Structure of a herpesvirus nuclear egress complex subunit reveals an interaction groove that is essential for viral replication

Kendra E. Leigh[1], Mayuri Sharma[1], My Sam Mansueto[1,2], Andras Boeszoermenyi, David J. Filman, James M. Hogle, Gerhard Wagner[3], Donald M. Coen, and Haribabu Arthanari[3]

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Contributed by Gerhard Wagner, June 10, 2015 (sent for review May 15, 2015; reviewed by Wade Gibson and Dmitri N. Ivanov)

Herpesviruses require a nuclear egress complex (NEC) for efficient transit of nucleocapsids from the nucleus to the cytoplasm. The NEC orchestrates multiple steps during herpesvirus nuclear egress, including disruption of nuclear lamina and particle budding through the inner nuclear membrane. In the important human pathogen human cytomegalovirus (HCMV), this complex consists of nuclear membrane protein UL50, and nucleoplasmic protein UL53, which is recruited to the nuclear membrane through its interaction with UL50. Here, we present an NMR-determined solution-state structure of the murine CMV homolog of UL50 (M50; residues 1–168) with a strikingly intricate protein fold that is matched by no other known protein folds in its entirety. Using NMR methods, we mapped the interaction of M50 with a highly conserved UL53-derived peptide, corresponding to a segment that is required for heterodimerization. The UL53 peptide binding site mapped onto an M50 surface groove, which harbors a large cavity. Point mutations of UL50 residues corresponding to surface residues in the characterized M50 heterodimerization interface substantially decreased UL50–UL53 binding in vitro, eliminated UL50–UL53 colocalization, prevented disruption of nuclear lamina, and halted productive virus replication in HCMV-infected cells. Our results provide detailed structural information on a key protein–protein interaction involved in nuclear egress and suggest that NEC subunit interactions can be an attractive drug target.

herpesvirus | cytomegalovirus | drug target | protein–protein interactions | NMR

Herpesviruses encompass a large family of infectious agents, including important veterinary and human pathogens (1). Among the latter is human cytomegalovirus (HCMV), which can cause serious disease, particularly in immunocompromised individuals and newborns (2). Despite the importance of HCMV in these medically vulnerable populations, currently available treatment options suffer from issues with toxicities, drug resistance, and/or pharmacokinetics (2, 3), motivating the identification of new drug targets.

All herpesviruses of mammals, birds, and reptiles undergo a remarkable process known as nuclear egress as part of the viral lifecycle. It is generally accepted that, after assembly in the nucleus, the viral nucleocapsid undergoes envelopment to cross the inner nuclear membrane (INM) followed by deenvelopment to cross the outer nuclear membrane, resulting in release into the cytoplasm for continuation of the virion maturation process (4). Nuclear egress is orchestrated by a highly conserved, heterodimeric nuclear egress complex (NEC), which recruits one or more protein kinases to disrupt the nuclear lamina, permitting access of nucleocapsids to the INM, where the NEC induces budding of the nucleocapsid into the perinuclear space (5–13). In HCMV, the NEC is comprised of UL50, which is an INM protein, and UL53, which is a nucleoplasmic protein that is brought to the INM by its interaction with UL50. These two proteins and their murine CMV (MCMV) homologues, M50 and M53, are essential for replication and nuclear egress (8, 14–17) of their respective viruses. Although a process similar to herpesvirus nuclear egress was recently described for movement of ribonucleoprotein particles during *Drosophila* myogenesis (18), no host cell homolog of the NEC that would serve as mediator of this mechanism has yet been identified. Furthermore, no structural information currently exists for any NEC subunit across the *Herpesviridae* family.

To gain a better molecular understanding of herpesvirus nuclear egress, we used NMR methods to solve the structure of the conserved half of MCMV M50 and map residues on the surface of M50 that are involved in interactions with the other NEC subunit. We then tested the importance of several of these residues for heterodimerization of both the MCMV and HCMV NECs by looking at the effect of single-alanine mutations in both M50 and UL50 on binding affinity and replication of HCMV by looking at the effect of mutations in the context of NEC localization, nuclear lamina disruption, and virus production. Our results identified a subunit interaction interface with features that suggest that it could be an attractive antiviral drug target.

Significance

Human cytomegalovirus (HCMV) is an important human pathogen. Current anti-HCMV therapies suffer from toxicities, drug resistance, and/or pharmacokinetic limitations. A possible antiviral drug target is a two-subunit complex that orchestrates nuclear egress, an essential, unusual mechanism by which nucleocapsids move from the nucleus to the cytoplasm during viral replication. We solved the structure of the conserved core of one subunit of the complex, mapped the primary interaction interface with the other subunit, and tested the importance of specific residues for subunit interactions and viral replication. The combined biophysical and biological analyses presented here develop molecular understanding of nuclear egress and identify a groove that includes a large cavity on the subunit as an attractive target for yet to be identified inhibitors.

Author contributions: J.M.H., G.W., D.M.C., and H.A. designed research; K.E.L., M.S., M.S.M., A.B., D.J.F., and H.A. performed research; K.E.L., M.S., M.S.M., A.B., D.J.F., and H.A. analyzed data; and K.E.L., M.S., J.M.H., G.W., D.M.C., and H.A. wrote the paper.

Reviewers: W.G., Johns Hopkins School of Medicine; and D.N.I., University of Texas Health Science Center at San Antonio.

The authors declare no conflict of interest.

Data deposition: The coordinates for the M50 structures have been deposited in the Protein Data Bank, www.rcsb.org/pdb/home/home.do (PDB ID code 5A3G), and the NMR assignments for M50 have been deposited in BioMagResBank, www.bmrb.wisc.edu (accession no. BMRB-26580).

