Two isoforms of the protein kinase pUL97 of human cytomegalovirus are differentially regulated in their nuclear translocation

Rike Webel,1 Jens Milbradt,1 Sabrina Auerochs,1 Vera Schregel,1† Christian Held,2 Katharina Nöbauer,3 Ebrahim Razzazi-Fazeli,3 Christophe Jardin,4 Thomas Wittenberg,2 Heinrich Sticht4 and Manfred Marschall1

The pUL97 protein kinase encoded by human cytomegalovirus is a multifunctional determinant of the efficiency of viral replication and phosphorylates viral as well as cellular substrate proteins. Here, we report that pUL97 is expressed in two isoforms with molecular masses of approximately 90 and 100 kDa. ORF UL97 comprises an unusual coding strategy in that five in-frame ATG start codons are contained within the N-terminal 157 aa. Site-directed mutagenesis, transient expression of point and deletion mutants and proteomic analyses accumulated evidence that the formation of the large and small isoforms result from alternative initiation of translation, with the start points being at amino acids 1 and 74, respectively. In vitro kinase assays demonstrated that catalytic activity, in terms of autophosphorylation and histone substrate phosphorylation, was indistinguishable for the two isoforms. An analysis of the intracellular distribution of pUL97 by confocal laser-scanning microscopy demonstrated that both isoforms have a pronounced nuclear localization. Surprisingly, mapping experiments performed to identify the nuclear localization signal (NLS) of pUL97 strongly suggest that the mechanism of nuclear transport is distinct for the two isoforms. While the extreme N terminus (large isoform) comprises a highly efficient, bipartite NLS (amino acids 6–35), a second sequence apparently conferring a less efficient mode of nuclear translocation was identified downstream of amino acid 74 (small and large isoforms). Taken together, the findings argue for a complex mechanism of nuclear translocation for pUL97 which might be linked with fine-regulatory differences between the two isoforms.

INTRODUCTION

Human cytomegalovirus (HCMV) is a clinically important, ubiquitous human pathogen (family Herpesviridae) that causes severe systemic diseases in immunosuppressed patients and prenatally infected children. HCMV-associated pathogenicity correlates with the efficiency of viral replication in permissive tissues (Mocarski et al., 2007; Sinzger et al., 2008). HCMV replication is restricted to specific host-cell types and is dependent on the balance of interactions between viral and cellular proteins. In this regard, various virus- or cell-encoded protein kinases display crucial functions during HCMV replication and are interlinked in several aspects of their regulatory properties (Prichard, 2009; Lee & Chen, 2010). The serine/threonine-type protein kinase pUL97 is an important determinant of efficient HCMV replication. Deletion of ORF UL97 from the viral genome, or pharmacological inhibition of the kinase activity, results in a rigorous reduction of virus replication efficiency by a factor of 100–1000 (Prichard et al., 1999; Wolf et al., 2001; Marschall et al., 2002; Biron et al., 2002; Herget et al., 2004). pUL97 exerts its influence on HCMV replication by phosphorylation of both viral and
cellular proteins, such as pUL44 (Marschall et al., 2003; Krosky et al., 2003), pUL69 (Thomas et al., 2009), pp65 (Becke et al., 2010), nuclear lamins (Marschall et al., 2005; Hamirally et al., 2009; Milbrandt et al., 2010), p32 (Marschall et al., 2005), Rb protein (Hume et al., 2008; Prichard et al., 2008), RNA polymerase II (Baek et al., 2004) and EF-16 (Kawaguchi et al., 1999; Romaker et al., 2006). Furthermore, pUL97 possesses a strong autophosphorylation activity, which is markedly influenced by the formation of pUL97 homodimers or oligomers via its self-interaction domain (Schregel et al. 2007).

