicity prediction programs. Experiments using animal sera showed that these sera will probably miss some of the antigenic determinants important in the human immune response. However, this random cloning and screening approach, too, is restricted to determinants which can be stably maintained in *E. coli* and to continuous epitopes which may represent only a minority of the relevant B-cell epitopes. Formation of discontinuous epitopes which would require correct three dimensional folding of the CMV-part of the fusion protein is rather improbable. Also epitopes which require posttranslational modifications, e.g., glycosylation, could not be detected. However, fusion proteins can be used to provide determinant-specific antisera and antigens for diagnostic purposes [13, 21, 34, and unpublished results]. Fusion proteins which can easily and reproducibly be produced allow testing of the immune response to individual determinants. This could be of diagnostic value because reactivity to individual CMV proteins (e.g., 150 k and intermediate molecular weight proteins) has been reported to be more significant for active CMV infection than the general anti-CMV response measured in conventional ELISA-assays [18]. Fusion proteins and corresponding synthetic peptides will allow accurate determination of antigenic epitopes. A more extensive analysis of antigenic determinants and testing of the biological significance of individual epitopes – including the ones responsible for cellular immune response – may eventually lead to a rational design of a CMV subunit vaccine.

Acknowledgments

We gratefully acknowledge the generous gift of CMV antigens and antisera by Dr. B. Fleckenstein, Erlangen, Dr. J. Middeldorp, Oss, Dr. T. H. The, Groningen, and Dr. M. P. Landini. We are grateful to Drs. K. Stanley and H. Mayer for the gift of pEX1 and pEXStuI, respectively. We thank M. Fitzek and Dr. König for help with the immunization and B. Seeger-Kunth for carefully typing the manuscript.

References

1. Alford CA, Britt WJ (1985) Cytomegalovirus. In: Fields BN (ed) Virology. Raven Press, New York, pp 629–657

2. Baer R, Bankier AT, Biggin MD, Deininger PL, Farrell PJ, Gibson TJ, Hatfull G, Hudson GS, Satchwell SC, Seguin C, Tuffnell PS, Barrell BG (1984) DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature 310: 207–211
3. Bernhard HU, Helsinki DR (1980) Bacterial plasmid cloning vehicles. In: Setlow JK, Holländer A (eds) Genetic engineering: principles and methods. Plenum, New York, pp 133–168
4. Cranage MP, Kouzarides T, Bankier AT, Satchwell S, Weston K, Tomlinson P, Barrell B, Hart H, Bell SE, Minson AC, Smith GL (1986) Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. EMBO J 5: 3057–3063
5. Cranage MP, Smith GL, Bell SE, Hart H, Brown CA, Bankier T, Tomlinson P, Barrell BG, Minson TC (1988) Identification and expression of a human cytomegalovirus glycoprotein with homology to the Epstein-Barr virus BXLF2 product, varicella-zoster virus gpIII, and herpes simplex virus type 1 glycoprotein H. J Virol 62: 1416–1422
6. Davison AJ, Scott JE (1986) The complete DNA sequence of varicella-zoster virus. J Gen Virol 67: 1759–1816
7. Fleckenstein B, Müller I, Collins J (1982) Cloning of the complete human cytomegalovirus genome in cosmids. Gene 18: 39–46
8. Gibson W (1983) Protein counterparts of human and simian cytomegaloviruses. Virology 128: 391–406
9. Gottesman ME, Adhya S, Das A (1980) Transcription antitermination by bacteriophage lambda N gene product. J Mol Biol 140: 57–75
10. Gubler U, Hoffmann BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25: 263–269
11. Hattori M, Sakaki Y (1986) Dideoxy sequencing method using denatured templates. Anal Biochem 152: 232–238
12. Hopp TP, Woods KR (1981) Prediction of antigenic determinants from amino acid sequences. Proc Natl Acad Sci USA 78: 3823–3828
13. Jahn G, Kouzarides T, Mach M, Scholl BC, Plachter B, Traupe B, Preddie E, Satchwell SC, Fleckenstein B, Barrell BG (1987) Map position and nucleotide sequence of the gene for the large structural phosphoprotein of human cytomegalovirus. J Virol 61: 1358–1367
14. Kemble GW, McCormick AL, Pereira L, Mocarski ES (1987) A cytomegalovirus protein with properties of herpes simplex virus ICP8: partial purification of the polypeptide and map position of the gene. J Virol 61: 3143–3151
15. Koenen M, Rüther U, Müller-Hill B (1982) Immunoenzymatic detection of expressed gene fragments cloned in the lac Z gene of *E. coli*. EMBO J 4: 509–512
16. Kouzarides T, Bankier AT, Satchwell SC, Weston K, Tomlinson P, Barrell BG (1987) Large scale rearrangement of homologous regions in the genomes of HCMV and EBV. Virology 157: 397–413
17. Kouzarides T, Bankier AT, Satchwell SC, Weston K, Tomlinson P, Barrell BG (1987) Sequence and transcription analysis of the human cytomegalovirus DNA polymerase gene. J Virol 61: 125—133
18. Landini MP, Mirolo G, Baldassari B, La Placa M (1985) Human immune response to cytomegaloviral structural polypeptides studied by immuno blotting. J Med Virol 17: 303–311
19. Lindenmaier W, Dittmar KEJ, Hauser H, Necker A, Sebald W (1985) Isolation of a functional human interleukin 2 gene from a cosmid libary by recombination in vivo. Gene 31: 33–39
20. Loon AM van, Van der Logt JTM, Van der Vee J (1981) Enzyme linked immunosorbent assay for measurement of antibodies against cytomegalovirus and rubella virus at a single serum dilution. J Clin Pathol 4: 51–60