[1]K.E.L., M.S., and M.S.M. contributed equally to this work.

[2]Present address: Early Discovery Pharmacology, Merck Research Laboratories, Boston, MA 02115.

[3]To whom correspondence may be addressed. Email: gerhard_wagner@hms.harvard.edu or hari_arthanari@hms.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511140112/-/DCSupplemental.

Results and Discussion

Solution-State NMR Structure of M50 (1–168). MCMV M50 is a 316-aa protein with a predicted C-terminal transmembrane region (11) and an N-terminal one-half that is highly conserved among herpesvirus homologs (Figs. S1*A* and S2) (16, 19). The NMR structure of this N-terminal conserved core (M50; residues 1–168) was determined using ∼1,400 NOE-derived distance constraints, of which ∼350 were long-distance constraints. Most of the long-range constraints were derived from 3D time-shared $^{15}N/^{13}C$-resolved nuclear Overhauser effect spectroscopy (NOESY) and 4D ^{13}C-resolved heteronuclear multiple quantum coherence (HMQC)-NOESY-HMQC experiments performed on an isoleucine, leucine, valine (ILV)-labeled sample, where the terminal methyl groups of these residues were ^{13}C-labeled and protonated in an otherwise deuterated background. In addition to NOEs, backbone dihedral constraints derived from the chemical shifts using TALOS+ (20) and orientation constraints derived from backbone amide residual dipolar couplings (RDCs) were used in the structure calculation. The structure was consequently refined in explicit water with 110 N-H RDCs. The RDC-refined and NOE-determined structures of M50 showed essentially indistinguishable secondary structure content and orientation of β-strands (Fig. S1*C*). The backbone rmsd for residues with secondary structure for the 15 lowest energy structures is 1.2 Å, and 98.8% of the residues in these 15 structures occupy the most favored and additionally allowed region in the Ramachandran plot. The complete statistics of the structure calculation as well as adherence to Ramachandran parameters are summarized in Table S1, and Fig. 1*A* displays an ensemble of the 15 lowest energy structures. Longitudinal and transverse relaxation times (T_1 and T_2) as well as hetero-NOE data (Fig. S1*D*) are consistent with a relatively rigid 18-kDa protein, with an average T_2 of 50 ms.

The structure of M50 (1–168) is comprised of a spreading β-sandwich, with nine clearly defined β-strands interspersed by six α-helices. The distinct V shape of the β-sandwich and the positions of the helices (Fig. 1*B*) led to the structure being described by a colleague as a "β-taco." Three of the α-helices are 10 aa or more in length, including one at the N terminus consistent with bioinformatics predictions (19), and the rest are short helices occurring between β-strands. Five of nine β-strands form one side (face A) of the fold, whereas the other four constitute the other side (face B). The opposing β-sheets are intricately woven together (Fig. 1*C*). β-strands 1 and 2 form an antiparallel sheet within face A, and β-strand 3 crosses over to face B and is returned back to face A as β-strand 4, which loops back into β-strand 5. β-strand 5 then traverses to face B through a 15-residue segment to form β-strand 6, which forms the edge of face B. β-strand 6 returns back to face A as β-strand 7 to complete the five antiparallel β-strands that make up face A. β-strand 7 then crosses over to face B through α-helix 5 to form the central two antiparallel β-strands 8 and 9, which complete the four β-strands of face B. β-strand 9 terminates in the final C-terminal α-helix 6. The strands ultimately cross five times between the two faces to form the β-taco fold.

Relationships to Other Proteins. Motivated by the intricately interwoven topology of the M50 (1–168) structure, a number of databases were scanned to search for possible structural homologs. The CATHEDRAL Algorithm (based on the CATH database) (21), the Dali structural comparison database (22), FATCAT (23, 24), VAST (25), PDBeFold (26), and Skan (27, 28) (Table S2) returned hits that failed to encompass the overall M50 (1–168) fold. VAST and PDBeFold returned hits that approximated the dual β-sheet nature of the structure but with additional secondary structural differences. None of the databases returned identical top hits, and regions of structural similarity often encompassed less than one-half of the M50 (1–168) fold. Visual comparison of the overall protein topologies as well as attempts to align the structures revealed major differences, even within areas of structural similarity (Table S2). These results across multiple databases indicate that, although portions of the M50 (1–168) structural fold may be familiar from other known proteins, so far, the overall M50 (1–168) fold cannot be matched to any previously determined structure in the Protein Data Bank (PDB).

To investigate whether any proteins of unknown structure might have a fold similar to that of M50 (1–168), we used threading with Phyre2 (29) in BackPhyre mode to screen several bacterial and eukaryotic sequence databases against the M50 (1–168) coordinates. These searches returned no confidence