A number of publications have demonstrated that pUL97 accumulates in the nucleus of HCMV-infected or transiently transfected cells (Michel et al., 1996; van Zeijl et al., 1997; Marschall et al., 2003; 2005; Prichard et al., 2005; Romaker et al., 2006; Milbrandt et al., 2009; Rechter et al., 2009). Using immunofluorescence analysis of pUL97–GFP fusion constructs, it was postulated that a nuclear localization signal (NLS) exists in the N terminus of pUL97, located between amino acids 48 and 110 (Michel et al., 1998). However, the exact sequence of the NLS has not been identified to date. In general, nuclear import of NLS-bearing proteins can be realized through several independent transport mechanisms. The conventional importin α/β pathway is described for classical NLS sequences, which are mostly composed of consecutive basic residues within a short stretch of amino acids. Importin α recognizes two specific binding domains of the classical NLS and attaches to importin β which assists interaction with the nuclear pore complex, thereby leading to nuclear translocation (Görlich et al., 1994, 1996, 1996; Moroianu et al., 1995; Conti et al., 1998; Herold et al., 1998). Some proteins carry an arginine-rich NLS and exclusively require importin β for nuclear transportation (Truant & Cullen, 1999). Alternatively, nuclear localization can be achieved by a complex non-conventional NLS that interacts with importin α, e.g. the 282 aa of the HCMV multifunctional regulator pUL84 essential for nuclear transport (Lischka et al., 2003). Nuclear import can also be determined by multiple NLS sequences, as described for the herpes simplex virus type 1 (HSV-1) thymidine kinase (Degrèv et al., 1998, 1999). Characterization of site-directed and deletion mutants of HSV-1, combined with crystallographic structure analyses, revealed a nonapeptide in the first 34 aa of the protein and two important domains (amino acids 236–237 and 317–318) that are required for nuclear transport (Degrèv et al., 1998, 1999). Interestingly, three isoforms of this thymidine kinase are generated by the use of three different in-frame ATG start codons for translational initiation. All three isoforms show nuclear localization and possess kinase activity (Marsden et al., 1983; Haarr et al., 1985). However, combined experimental data suggest that only the full-length isoform of HSV-1 thymidine kinase may comprise the complete N-terminal NLS sequence.

This situation is reminiscent of the coding strategy for the HCMV protein kinase pUL97. ORF UL97 contains five in-frame ATG start codons within the first 157 aa (Fig. 1a), which raised a question as to whether more than one start codon is used in pUL97 synthesis. This point is in context with the early finding that pUL97 was detectable as a characteristic double band at 90/100 kDa (Marschall et al., 2003). More recently, it was shown in detail that N-terminally intact recombinant versions of pUL97 were consistently observed as two different isoforms (Schregel et al., 2007). For this study, we focused on investigating the two pUL97 isoforms, their catalytic activity and the regulation of their nuclear translocation. The putative role of the individual isoforms in the multiple functions of pUL97 is discussed.

RESULTS AND DISCUSSION

pUL97 occurs in two isoforms

As previously published, transient expression of pUL97 is characterized by the occurrence of two prominent products of approximately 90 and 100 kDa (Marschall et al., 2003; Schregel et al., 2007). We addressed this issue in detail, particularly by way of the comparative analysis of proteins derived from HCMV-infected cells, infectious virions and transiently transfected cells. In primary human foreskin fibroblasts (HFFs) infected with HCMV strain AD169, pUL97-specific bands became detectable at 2 days post-infection (p.i.) (Fig. 2a, lanes 1–3) and quantitatively increased up to 7 days (Fig. 2a, lanes 4–6). This was comparable with a typical early–late kinetic of expression as reported by Michel et al. (1996). In Fig. 2(a), the two protein bands with molecular masses of 90 and 100 kDa were consistently detectable. When the virus inoculum used for infection and the progeny virus released into the culture medium 12 days p.i. were analysed by Western blotting (Wb), no quantitative difference in the occurrence of the pUL97 double band was observed (Fig. 2a, lanes 7–8). When virion-associated pUL97 was immunoprecipitated and subjected to a pUL97-specific in vitro kinase assay, high activity was detected with respect to autophosphorylation for both versions of the protein (Fig. 2a, lane 9). These findings strongly suggest that physiologically produced pUL97 regularly occurs in two isoforms which are both catalytically active.

Molecular mechanism of formation of pUL97 isoforms: alternative initiation of translation

Three possible mechanisms for the formation of pUL97 isoforms were taken into consideration: (i) a splicing mechanism, (ii) a post-translational modification (such as phosphorylation or proteolytic processing) and (iii) alternative translational initiation. First, to investigate a splicing-based system of regulation of gene expression, we tried to detect alternative UL97-specific RNAs for several positions along the coding region. However, reverse transcription PCR (RT-PCR; Reverse Transcription System, Promega) of UL97 RNAs produced in HCMV-infected or
UL97 plasmid-transfected cells did not provide any evidence for alternative splicing products. Instead, all of the RT-PCR products obtained originated from unspliced transcripts of the UL97 region (not shown), a finding which is consistent with an earlier analysis of this transcriptional unit (Wing & Huang, 1995).

Secondly, the possibility of post-translational modification by phosphorylation (i.e. the generation of a large hyperphosphorylated isoform from a smaller hypophosphorylated precursor protein) appeared improbable because of the fact that recombinantly expressed pUL97 (detected by Wb) and the in vitro autophosphorylation products of pUL97 (detected by in vitro kinase assay) were indistinguishable from each other. On the basis of SDS-PAGE mobility, the 90/100 kDa double-band was detected in both cases (Figs 2a and 3; Schregel et al., 2007; Thomas et al., 2009). Importantly, the catalytically inactive K355M mutant of pUL97 (which was found to be devoid of autophosphorylation activity) was also expressed as a double band. Also the two protein species of pUL97(K355M)–Flag both showed a slightly faster mobility in SDS-PAGE than their wild-type counterparts (Schregel et al., 2007). This indicated on the one hand that the faster mobility (i.e. lower molecular mass) was actually because of a lack of autophosphorylation and on the other hand that autophosphorylation was not the mechanism of isoform generation (Fig. 2c, lanes 1 and 5). It also seemed unlikely that post-translational proteolytic cleavage was a basis of isoform generation. This was because of the finding that fusion constructs, containing various portions of the entire pUL97 (Fig. 4a-d), did not show any sign of specific cleavage when analysed by Wb (not shown).

