Supporting Information

A Small Covalent Allosteric Inhibitor of Human Cytomegalovirus DNA Polymerase Subunit Interactions

Han Chen, Molly Coseno, Scott B. Ficarro, My Sam Mansueto,
Gloria Komazin-Meredith, Sandrine Boissel, David J. Filman, Jarrod A. Marto,
James M. Hogle, Donald M. Coen*

* Correspondence: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Don_Coen@hms.harvard.edu

Table of Contents

Experimental Procedures
Results
References
Figure S1. NMR and LC-MS analyses of SGM8
Figure S2. Activity of SGM8 on long-chain DNA synthesis by UL54 /UL44.
Figure S3. SGM8-binding to UL44 in the crystal structure
Figure S4. The ligand-free conformation of the HCMV UL44 connector loop is altered in different ways by SGM8 and by binding of the C-terminal tail of HCMV Pol.
Figure S5. ^{1}H NMR spectrum of SGM8A
Figure S6. ^{1}H NMR spectrum of SGM8-pentylamine
Figure S7. Binding of peptide P54 to wild type and K60A UL44
Table S1. Crystallographic data collection and refinement statistics for the UL44-SGM8 complex and apo-UL44
Table S2. Anti-HCMV and cytotoxicity activities of SGM8.
Table S3. Screened libraries for HTS

Experimental Procedures

Peptides and proteins

Peptides P54, [1] P30, [2] and P21,[3] both unlabeled and labeled at the N-terminus with the pentafluorescein-derivative Oregon Green 514 (Molecular Probes), were prepared at the Biopolymers Facility of the Department of Biological Chemistry and Molecular Pharmacology (Harvard Medical School).

Wild-type GST-UL44ΔC290 and MBP-UL42ΔC340 fusion proteins, and recombinant human PCNA, were prepared as previously described.[3-6] GST-UL44ΔC290 K60A was obtained using the Quik-Change Mutagenesis Kit (Stratagene) by amplification of the pD15-UL44ΔC290 plasmid[7] using forward primer AGTCACTGCGTGTCAGCCATC ACTTTT, and reverse primer GCATGAGCTGTTAAAAGTGATGGCTG. Wild-type GST-UL54 protein was purified as previously described.[8]

Compounds and FP assays

For high-throughput screening, compounds and extracts from ICCB-L (Harvard Medical School) libraries, listed in Table S3 were tested against UL44-P54 interactions and UL42-P30 interactions using the FP assays described previously.[1, 2] Selected compounds were also tested against PCNA-P21 interactions in a similar FP assay except 11 nM Oregon Green-labeled peptide P21 and 0.25 μM PCNA were used. Compound SGM8 was subsequently purchased from Maybridge and from Aberjona Laboratories. For assaying the time-dependence of SGM8 inhibition, reactions were assembled on ice and started by addition of various concentration of SGM8, then incubated at room temperature for 10, 20, 30, or 60 min prior to measurement of FP values on an EnVision

plate reader (Perkin Elmer). IC_{50} values were calculated using GraphPad Prism (Version 6).

Characterization of SGM8 (5-((dimethylamino)methylene-3-(methylthio)-6,7-dihydrobenzo[c]thiophen-4(5H)-one)

^{1}H NMR (Fig. S1A, top) and ^{13}C NMR spectra (Fig. S1A, bottom) were recorded on a Varian 300 MHz NMR spectrometer and a Bruker Avance III 500 MHz NMR spectrometer, respectively. Chemical shifts are reported in parts per million referenced with respect to residual solvent CD_3CN 1.94 ppm (^{1}H NMR) or DMSO-d_6 2.49 ppm (^{13}C NMR). High resolution mass spectrometry (HRMS) was performed on an Agilent 6530 QTOF LCMS (ESI). The purity of SGM8 in acetonitrile (Fig. S1B) and the stability of SGM8 in HEPES buffer (pH 7.5) containing 1 mg/mL BSA following various lengths of time of incubation at room temperature (Fig. S1C) were assessed by LC-MS using an Agilent 1200 series HPLC with 6130 ESI mass spectrometer. The compound was eluted through a Phenyl-Hexyl column (Luna) with a gradient of 10 - 90% acetonitrile in water at a flow rate of 0.7 mL/min.

^{1}H NMR (300 MHz, CD_3CN) δ 7.55 (s, 1H), 6.93 (s, 1H), 3.17 (s, 6H), 2.94 (t, J = 6 Hz, 2H), 2.80 (t, J = 6 Hz, 2H), 2.59 (s, 3H)

^{13}C NMR (125 MHz, DMSO-d_6) δ 182.42, 149.61, 147.67, 142.43, 132.38, 115.40, 103.09, 43.69, 26.33, 24.47, 17.79

HRMS (ESI) m/z calculated for $C_{12}H_{15}NOS_2$ $[M+1]^+$ = 254.0668, found 254.0669

Effect of SGM8 on interaction of UL44 with UL54.