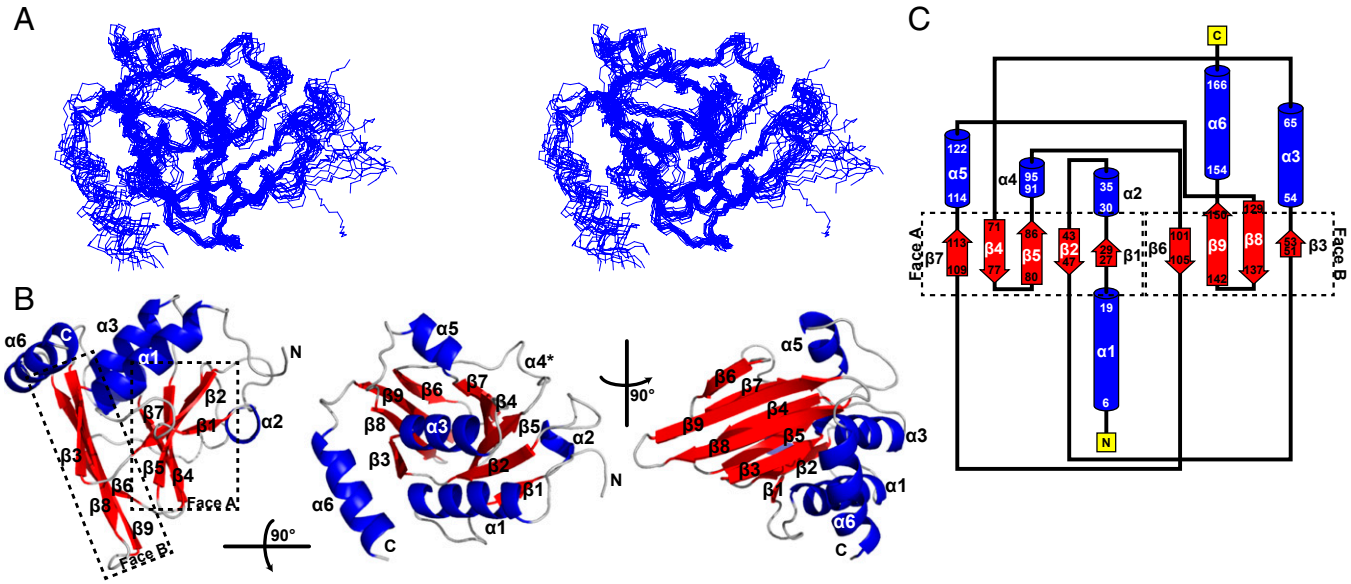

Fig. 1. Solution-state structure of M50 (1–168) as determined by NMR. (*A*) The 15 lowest energy structures generated by the AMBER water refinement are shown overlaid using backbone atoms for residues 10–158 in side-by-side stereoview (Protein Data Bank ID code 5A3G). (*B*) Three viewpoints of a ribbon diagram of M50 (1–168) showing nine β-strands (red) and six α-helices (blue). Faces A and B of the β-taco are delineated in *Left* by the dotted outlines and labeled. *Helix-α4 is recognized as helical by PROCHECK (50) but not by default PyMOL (51) secondary structure identification. (*C*) Schematic of the topology of M50 (1–168).

scores higher than ~42% (for an *Arabidopsis* protein of unknown function) and fit no more than 43 amino acids (of 244 total) for a *Caenorhabditis elegans* protein (F18E9.8). In no case did the polypeptide segments of M50 that were fit form a compact domain. However, threading of viral protein sequences with Phyre2 predicted known homologs of M50, including HCMV UL50 (Fig. S3) (100% confidence score), to have structures highly similar to that of M50 (1–168). Threading of the sequence of the more distantly related HSV-1 homolog of M50 (UL34) onto the M50 structure returned an ~98% confidence score (Fig. S3). Although lack of energy optimization in the Phyre2 calculations limits the likelihood that the actual structural homologies are quite so high, the striking scores support the suggestion that our structural understanding of M50 (1–168) also applies to its homologs in a variety of herpesviruses.

Titration of a Conserved Peptide Identifies a Potential Binding Groove with a Large Cavity. An important function of M50 and its homologs is to interact with their nucleoplasmic partners. To gain insight into this interaction, we hypothesized that the strong sequence conservation between both subunits of the HCMV and MCMV NECs would allow M50 to bind a peptide derived from a region of HCMV UL53 that was previously shown to be important for heterodimer formation (Figs. S1*B* and S4*A*) (30). We titrated this synthetic UL53 peptide into ^{15}N-labeled M50 (1–168) and monitored the binding using a series of ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) experiments. The titration resulted in line broadening and associated reductions of peak intensity of specific resonances in the M50 (1–168) spectrum, indicating that the UL53 peptide binding was in an intermediate exchange regime (Fig. 2 *A* and *B*). When the resonances that displayed substantial broadening are mapped onto the structure of M50 (1–168) (Fig. 2 *C* and *D*), the affected residues form a continuous grooved region that encompasses most of helix-α3 and -α6 and parts of helix-α5 and strand-β8 and -β9. A solvent-accessible cavity identification program (SCREEN) (28, 31) identified eight potential cavities on the M50 (1–168) structure. The number one ranked cavity and largest by surface area runs parallel to the C-terminal helix (α6) and is contained within the area identified by the titration experiment. These results suggested that the peptide was binding within this groove, forming the core of the NEC heterodimerization interface.

Importance of Residues in the Groove and Cavity for Heterodimerization. To test whether residues within the groove and cavity are important for heterodimerization, we individually substituted nine different M50 (1–168) residues identified in the titration experiment with alanine. We also substituted seven of the corresponding residues and three others in HCMV UL50 (1–169) with alanine. Using isothermal titration calorimetry (ITC), we assayed the M50 mutants and the UL50 mutants along with the WT versions of these proteins for their ability to bind M53 (residues 103–333) or UL53 (residues 50–292), respectively. The substitutions resulted in a range of effects, with some causing reductions in binding affinity by a factor greater than 100 (Table 1 and Fig. S4 *B* and *C*). Substitutions that alter residues between positions 52 and 60 of either protein (at the bottom of the groove in Fig. 2 *C* and *D*) all decreased affinity by at least fivefold. This segment of UL50 and M50, particularly E56 and Y57, and the corresponding segment of their pseudorabies virus and Kaposi's sarcoma herpes virus homologs have previously been shown to be important for interaction with the other NEC subunit in less quantitative coimmunoprecipitation and colocalization assays (16, 19, 32, 33). We also found that several substitutions that alter residues between positions 121 and 129 of M50 and 125 and 139 of UL50 (at the top of the groove and within the cavity in Fig. 2 *C* and *D*) also decreased binding, with substitution of a leucine (L129 in M50 and L130 in UL50) within the cavity having at least a 27-fold effect. Finally, substitution of a residue (N154A) within the loop