Thirdly, the possibility of an alternative mode of translational initiation was experimentally addressed. For this purpose, specific point and deletion mutants of pUL97 were produced. Wb analysis of the transient expression products revealed that wild-type pUL97 occurred as the typical double band while mutation of the first start codon of translation (M1L) abrogated the formation of the large isoform and exclusively allowed expression of the small isoform (Fig. 2b, c, lanes 1 and 6). Interestingly, deletion of the N-terminal 110, 156 or 180 aa of pUL97 led to the formation of slightly smaller products than the small isoform expressed by the M1L mutant (Fig. 2b, lanes 4–6; Fig. 2c, lanes 3, 4 and 6), thus suggesting that the small...
isoform is initiated upstream of M111. In addition, expression of deletion mutant 38–707 (Fig. 2b, lane 2) produced two pUL97 bands, one of these apparently identical to the small isoform, the second with a molecular mass consistent with translational initiation at M38. This finding argues that the small isoform is initiated downstream of M38, most probably at M74 (Fig. 1a). Importantly, expression of mutant 74–707 resulted in a protein with SDS-PAGE mobility identical to the M1L product (Fig. 2b, lanes 3 and 6), supporting the idea of M74 initiation for the small isoform. In this context it should be noted that sequence analyses demonstrated a conservation of M74 among human cytomegaloviruses, but not among simian and other cytomegaloviruses (Supplementary Fig. S1, available in JGV Online; Romaker et al., 2006). Interestingly, in the case of rat and mouse cytomegaloviruses only single protein bands of the pUL97 homologues were identified (pR97 and pM97, respectively), suggesting that a small isoform may not be produced (Romaker et al., 2006; M. Marschall, unpublished).

Fig. 2. Isoforms of pUL97 in HCMV-infected primary fibroblasts, HCMV virions and transiently transfected cells. (a) HCMV-infected HFFs were used for Wb analysis. As controls, virus inoculum (lane 8) and released progeny virus (lane 7) were harvested by centrifugation and assayed in parallel to visualize virion-derived pUL97. Staining with mAb-β-actin was performed to control for loading (lanes 1–6) and the absence of cell debris in virion preparations (lanes 7–8). The kinase activity of pUL97 derived from virions was determined in parallel (lane 9, sample identical to lane 8). d, Day p.i.; IVKA, in vitro kinase assay. (b, c) Plasmid-transfected 293T cells were used for Wb analysis as indicated. (d) N terminus of pUL97 (amino acids 1–16) was detected using pAb-PepAs1–16 (lanes 1–3); expression control staining with mAb-Flag (lanes 4–6). *, Cross-reaction with a cellular protein (loading control). F, Flag tag.

Fig. 3. Autophosphorylation and substrate phosphorylation of the two pUL97 isoforms. (a) 293T cells were transfected with the constructs indicated and lysates were used for in vitro kinase assays (IVKA) to detect histone phosphorylation (phos.) and autophosphorylation (autophos.). The ratio of concentrations of the large and small isoforms was achieved by using varying concentrations of the constructs pUL97(M38/47/111/157)–F and pUL97(M1L)–F as indicated and was verified by an expression control Wb (b). F, Flag tag; RFP, red fluorescent protein.

isoform is initiated upstream of M111. In addition, expression of deletion mutant 38–707 (Fig. 2b, lane 2) produced two pUL97 bands, one of these apparently identical to the small isoform, the second with a molecular mass consistent with translational initiation at M38. This finding argues that the small isoform is initiated downstream of M38, most probably at M74 (Fig. 1a). Importantly, expression of mutant 74–707 resulted in a protein with SDS-PAGE mobility identical to the M1L product (Fig. 2b, lanes 3 and 6), supporting the idea of M74 initiation for the small isoform. In this context it should be noted that sequence analyses demonstrated a conservation of M74 among human cytomegaloviruses, but not among simian and other cytomegaloviruses (Supplementary Fig. S1, available in JGV Online; Romaker et al., 2006). Interestingly, in the case of rat and mouse cytomegaloviruses only single protein bands of the pUL97 homologues were identified (pR97 and pM97, respectively), suggesting that a small isoform may not be produced (Romaker et al., 2006; M. Marschall, unpublished).
observations). In the case of pUL97, another intriguing finding was the expression pattern of a fourfold missense mutant, M38/74/111/157, which exclusively yielded the large isoform (Fig. 2b, lane 7). This strongly suggested that it was none of these four start codons that was responsible for translational initiation of the large isoform of pUL97, but rather that it was M1 specifically (see above). This hypothesis was confirmed by the finding that a peptide-specific antibody (amino acids 1–16) recognized the large, but not the small, isoform of pUL97 (Fig. 2d). Taken together, these experiments point to the importance of M1 and M74 for the generation of the two pUL97 isoforms. This idea was strengthened further by the finding that only the large isoform was detectable upon transfection of an N-terminally tagged expression construct by the use of a tag-specific antibody (Fig. 2c, lane 2), pointing to the exclusive presence of a complete N terminus in the large isoform, but the absence of this tagged N terminus in the small isoform.