21. Mach M, Utz U, Fleckenstein B (1986) Mapping of the major glycoprotein of human cytomegalovirus. J Gen Virol 67: 1461–1467
22. Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory Cold Spring Harbor, NY
23. Martinez J, St Jeor SC (1986) Molecular cloning and analysis of three cDNA clones homologous to human cytomegalovirus RNAs present during late infection. J Virol 60: 531–538
24. Maxam A, Gilbert W (1980) Sequencing end labelled DNA with base specific chemical cleavages. In: Grossman L, Moldave K (eds) Methods of enzymology, vol 65. Academic Press, New York, pp 499–560
25. McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, McNab D, Perry LJ, Scott JE, Taylor P (1988) The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J Gen Virol 69: 1531–1574
26. McGeoch DJ (1989) The genomes of the human herpesviruses: contents, relationships, and evolution. Annu Rev Microbiol 43: 235–265
27. Meyer H, Bankier AT, Landini MP, Brown CM, Barrell BG, Rüger B, Mach M (1988) Identification and procaryotic expression of the gene coding for the highly immunogenic 28-kilodalton structural phosphoprotein (pp28) of human cytomegalovirus. J Virol 62: 2243–2250
28. Mocarski ES, Pereira L, Michael N (1985) Precise localization of genes on large animal virus genome use of λgt II and monoclonal antibodies to map the gene for a cytomegalovirus protein family. Proc Natl Acad Sci USA 82: 1266–1270
29. Nagai K, Thøgersen HC (1984) Generation of globin by sequence specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature 309: 810–812
30. Nowak B, Gmeiner A, Sarnow P, Levine AJ, Fleckenstein B (1984) Physical mapping of human cytomegalovirus genes: identification of DNA sequences coding for a virion phosphoprotein of 71-kDa and a viral 65-kDa polypeptide. Virology 134: 91–102
31. Re MC, Landini MP, Coppolecchia P, Furlini G, La Placa M (1985) A 28.000 molecular weight human cytomegalovirus structural polypeptide studied by means of a specific monoclonal antibody. J Gen Virol 66: 2507–3511
32. Robson L, Gibson W (1989) Primate cytomegalovirus assembly protein: genome localization and nucleotide sequence. J Virol 63: 669–676
33. Rüger B, Klages S, Walla B, Albrecht J, Fleckenstein B, Tomlinson P, Barrell B (1987) Primary structure and transcription of the genes coding for the two virion phosphoproteins pp65 and pp71 of human cytomegalovirus. J Virol 61: 446–453
34. Scholl BC, von Hintzenstern J, Borisch B, Traupe B, Bröker M, Jahn G (1988) Prokaryotic expression of immunogenic polypeptides of the large phosphoprotein (pp150) of human cytomegalovirus. J Gen Virol 69: 1195–1204
35. Stanley KK, Luzio JP (1984) Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J 3: 1429–1434
36. Stinski MF (1977) Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J Virol 23: 751–767
37. Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA 76: 4350–4354
38. Verweij CL, de Vries CJM, Distel B, van Zonneveld AJ, van Kessel AG, van Mourik JA, Pannekoek H (1985) Construction of a cDNA coding for human von Willebrand factor using antibody probes for colony-screening and mapping of the chromosomal gene. Nucleic Acids Res 13: 4699–4717
40. Weston K, Barrell BG (1986) Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. J Mol Biol 192: 177–208

41. Wingender E, Bercz G, Blöcker H, Frank R, Mayer H (1989) Expression of human parathyroid hormone in Escherichia coli. J Biol Chem 264: 4367–4373

Authors' address: Dr. W. Lindenmaier, GBF, Mascheroder Weg 1, D-3300 Braunschweig, Federal Republic of Germany.

Received January 31, 1990