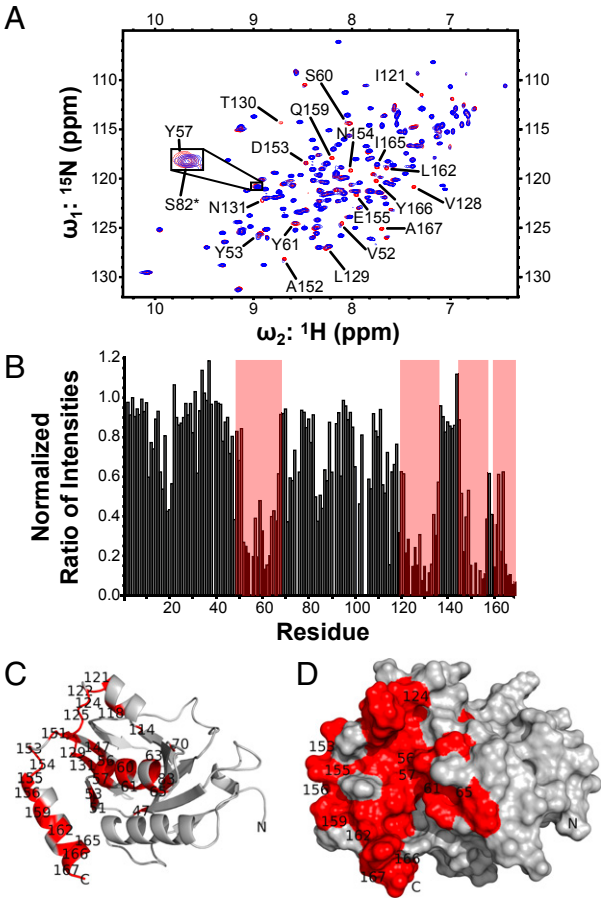

Fig. 2. UL53 peptide binding site. (*A*) Heteronuclear single-quantum coherence spectrum of ^{15}N M50 (1–168) with equimolar unlabeled UL53 peptide titrated into the sample. The expanded view shows Y57 as an example of a peak that shows chemical shift perturbation on the addition of the peptide and S82 (marked with an asterisk) as an example of a peak that shows no chemical shift perturbation. (*B*) Bar graph of the normalized ratio of peak intensities with equimolar unlabeled UL53 peptide present to peak intensities with no unlabeled UL53 peptide present for each residue. Areas highlighted by red boxes are residues that show marked reduction in peak intensity on addition of the peptide, and these areas correspond to the residues shown in red in *C* and *D*. (*C*) Ribbon representation of M50 (1–168), with residues showing a nonzero normalized ratio of peak intensity lower than 0.4 in titration experiments highlighted in red. (*D*) Surface representation of M50 (1–168), with residues showing a nonzero normalized ratio of peak intensity lower than 0.4 in titration experiments highlighted in red. Only a subset of residues has been labeled in *C* and *D* for clarity of presentation.

connecting β9 to α6 in M50 also substantially decreased binding. Deletions of residues with the corresponding segments of the HSV-1 and Kaposi's sarcoma herpes virus homologs decrease subunit interactions based on pull-down and colocalization experiments (33, 34). Our analysis provides a structural basis for these previous findings and identifies specific residues within a cavity of M50 and UL50 that are crucial for subunit interactions.

Previous studies indicated that the interaction interface on M53 and UL53 includes a predicted amphipathic helical segment interrupted by a proline comprising residues 60–82 (Fig. S4*A*) and requires both hydrophobic residues and charged residues within this predicted helix (17, 30). We speculate that at least part of this helical segment fits within the groove in such a way as to allow burying of bulky hydrophobic sidechains present in the peptide (Fig. S4*A*). We found that replacement of certain large hydrophobic residues or neutral polar residues on M50 and UL50 with a smaller hydrophobic residue, alanine, resulted in reduced binding

Table 1. Effects of substitutions on subunit interactions and viral replication

M50 (1–168) mutation	K_d (μM)*	Factor weaker than WT†	UL50 (1–169) mutation	K_d (μM)	Factor weaker than WT	HCMV replication phenotype‡
WT	0.97 ± 0.12	1	WT	0.29	1	WT
V52A	NB	>100	V52A	1.4 ± 0.72	5	5- to 10-fold lower yield
E56A	N/A	N/A	E56A	30–90	>100	Nonviable
Y57A	4.1 ± 0.35	5	Y57A	NB	>100	N/A
S60A	17	20	S60A	N/A	N/A	N/A
Y61A	Insoluble	N/A	Y61A	Insoluble	N/A	N/A
			E63A	1.1 ± 0.33	4	N/A
			R65A	0.55 ± 0.08	2	N/A
I121A	98	>100	I122A	N/A	N/A	
C124A	76	87	S125A	2.2	7	5- to 10-fold lower yield
V128A	0.89	1	V129A	0.32	1	N/A
L129A	60–85	68–100	L130A	7.9	27	Nonviable
			C133A	0.67 ± 0.30	1	N/A
D153A	0.34 ± 0.14	0.4	E154A	0.51 ± 0.24	2	N/A
N154A	16 ± 2.5	16	N155A	N/A	N/A	N/A