Fig. 4. Identification of NLS of pUL97 using an established NLS mapping system. HeLa cells were transfected with the constructs indicated and analysed by confocal laser-scanning microscopy. (a) Five NLS mapping constructs covering the entire ORF UL97 were analysed for intracellular localization [comparison with parental construct GFP–β-gal (pHM830) and an empty vector control (pcDNA3.1)]. (b) Fine-mapping of the N-terminal region 1–110 was performed by the use of additional constructs as indicated. (c) Further examination of the constructs 111–180 and 111–250 with regard to intermediate phenotypes of localization. (d) Constructs representing fusions of β-gal (in absence of GFP) with the large or small isoforms of pUL97 were analysed in parallel by immunostaining with pAb-UL97.
Proteomic analyses were performed to confirm the identity of the N terminus of the small isoform of pUL97. For this purpose, the expression construct for pUL97(M1L)–Flag was used to purify recombinant Flag-tagged protein corresponding to the small isoform. This protein was subjected to tryptic digestion and analysed by mass spectrometry (ISD MALDI–TOF/TOF) leading to the identification of 33 pUL97-specific peptides (Supplementary Fig. S2, available in JGV Online). The absence of scored peptides in the N-terminal region (amino acids 1–107) is consistent with initiation of pUL97 expression at M74. Attempts to sequence the small isoform of pUL97 using N-terminal Edman microsequencing were unsuccessful, indicating that the N-terminal amino acid is blocked by acetylation or another modification.

**In vitro autophosphorylation and substrate phosphorylation is similarly detectable for both pUL97 isoforms**

The small and large isoforms of pUL97 were transiently expressed by the use of plasmids coding for pUL97(M1L)–Flag and pUL97(M38/74/111/157)–Flag. Co-expression of the two proteins was achieved by co-transfection of a variety of concentrations of the two plasmids, and expression levels were monitored by Western (Wb) analysis (Fig. 3b). The determination of the *in vitro* kinase activity of immunoprecipitated Flag-tagged proteins revealed that the activity was similarly detectable for the large and small isoforms of pUL97 (Fig. 3a). Autophosphorylation and substrate phosphorylation (histones) were quantitatively and qualitatively indistinguishable between the two isoforms (Fig. 3a, lanes 2 and 8) and were not different from wild-type pUL97–Flag that was used as a positive control (Fig. 3a, lane 1). Also, variation in the relative amounts of the two proteins did not result in detectably different phosphorylation characteristics (Fig. 3a, lanes 3–7). This indicates that the two pUL97 isoforms do not appear to be distinct from each other in terms of their basic catalytic activity. However, these *in vitro* kinase experiments cannot exclude the possibility that a specific physiological substrate of pUL97, either viral or cellular, is preferentially phosphorylated by one of the two isoforms.

**Nuclear localization is detectable for wild-type pUL97 and the two individually expressed large and small isoforms of pUL97**

The intracellular localization of pUL97 has been analysed in detail by previous studies (Michel et al., 1996, 1998; Wolf et al., 1998; Marschall et al., 2003, 2005; Prichard et al., 2005; Romaker et al., 2006; Hamirally et al., 2009; Milbradt et al., 2009, 2010). A predominantly nuclear and perinuclear localization of pUL97 occurs in HCMV-infected cells (Michel et al., 1996; Marschall et al., 2003), with concentrations of the protein kinase occurring at the nuclear rim during the late phase of infection (Milbradt et al., 2010; Marschall et al., 2005; M. Marschall, unpublished data). Additional cytoplasmic functions of pUL97 (Kawaguchi et al., 1999; Prichard et al., 2005; Azzeh et al., 2006; Romaker et al., 2006) suggest a complex regulation system of pUL97 intracellular transport. Here, we investigated transiently expressed wild-type pUL97 in comparison to the two individually expressed large and small isoforms of pUL97. Confocal laser-scanning microscopy showed that pUL97 was predominantly localized in the nucleus and that there was no detectable difference in the localization of the large and small isoforms (Supplementary Fig. S3, available in JGV Online). Against this background, experiments were performed to identify the exact location of the NLS responsible.