N/A, not generated or tested (the M50 and UL50 Y61A mutant proteins were insoluble); NB, no detectable binding.
*K_D values ± SD were derived from ITC data for binding of M53 (residues 103–333) to WT and single point mutant M50 (residues 1–168) proteins and binding of UL53 (residues 50–292) to WT and single point mutant UL50 (residues 1–169) proteins.
†Factors weaker than WT values were calculated by dividing the K_d value of the mutant by the K_d value of the WT protein.
‡Mutant BACs that produced no detectable spread or virus were considered nonviable. Factor reductions in yield of viable mutants were calculated by dividing the supernatant viral titer of the mutant virus by the supernatant viral titer of WT virus at day 7 postinfection (Fig. S5*A*).

of a number of mutants to the nucleoplasmic partner (Table 1). Replacement of the negatively charged residue E56 also reduced binding. In contrast, replacement of a different negatively charged residue in M50, D153, with alanine resulted in a marginal increase in binding. Quantitative differences in binding affinity caused by substitutions of homologous residues (Table 1) show that there are variations in the exact mode of binding, even between closely related proteins, such as M50 and UL50; however, several of the residues that we have identified as important, such as E56 and L129 (L130 in UL50), are identical or similar across widely divergent herpesviruses (Fig. S2), which is consistent with similarities in binding mechanisms.

Importance During HCMV Replication. To examine the importance of individual residues in the groove and cavity during HCMV replication, four substitutions in UL50, which have varying importance for heterodimerization in vitro, were introduced into a GFP-encoding BAC derived from HCMV strain AD169, pBADGFP (35), or modifications of this BAC, in which sequences encoding an FLAG tag have been fused to UL53 coding sequences (8). Electroporation of BACs containing WT *UL50* and mutants V52A and S125A that had 5- to 10-fold reductions in UL53 binding in vitro (Table 1) into human foreskin fibroblasts (HFFs) resulted in a spreading infection, with increasing numbers of GFP-expressing cells showing cytopathic effect, and production of infectious virus. For V52A and S125A, after infection at a low multiplicity, the resulting mutant viruses displayed 5- to 10-fold reductions at 5–7 days postinfection, with S125A maintaining a reduced yield and V52A attaining a near-WT yield at 9 days postinfection (Table 1 and Fig. S5 *A* and *B*). In contrast, electroporation of BACs containing either a *UL50* null mutation [as previously observed (8)], the E56A [as also observed recently (36)], or L130A substitutions that reduced binding to UL53 in vitro by 27- to >100-fold (Table 1) resulted in only individual cells expressing GFP and no detectable

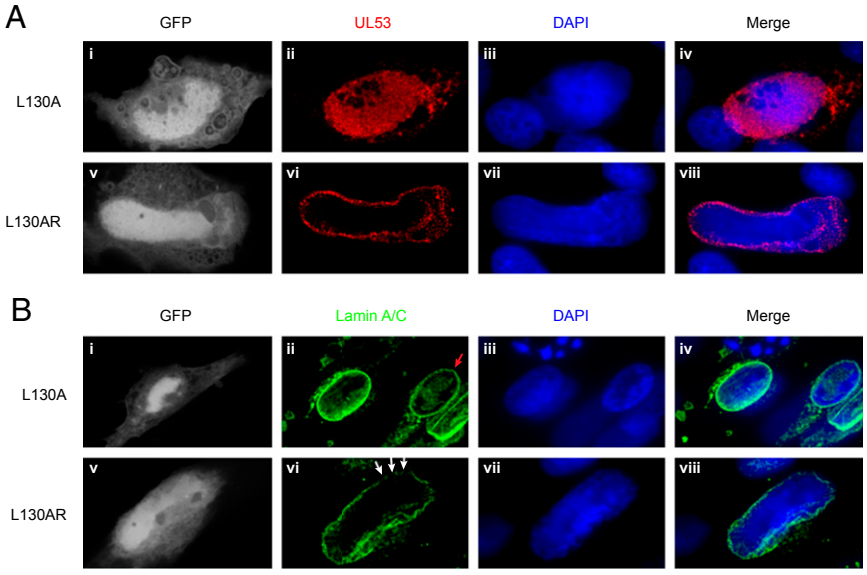

Fig. 3. Effects of substitutions on distribution of UL53 and disruption of nuclear lamina. HFFs were electroporated with (*i–iv*) the mutant L130A UL53-FLAG pBADGFP or (*v–viii*) the rescued derivative L130AR UL53-FLAG pBADGFP. Cells were fixed on day 7, stained with DAPI, and stained with either (*A*) anti-FLAG antibody (red) or (*B*) antibody against lamin A/C (green). Cells positive for GFP were visualized by confocal microscopy. In *B*, the red arrow points to an uninfected cell, and the white arrows point to gaps in nuclear lamina in cells infected with the rescued derivative.

infectious virus. To determine if the effects on virus replication were caused by the introduced substitutions, we constructed rescued derivatives of the mutant BACs, in which each mutation was replaced with WT sequences. For the rescued derivatives of the E56A and L130A mutants (E56AR pBADGFP and L130AR pBADGFP, respectively), electroporation resulted in spread and cytopathic effect. The resulting viruses, and the viruses generated by electroporation of the rescued derivatives of V52A and S125A (V52AR and S125AR) all replicated indistinguishably from WT virus after infection at a low multiplicity (Fig. S5*B*). Thus, there was a strong positive correlation between the effects of substitutions on UL53 binding in vitro and their effects on viral replication in cells.