**Identification of novel NLS sequences within pUL97**

We utilized an established assay system to identify NLS elements within pUL97 (Sorg & Stamminger, 1999). For this purpose, five fragments covering the entire ORF UL97 were separately inserted into a GFP–β-galactosidase (β-gal) fusion cassette and transiently transfected into HeLa cells to determine the functionality of the inserted fragments as effective nuclear translocation elements. While the original GFP–β-gal carrier protein showed a clear cytoplasmic localization pattern (Fig. 4a, panels 4–6), the insertion of pUL97 amino acids 1–110 effected a nuclear translocation (Fig. 4a, panels 7–9). This indicated the presence of a typical NLS within fragment 1–110. The three fragments spanning regions between amino acids 181 and 707 did not have any effect on the cytoplasmic localization of the carrier protein, indicating that none of these fragments contained a functional NLS (Fig. 4a, panels 13–21). Interestingly, the remaining pUL97 fragment containing amino acids 111–180 displayed an intermediate phenotype. While for most of the cells cytoplasmic localization of the construct was detected (Fig. 4a, panels 10–12), a minor fraction of cells (<5%) showed nucleo-cytoplasmic localization (Fig. 4c, panels 1–3). This intermediate phenotype was even more evident for the construct containing pUL97 fragment 111–250, which showed nucleo-cytoplasmic localization in more than 50% of cells (Fig. 4c, panels 4–6).

These differences in intracellular localization were interpreted with regard to a computational analysis indicating the presence of three candidate NLS sequences in pUL97. The predicted NLS sequences all belong to the bipartite NLS type, which is characterized by having two clusters of basic residues separated by a 10–12 aa linker. NLS candidate sequences span residues 6–35, 164–198 and 190–213 of pUL97 (not shown). The first candidate NLS (residues 6–35; Fig. 1b) is evolutionarily conserved among the pUL97 homologues from all human, chimpanzee, rhesus and cercopithecine cytomegaloviruses analysed, while the conservation of the second and third candidates (residues 164–198 and 190–213, respectively) is mainly restricted to HCMV strains (Supplementary Fig. S1; Cunningham et al., 2010). This might indicate that only the first NLS is functional, because of an earlier observation that evolutionary non-conserved sequence
motifs are frequently non-functional (Dinkel & Sticht, 2007).

The nuclear localization of the 1–110 fragment indicated the functionality of the N-terminal NLS (residues 6–35), while the cytoplasmic localization of the 181–365 fragment demonstrated that the predicted NLS from residues 190–213 is non-functional (at least when separated from the overlapping sequence candidate 164–198). In contrast to this, the candidate NLS from residues 164–198 (flanked by the overlapping sequence candidate 190–213; Fig. 1b) appeared to be responsible for the partial nuclear localization of the 111–250 fragment. This idea is consistent with the finding that the NLS function was largely absent from the shorter 111–180 fragment. However, when the exact fragment 164–198 was inserted into a GFP–β-gal fusion construct, the NLS function was still not sufficiently provided (i.e. localization was mostly cytoplasmic; not shown). This finding indicated that certain flanking residues, additional to the predicted minimal NLS of residues 164–198, are functionally required. In particular, it appears probable that the two overlapping NLS candidates, 164–198 and 190–213, are both required to confer efficient nuclear transport. Moreover, it should be stressed that an additional computational prediction was directed to the flanking vector sequences of the NLS mapping constructs and this clarified that these did not show any tendency to affect or mimic NLS sequences. This suggested that the analysis was not biased by the chosen reporter system.