We then assessed the effects of the lethal substitutions E56A and L130A on the subcellular distribution of UL53 during infection by comparing the mutant viruses with their rescued derivatives. UL53 showed an almost exclusive localization to the nuclear rim in cells infected with the rescued derivatives, similar to what is seen with WT HCMV. However, in E56A and L130 mutant-infected cells, UL53 was distributed throughout the nucleus, similar to what is seen in cells infected with a *UL50* null mutant. (Fig. 3*A* and Fig. S5*C*) (8). Furthermore, cells infected with the mutants L130A or E56A, displayed oval nuclei and almost intact nuclear lamina comparable with mock-infected cells or cells infected with a *UL50* null mutant, whereas cells infected with the rescued derivatives cells showed deformation of nuclear shape, ruffling and thinning of the nuclear lamina, and generation of gaps in the lamina that are characteristic of WT HCMV infection (Fig. 3*B* and Fig. S5*D*). Thus, substitutions of individual residues in the groove and cavity that are lethal for HCMV prevent UL50–UL53 colocalization and the NEC-mediated disruption of the nuclear lamina that accompanies nuclear egress. These results taken together strongly suggest that HCMV NEC subunit interactions are essential for lamina disruption and replication.

Second Functional Surface. Previous work with MCMV has identified dominant negative mutants of M50, which would be consistent with the protein being multifunctional. One of these functions is almost certainly its interaction with the nucleoplasmic subunit (M53), explaining the identified dominant negative phenotypes of insertions between E56 and Y57 and between D125 and K126 of M50 (37), because these residues are located within the core interaction interface described here. A different function may be defined by a dominant negative insertion between Y40 and S41 (37). These residues are found in a loop just after helix-α2. This loop is quite distal to the core interaction surface, and other insertions nearby (between K36 and N37 and between C43 and D44) permit M53 binding (16). Interestingly, substitutions of HSV-1 UL34 residues corresponding to the β1-strand of M50, which is in the vicinity of this loop, also result in a dominant negative phenotype without affecting subunit interactions as assessed by colocalization (38, 39). Thus, these dominant negative mutants seem to identify at least one additional functional surface of this NEC subunit.

Implications for Drug Discovery. Our finding of lethal single-amino acid substitutions within the binding groove/cavity that disrupt NEC subunit interactions suggests that small molecules may be found that mimic this effect, thus ablating viral replication. This interaction interface consists primarily of hydrophobic and aromatic residues interrupted by a few charged amino acids, presenting an ideal landscape for drug design with complementing hydrophobic groups and hydrogen bond donors and acceptors to anchor a small-molecule drug. The complete conservation across human herpesviruses of residues, such as M50 E56, and the conservation of the hydrophobic character of others, such as M50 L129 (UL50 L130) (Fig. S2), which are important for binding and virus viability, raise the possibility that small molecules that bind these sites could have antiviral activity against a broad spectrum of herpesviruses. Such compounds might have considerable specificity given the apparent lack of cellular homologs of this subunit of the NEC.

Although nuclear egress was once thought to be limited to herpesviruses, recently, a broadly similar process has been identified for the export of ribonucleoprotein complexes in *Drosophila* (18, 40), suggesting that insight into herpesvirus nuclear egress may be more generally applicable than previously thought. Still, the unusual structure presented here of a crucial player in nuclear egress underscores the unusual nature of this process. Insights from this structure should apply to widely divergent herpesviruses. Moreover, a combined biophysical and genetic analysis of one of its key functions has identified a target that may permit the rational development of much needed, new therapeutics for diseases caused by HCMV and other herpesviruses.

Materials and Methods

Protein Expression and Purification. Unlabeled and uniformly labeled (^{15}N, ^{13}C, and/or ^{2}H) samples of C-terminal truncations of UL50, UL53, M50, and M53 as well as alanine mutants of UL50 and M50 were derived from constructs created and expressed using methods similar to those described in the work by Sam et al. (30). Isoleucine, leucine, valine-methyl–labeled samples were prepared using α-ketobutyrate and α-ketoisovalerate precursors as described (41, 42). More details are in *SI Materials and Methods*.

Structure and Interface Determination by NMR. Standard NMR spectra for structure determination in addition to 3D ^{13}C,^{15}N time-shared NOESY (43) and 4D ^{13}C-HMQC-NOESY-HMQC experiments (44) were acquired at 291.15 K with ∼150- to 300-μM M50 (1–168) protein samples in 25 mM $NaPO_4$ (pH 6.5), 150 mM NaCl, and 1 mM DTT. All NMR data were processed using NMRPipe (45), and they were visualized and analyzed using CARA (46, 47) and CcpNmr Analysis (48). The structure calculation was run using CYANA. Dihedral angle constraints were extracted from the backbone chemical shifts using the program TALOS+ (20). The 50 lowest energy structures were selected for refinement using ^{15}N-H RDC constraints in explicit water with AMBER through the WeNMR web interface (49). The core heterodimerization interface was mapped using transverse relaxation-optimized spectroscopy (TROSY)-based ^{15}N-HSQC titration experiments with UL53 peptide (residues 58–85). More details are in *SI Materials and Methods*.

Structural Database Search. The PDB coordinate file for the M50 (residues 1–168) fold was submitted to the CATHEDRAL Algorithm (based on the CATH database) (21), the Dali structural comparison database (22), FATCAT (23, 24), VAST (25), PDBeFold (26), and Skan, run through the MarkUs protein annotation server (27, 28), to search for structural homologs using default parameters. More details on top hits are in Table S2.

Structure Threading. All threading calculations were done using the Phyre2 web portal (29). More details are in *SI Materials and Methods*.

ITC. ITC experiments were carried out under conditions similar to those described previously in the work by Sam et al. (30). More details are in *SI Materials and Methods*.