The finding that pUL97 fragment 1–110 contains a classical NLS (Fig. 4a, panels 7–9) is principally compatible with a previously reported NLS in this region (Michel et al., 1998). However, our study strongly suggests that the exact location of this NLS is not within region 48–110 as initially postulated. To narrow down the location of the NLS experimentally, further truncated constructs were generated and analysed. Our data provide evidence that neither region 48–110 nor 1–47 alone is sufficient to confer nuclear translocation (Fig. 4b, panels 7–12). On the other hand, a C-terminal truncation of the 1–110 element to amino acids 1–85 and 1–60 did not destroy NLS activity. Both constructs showed a predominantly nuclear phenotype in the majority of cells (Fig. 4b, panels 1–6). These observations are consistent with the presence of a strong NLS, predicted to span residues 6–35 of pUL97 (Fig. 1b). Further evidence for the location of this NLS was provided by a construct in which the nine N-terminal amino acids were deleted (construct 10–110; Fig. 4b, panels 13–15). In this construct, the first part of the bipartite NLS is missing, resulting in partially reduced nuclear translocation activity. Interestingly, the cytoplasmic localization of pUL97 fragment 1–47 is in apparent contradiction to the location of the NLS C-terminal end at residue 35. However, investigation of the secondary structure of pUL97 offers an explanation for this finding. Secondary structure prediction (Fig. 1b) revealed that the C-terminal part of the NLS belongs to a long α-helix that, approximately, spans residues 31–64. The fact that pUL97 fragment 1–60, but not fragment 1–47, was transported to the nucleus suggested that the proper α-helical structure contributes to full NLS activity, while a broken helix appears to handicap the NLS function. The importance of structural requirements within the N terminus of pUL97, in particular the predicted α-helix between residues 31–64, was further substantiated by experiments with additional reporter constructs carrying fragments 1–55 and 1–65 of pUL97. Construct 1–65 was thought to contain the predicted α-helix, while construct 1–55 was supposed to be impaired in forming this structure. In fact, and as shown in Fig. 5, a significant difference in the efficiency of nuclear import was measured for the two constructs (Student’s t-test). For construct 1–65, a substantially higher portion of cells showing mainly nuclear localization was detected (65.2 %) than for construct 1–55 (28.0 %, Fig. 5a). A detailed quantitative analysis was performed on the basis of a semi-automatic interactive cell segmentation (Fig. 5b). The data confirm the pronounced nuclear localization of construct 1–65 in the majority of cells (class 1, intensity ratio 0–0.59) compared with the predominant nucleo-cytoplasmic staining of construct 1–55 (class 2, intensity ratio 0.59–1.19). Finally, we found that the exact NLS sequence 6–35 (i.e. without the downstream helical region), when inserted into a GFP–β-gal fusion construct, showed almost perfect nuclear localization (Supplementary Fig. S4, available in JGV Online). This strengthened the idea that the absence of any incompletely structured flanking regions, which may be inhibitory rather than supportive for NLS function, resulted in very efficient nuclear transport of construct 6–35. As observed for constructs 1–47 and 1–55, the flanking regions obviously provided a suboptimal efficiency of nuclear transport which may relate to a partial or full masking of the NLS function. The observation of efficient transport of the 1–65 construct suggests that the formation of an α-helix spanning residues 31–65 (Fig. 1b) fixes the flanking residues in a defined conformation, thereby preventing their masking NLS function. Due to the C-terminal truncation, the respective helix cannot be formed in the shorter 1–47 and 1–55 constructs. Taken together, these findings argue for the combination of a classical bipartite NLS and the embedding of this NLS into an intact secondary structure.

The regulation of nuclear translocation is distinct between the two pUL97 isoforms

The findings of Fig. 4(a–c) led to the assumption that the two isoforms of pUL97 may translocate to the nucleus through two different mechanisms, since both are localized to the nucleus but only the large isoform possesses the efficient NLS in position 6–35. To address this question, constructs were generated containing the entire wild-type pUL97 or isoform-specific versions of pUL97, which were N-terminally fused to β-gal (in the absence of an additional GFP fusion in order to limit the overall sizes of the products; Fig. 4d). Very strong and complete nuclear
Fig. 5. Quantification of the efficiency of nuclear transport of NLS mapping constructs carrying N-terminal fragments 1–55 or 1–65 of pUL97. (a) Confocal laser-scanning microscopy was performed with transfected HeLa cells. A qualitative and quantitative evaluation of the efficiency of nuclear transport is presented. (b) Class assignment was performed to quantify the intensity ratio of pUL97 fusion constructs inside and outside of the nucleus. An intensity ratio of $R<1$ indicates localization of the pUL97 construct mainly in the nucleus (class 1), whereas $R \approx 1$ describes a nucleo-cytoplasmic distribution (class 2) and $R>1$ localization mainly to the cytoplasm.
localization of the construct containing pUL97 in its wild-type version (1–707) was observed (Fig. 4d, panels 1–3). A construct exclusively representing the large isoform of pUL97 (M38/74/111/157–707) either displayed complete nuclear localization or combined nucleo-cytoplasmic localization (Fig. 4d, panels 4–6). In contrast, two constructs representing the small isoform of pUL97 (M1L–707 and 74–707) not only showed nucleo/cytoplasmic staining, but also purely cytoplasmic localization in a subpopulation of cells (up to 15% of cells in independent determinations; Fig. 4d, panels 7–12). Such exclusive cytoplasmic localization was never detected for wild-type 1–707 or mutant M38/74/111/157–707. The observation that constructs containing wild-type pUL97 or the large isoform of pUL97 predominantly mediate nuclear translocation while constructs containing the small isoform also show partially cytoplasmic localization supports the idea that the small isoform cannot support nuclear transport at full efficiency.