Viruses. pBADGFP-based BACs were electroporated into HFFs as described previously (8) to test for infectivity and when spread occurred, to generate infectious viruses (constructs are listed in Table S3, and constructions and primer sequences are listed in Table S4 and are described in *SI Materials and Methods*). HCMV replication after infection at a multiplicity of infection (MOI) of 0.1 was assessed as described previously (8). Titration was done by infecting 1×10^5 HFF cells per well (in a 24-well plate) with serial dilutions of harvested virus samples for 1 h, after which the inocula were replaced with media containing methylcellulose. After 14 days, the monolayers were stained with crystal violet, and plaques were counted using a dissecting microscope. Titers represent average values from duplicate samples.

Immunofluorescence. HFFs were mock-electroporated or electroporated with pBADGFP constructs (WT, UL50 null, L130A, L130AR, E56A, and E56AR) and treated identically as described previously (8).

ACKNOWLEDGMENTS. We thank Sven G. Hyberts for the hmsIST processing of the 4D NOESY; Remy Sounier and Jim Zhen-Yu Sun for assistance with the residual dipolar coupling refinement; Ekaterina Heldwein for the description

of the fold as a β-taco; Natalia Reim, Jackie Havens, and Stephen Marvel-Coen for technical assistance; the Nikon Imaging Center at Harvard Medical School and its staff for assistance with acquisition and analysis of immunofluorescence data; and Rafael E. Luna for his comments on the manuscript. This study was funded by NIH Grants R01AI026077 (to J.M.H. and D.M.C.), P01GM047467 (to G.W.), P01CA68262 (to G.W.), and P41 EB002026 (to G.W.), and NIH National Institute of Diabetes and Digestive and Kidney Diseases Grant K01DK085198 (to H.A.).