Conclusions

This study provides the first evidence for the generation of two isoforms of pUL97 produced in HCMV-infected cells and transiently transfected cells, and contained within virions. A functional characterization of the two isoforms suggests the following: (i) the large isoform is a full-length product of the entire ORF UL97 initiated at start codon M1, (ii) an alternative translation initiation site is responsible for the production of the small isoform, most probably at start codon M74, (iii) both isoforms possess full catalytic activity, (iv) both isoforms show nuclear localization and (v) the regulation of nuclear translocation of the isoforms appears to be based on independent mechanisms. A bipartite NLS could be identified at residues 6–35, thus specifically regulating the nuclear transport of the large isoform. Furthermore, a second putative NLS region was identified downstream of amino acid 74, which was present in both the large and small isoforms (residues 164–198/190–213). This second region apparently confers a less efficient mode of nuclear translocation and might be specifically responsible for the nuclear transport of the small isoform. In addition to classical NLS function, the protein interactions of pUL97 could also contribute to its nuclear localization. An indirect nuclear transport mechanism could result from interaction of pUL97 with a nuclear cellular protein, such as the retinoblastoma protein (Rb) which contains a strong importin-α-driven NLS (Fontes et al., 2003). However, although motifs involved in the regulation of Rb phosphorylation through pUL97 were described (Gill et al., 2009; Prichard, 2009; Prichard et al., 2008), a direct binding of Rb to pUL97 has not been demonstrated yet. Another factor impacting on the regulation of nuclear translocation may be provided through the pronounced autophosphorylation activity of pUL97. It appears possible that the activity of the NLS sequences described might be modulated by autophosphorylation of serine or threonine residues contained within these regions. However and interestingly, we did not detect any alteration of the localization of GFP–UL97–β-gal fusion constructs carrying NLS sequences 6–35 or 164–198 when we co-expressed active pUL97-F (F, Flag tagged) or the inactive mutant pUL97(K355M)-F (Supplementary Fig. 5, further data not shown). Nevertheless, this finding cannot fully exclude a phosphorylation-specific fine regulation of nuclear translocation. In summary, our study contributes novel aspects to the understanding of pUL97 functionality and points to some yet unknown fine-regulatory differences between the two pUL97 isoforms regarding nuclear transport. Future analyses of the two individual isoforms expressed in the genomic context of recombinant HCMV may further elucidate the role of this specific feature of pUL97 in the replication process of HCMV.

METHODS

Plasmid constructs. Construction of UL97 wild-type, point and deletion mutants pcDNA–UL97–F, pcDNA–UL97(K355M)–F (Marschall et al., 2001), pcDNA–UL97(181–707)–F, pcDNA–UL97(111–707)–F (Marschall et al., 2005), pcDNA–UL97(M1L)–F (Schregel et al., 2007) has been described. For generating the N-terminally tagged construct pcDNA–Flag–UL97 and the C-terminally tagged constructs pcDNA–UL97(38–707)–F, pcDNA–UL97(74–707)–F and pcDNA–UL97(157–707)–F, as well as constructs of UL97 in NLS test vector pHM830 (1–110, 111–180, 181–365, 366–459, 460–707, 1–85, 1–40, 41–48, 49–110, 111–250, 6–35 and 164–198), fragments were amplified from the template pcDNA–UL97–F. For amplification, primers containing EcoRI, HindIII, AfII or XbaI restriction sites were used (Supplementary Table S1, available in JGV Online). PCR was performed with Vent DNA polymerase (New England BioLabs) under standard conditions (Schregel et al., 2007). PCR products were subsequently inserted either into pcDNA3.1 (Invitrogen) or NLS test vector pHM830 producing GFP–UL97–β-gal constructs (Sorg, & Stamminger, 1999). The GFP expression cassette could be removed from the vector pHM830 by HindIII digestion in order to generate UL97–β-gal constructs (1–707, M38/74/111/157–707, M1L–707 and 74–707). A multiple point mutant of pUL97 (M38/74/111/157, i.e. carrying missense mutations M38A, M74L, M111L and M157A) was generated by site-directed mutagenesis (Kunkel, 1985; Supplementary Table S1) of pcDNA–UL97–Flag. A commercially available construct pDsRed1-N1 (BD Clontech) expressing red fluorescent protein was used as a control.

Oligonucleotides. Synthetic oligonucleotide primers for PCR were purchased from Biomers (Supplementary Table S1).

Cell culture, transfections and HCMV infections. Experimental settings for transient transfection (Lipofectamine 2000, Invitrogen; or PEI, Sigma) and HCMV infection of HFFs with HCMV strain AD169 at a multiplicity of 0.5 p.f.u. per cell were performed as described previously (Marschall et al., 2000; Schregel et al. 2007).