1. Pellett PE, Roizman B (2013) Herpesviridae. *Fields Virology*, eds Knipe DM, Howley PM (Wolters Kluwer/Lippincott Williams & Wilkins Health, Philadelphia), 6th Ed, Vol 2, pp 1802–1822.
2. National Center for Immunization and Respiratory Diseases Division of Viral Diseases (2010) *CDC: Cytomegalovirus (CMV) and Congenital CMV Infection: Overview* (Centers for Disease Control, Atlanta, GA).
3. Boppana SB, Fowler KB (2007) Persistence in the population: Epidemiology and transmisson. *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, eds Arvin AM, Campadelli-Fiume G, Mocarski E (Cambridge Univ Press, Cambridge, United Kingdom), pp 795–813.
4. Mettenleiter TC (2004) Budding events in herpesvirus morphogenesis. *Virus Res* 106(2):167–180.
5. Mettenleiter TC, Müller F, Granzow H, Klupp BG (2013) The way out: What we know and do not know about herpesvirus nuclear egress. *Cell Microbiol* 15(2):170–178.
6. Klupp BG, et al. (2007) Vesicle formation from the nuclear membrane is induced by coexpression of two conserved herpesvirus proteins. *Proc Natl Acad Sci USA* 104(17):7241–7246.
7. Desai PJ, Pryce EN, Henson BW, Luitweiler EM, Cothran J (2012) Reconstitution of the Kaposi's sarcoma-associated herpesvirus nuclear egress complex and formation of nuclear membrane vesicles by coexpression of ORF67 and ORF69 gene products. *J Virol* 86(1):594–598.
8. Sharma M, Kamil JP, Coughlin M, Reim NI, Coen DM (2014) Human cytomegalovirus UL50 and UL53 recruit viral protein kinase UL97, not protein kinase C, for disruption of nuclear lamina and nuclear egress in infected cells. *J Virol* 88(1):249–262.
9. Bigalke JM, Heuser T, Nicastro D, Heldwein EE (2014) Membrane deformation and scission by the HSV-1 nuclear egress complex. *Nat Commun* 5:4131.
10. Reynolds AE, et al. (2001) U(L)31 and U(L)34 proteins of herpes simplex virus type 1 form a complex that accumulates at the nuclear rim and is required for envelopment of nucleocapsids. *J Virol* 75(18):8803–8817.
11. Muranyi W, Haas J, Wagner M, Krohne G, Koszinowski UH (2002) Cytomegalovirus recruitment of cellular kinases to dissolve the nuclear lamina. *Science* 297(5582):854–857.
12. Park R, Baines JD (2006) Herpes simplex virus type 1 infection induces activation and recruitment of protein kinase C to the nuclear membrane and increased phosphorylation of lamin B. *J Virol* 80(1):494–504.
13. Lorenz M, et al. (2015) A single herpesvirus protein can mediate vesicle formation in the nuclear envelope. *J Biol Chem* 290(11):6962–6974.
14. Dunn W, et al. (2003) Functional profiling of a human cytomegalovirus genome. *Proc Natl Acad Sci USA* 100(24):14223–14228.
15. Yu D, Silva MC, Shenk T (2003) Functional map of human cytomegalovirus AD169 defined by global mutational analysis. *Proc Natl Acad Sci USA* 100(21):12396–12401.
16. Bubeck A, et al. (2004) Comprehensive mutational analysis of a herpesvirus gene in the viral genome context reveals a region essential for virus replication. *J Virol* 78(15): 8026–8035.
17. Lötzerich M, Ruzsics Z, Koszinowski UH (2006) Functional domains of murine cytomegalovirus nuclear egress protein M53/p38. *J Virol* 80(1):73–84.
18. Speese SD, et al. (2012) Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. *Cell* 149(4):832–846.
19. Milbradt J, et al. (2012) Specific residues of a conserved domain in the N terminus of the human cytomegalovirus pUL50 protein determine its intranuclear interaction with pUL53. *J Biol Chem* 287(28):24004–24016.
20. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
21. Sillitoe I, et al. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures. *Nucleic Acids Res* 41(Database issue):D490–D498.
22. Holm L, Rosenstrom P (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* 38(Web Server issue):W545–W549.
23. Ye Y, Godzik A (2003) Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics* 19(Suppl 2):ii246–ii255.
24. Ye Y, Godzik A (2004) FATCAT: A web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Res* 32(Web Server issue):W582–W585.
25. Gibrat JF, Madej T, Bryant SH (1996) Surprising similarities in structure comparison. *Curr Opin Struct Biol* 6(3):377–385.
26. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2288–2294.
27. Petrey D, Honig B (2003) GRASP2: Visualization, surface properties, and electrostatics of macromolecular structures and sequences. *Methods Enzymol* 374:492–509.
28. Petrey D, Fischer M, Honig B (2009) Structural relationships among proteins with different global topologies and their implications for function annotation strategies. *Proc Natl Acad Sci USA* 106(41):17377–17382.
29. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10(6):845–858.
30. Sam MD, Evans BT, Coen DM, Hogle JM (2009) Biochemical, biophysical, and mutational analyses of subunit interactions of the human cytomegalovirus nuclear egress complex. *J Virol* 83(7):2996–3006.
31. Nayal M, Honig B (2006) On the nature of cavities on protein surfaces: Application to the identification of drug-binding sites. *Proteins* 63(4):892–906.
32. Paßvogel L, Trübe P, Schuster F, Klupp BG, Mettenleiter TC (2013) Mapping of sequences in Pseudorabies virus pUL34 that are required for formation and function of the nuclear egress complex. *J Virol* 87(8):4475–4485.
33. Luitweiler EM, et al. (2013) Interactions of the Kaposi's Sarcoma-associated herpesvirus nuclear egress complex: ORF69 is a potent factor for remodeling cellular membranes. *J Virol* 87(7):3915–3929.
34. Liang L, Baines JD (2005) Identification of an essential domain in the herpes simplex virus 1 UL34 protein that is necessary and sufficient to interact with UL31 protein. *J Virol* 79(6):3797–3806.
35. Strang BL, et al. (2012) A mutation deleting sequences encoding the amino terminus of human cytomegalovirus UL84 impairs interaction with UL44 and capsid localization. *J Virol* 86(20):11066–11077.
36. Milbradt J, et al. (2014) Proteomic analysis of the multimeric nuclear egress complex of human cytomegalovirus. *Mol Cell Proteomics* 13(8):2132–2146.
37. Rupp B, et al. (2007) Random screening for dominant-negative mutants of the cytomegalovirus nuclear egress protein M50. *J Virol* 81(11):5508–5517.
38. Roller RJ, Bjerke SL, Haugo AC, Hanson S (2010) Analysis of a charge cluster mutation of herpes simplex virus type 1 UL34 and its extragenic suppressor suggests a novel interaction between pUL34 and pUL31 that is necessary for membrane curvature around capsids. *J Virol* 84(8):3921–3934.
39. Bjerke SL, et al. (2003) Effects of charged cluster mutations on the function of herpes simplex virus type 1 UL34 protein. *J Virol* 77(13):7601–7610.
40. Jokhi V, et al. (2013) Torsin mediates primary envelopment of large ribonucleoprotein granules at the nuclear envelope. *Cell Reports* 3(4):988–995.
41. Gardner KH, Kay LE (1997) Production and incorporation of 15N, 13C, 2H (1H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc* 119(32):7599–7600.
42. Rosen MK, et al. (1996) Selective methyl group protonation of perdeuterated proteins. *J Mol Biol* 263(5):627–636.
43. Frueh DP, et al. (2009) Time-shared HSQC-NOESY for accurate distance constraints measured at high-field in (15)N-(13)C-ILV methyl labeled proteins. *J Biomol NMR* 45(3):311–318.
44. Hiller S, Ibraghimov I, Wagner G, Orekhov VY (2009) Coupled decomposition of four-dimensional NOESY spectra. *J Am Chem Soc* 131(36):12970–12978.
45. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293.
46. Keller RLJ (2005) Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. PhD thesis (Swiss Federal Institute of Technology Zürich, Steinerberg, Switzerland).
47. Keller RLJ (2004) *The Computer Aided Resonance Assignment Tutorial* (CANTINA Verlag, Goldau, Switzerland).
48. Vranken WF, et al. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* 59(4):687–696.
49. Bertini I, Case DA, Ferella L, Giachetti A, Rosato A (2011) A Grid-enabled web portal for NMR structure refinement with AMBER. *Bioinformatics* 27(17):2384–2390.
50. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8(4):477–486.
51. Schrödinger LLC (2011) The PyMOL Molecular Graphics System, Version 1.4.1.
52. Hyberts SG, Milbradt AG, Wagner AB, Arthanari H, Wagner G (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J Biomol NMR* 52(4):315–327.
53. Kontaxis G, Clore GM, Bax A (2000) Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J Magn Reson* 143(1):184–196.
54. Schäfer N (1992) Automatische Signalerkennung in heteronuklearen 3D und 4D NMR Spektren. Diploma thesis (Swiss Federal Institute of Technology Zürich, Steinerberg, Switzerland).
55. Güntert P, Braun W, Wüthrich K (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J Mol Biol* 217(3):517–530.
56. Dosset P, Hus JC, Marion D, Blackledge M (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J Biomol NMR* 20(3):223–231.
57. Bhattacharya A, Tejero R, Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins* 66(4):778–795.
58. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26(2):283–291.
59. Tischer BK, Smith GA, Osterrieder N (2010) En passant mutagenesis: A two step markerless red recombination system. *Methods Mol Biol* 634:421–430.
60. Tischer BK, von Einem J, Kaufer B, Osterrieder N (2006) Two-step red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. *Biotechniques* 40(2):191–197.
61. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
62. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4(3):363–371.