In vitro kinase assay. Protein kinase pUL97 was immunoprecipitated from transfected 293T cells and the kinase activity was analysed in vitro (2.5 μCi of γ-32P–ATP) as described by Marschall et al. (2001). For substrate phosphorylation, a mixture of purified histones 1, 2a, 2b, and 3 (Roche) was added to the reactions. To determine the activity of pUL97 derived from virions, a sample of inoculum virus was harvested by high-speed centrifugation, lysed in coimmunoprecipitation buffer and immunoprecipitated with polyclonal antibody (pAb)–UL97 for analysis in an in vitro kinase reaction.
Western blot (Wb) analysis. Transfected 293T cells were lysed and protein extracts were subjected to SDS-PAGE followed by transfer to a nitrocellulose membrane as described previously (Thomas et al., 2009). Immunostaining was performed with the following pAb and mAb: pAb-UL97 (Marschall et al., 2003), mAb-β-actin (Ac-15, Sigma), mAb-Flag (M2, Sigma), pAb-PepAs1–16 (#1343; a pUL97-specific rabbit antiserum raised against peptide pUL97/1–16, MSSALRSRARSASLGT).

Protein purification, mass spectrometry and sequence analysis. The small isoform of pUL97, i.e. pUL97(M11)–Flag, was immunoprecipitated from transiently transfected 293T cells by a Flag purification procedure according to the protocol of the manufacturer (Anti-Flag M2 Affinity Gel; Sigma). The protein was further purified by separation through SDS-PAGE, electroblotted to PVDF membranes and used for Edman sequencing (Proteome Factory AG; Christian Radauer, Inst. Pathophys., Med. Univ., Vienna, Austria). Alternatively, the Coomasie blue-stained pUL97-specific band was excised from an SDS-PAGE gel, eluted and used for tryptic digestion followed by standard mass spectrometry (ISD MALDI–TOF/TOF). Mass spectra were collected using an Ultraflex II MALDI–TOF/TOF mass spectrometer (Bruker) in linear mode, i.e. using a mass range of 2–20 kDa.

Immunofluorescence assays and confocal microscopy. HeLa cells were seeded on coverslips and used for indirect immunofluorescence analysis as described previously (Milbradt et al., 2010). Expression of GFP fusion proteins was analysed directly without antibody staining. Confocal laser-scanning microscopy was performed using a Leica TCS SP5 microscope.

Semi-automatic interactive cell segmentation for the determination of nucleo-cytoplasmic intensity ratios. For an objective quantification of the degree of nuclear versus cytoplasmatic localization of GFP–UL97–β-gal NLS constructs, an interactive image segmentation method was applied to outline the boundaries of individual nuclei from the DAPI images and the boundaries of the corresponding cytoplasmatic areas from the GFP images. Based on the segmentation results, the mean nuclear intensity \( I_{nucleus} \) as well as the mean cytoplasmatic intensity \( I_{cytoplasm} \) could be automatically quantified in order to determine varying patterns of intensity relationships by analysing the intensity ratio \( R = \frac{I_{cytoplasm}}{I_{nucleus}} \). In addition to the measurement of intensities based on semi-automatic interactive cell segmentation and subsequent computation of intensity ratios, each pattern was also manually assigned to a specific class. The classes were defined according to the computed ratios of \( R \); class 1 describes cells with an intensity ratio of \( R<1 \) and indicates localization of the pUL97 construct mainly in the nucleus, class 2 describes cells with \( R=1 \) denoting a nucleo-cytoplasmatic distribution, while class 3 refers to cells with \( R>1 \) in which the localization is mainly cytoplasmatic. The data obtained showed that the objective measurement nicely corresponded to the manually assigned classes.

Computational analysis of secondary structures and NLS sequences. Prediction of pUL97 secondary structure was performed using the NPS@ consensus-prediction method (Combet et al., 2000). NLS candidates were predicted using NLS mapper (Kosugi et al., 2009) and globular protein domains were detected with GlobPlot (Linding et al., 2003). Putative NLS sequences located within globular domains were excluded from further analysis because they are not accessible to importin-mediated transport. GenBank accession numbers of cytomegalovirus genomic sequences are X17403 (human), AF480884 (chimpanzee), NC 006150 (rhesus) and DQ120516 (cercopithecine).

ACKNOWLEDGEMENTS
The authors wish to thank Christian Radauer (Inst. Pathophys., Med. Univ., Vienna, Austria) for support in protein analysis, Thomas Stammeringer (Institute Clin. Mol. Virology, Erlangen, Germany) for kindly providing vectors for the NLS mapping system, Claudia Dach (Fraunhofer Inst., Erlangen, Germany) for valuable contributions to the establishment of interactive cell segmentation methods and Gillian M. Scott (Univ New South Wales, Sydney, Australia) for reading the manuscript. This work was supported by Deutsche Forschungsgemeinschaft SFB 796 (C3, A2 and A4), MA 1289/4-1 and Bayerische Forschungsstiftung (grant 4SC-2010).

REFERENCES