The inhibitory effect of SGM8 against the interaction of UL44 and full-length UL54 was assessed by measuring the changes in the binding of free UL44 with GST-UL54 at steady

state in the presence of SGM8 using a BLItz system (fortéBIO) at the Harvard Medical School Center for Macromolecular Interactions. (Using this system, we obtained a K_d of 0.92 μM for the UL44-UL54 interaction, which, interestingly, is very similar to that obtained for the UL44-P54 interaction. [1]) Briefly, anti-GST biosensors (fortéBIO) were hydrated in binding buffer (50 mM HEPES (pH 7.5), 0.5 mM EDTA, 1 mM DTT, 150 mM NaCl, 4% glycerol, 100 μg/mL BSA, 0.02% Tween 20) for at least 10 min, and were also primed prior to the loading step following the manufacturer's instructions. GST-UL54 (33 μg/mL) was then loaded on the anti-GST biosensors for 120 s and washed in binding buffer plus 1% DMSO to reach a stable baseline value. UL44 (23 μg/mL), which was pre-incubated with various concentrations of SGM8 at room temperature for at least 20 min, was exposed to loaded biosensors for association until its binding reached equilibrium. The equilibrium signals recorded at each SGM8 concentration minus the baseline value were analyzed by Microsoft Excel, and the IC_{50} was calculated using GraphPad Prism (Version 6).

DNA polymerase assay

Long chain DNA synthesis by UL54 plus UL44 on a poly(dA)-oligo(dT) primer-template (Amersham Biosciences) with radiolabeled dTTP was assayed using a gel-based assay described previously[7], but with 15 nM GST-UL54 plus 40 nM GST-UL44 ΔC290 in the absence or presence of various concentration of SGM8. Incorporated radioactivity was quantified by a phosphorimager (Bio-Rad) and IC_{50} values were calculated using GraphPad Prism (Version 6). The DNA polymerase activity of UL54 alone on the same primer-template with radiolabeled dTTP in the absence or presence of various concentrations of SGM8 was measured using a filter-binding assay reported previously,[9]

but with 15 nM GST-UL54. Radioactivity was measured with a Tri-carb liquid scintillation analyzer (Perkin Elmer) and IC_{50} was calculated using GraphPad Prism (Version 6).

Determining the structure of unliganded UL44

0.2 μL of 8 mg/mL UL44 in storage buffer (20 mM Tris (pH 7.5), 500 mM NaCl, 0.1 mM EDTA, 2 mM DTT and 5% glycerol) and 0.2 μL of crystallization buffer (0.1 M HEPES (pH 7.0) and 18% PEG12K) were mixed and crystallized in sitting drops over the crystallization buffer at room temperature. The crystals were flash-frozen in liquid nitrogen. Diffraction data, in space group $P2_12_12_1$, were collected at the Argonne National Laboratory APS ID24C beamline at 100 K. Data were processed with HKL2000 and merged to yield an essentially complete data set extending to 2.7 Å resolution. Molecular replacement phases were generated using PHENIX, starting with. UL44ΔC290 (PDB code 1T6L). All subsequent refinement was performed using the same methods described for UL44-SGM8 complex.

Model reactions

1. Reaction of SGM8 with pentylamine

2.5 mM SGM8 was mixed with excess pentylamine (Aldrich) in DMSO (10%) and HEPES (0.1M, PH 7.5) at room temperature. Liquid chromatography-mass spectrometry (Waters) was used to monitor this reaction using a gradient of 0-98% acetonitrile in water containing 0.1% trifluoacetic acid (TFA) at a flow rate of 1 mL/min. Peaks were subjected to NMR analysis.

SGM8A (*Z*)-5-(hydroxymethylene)-3-(methylthio)-6,7-dihydrobenzo[*c*]thiophen-4(5*H*)-one

^{1}H NMR (300 MHZ, $CDCl_3$) δ 2.49 (t, *J* = 8.0 Hz, 2H), 2.63 (s, 3H), 2.78 (t, *J* = 8.0 Hz, 2H), 6.78 (s, 1H), 7.51 (s, 1H)

SGM8-pentylamine (3-(methylthio)-5-((pentylamino)methylene)-6,7-dihydrobenzo[c]thiophen-4(5H)-one

^{1}H NMR (300 MHZ, $CDCl_3$) δ 0.92 (t, *J* = 10 Hz, 3H), 1.34 (m, 4H), 1.58 (t, *J* = 10 Hz, 2H), 2.46 (t, *J* = 8.0 Hz, 2H), 2.62 (s, 3H), 2.73 (t, *J* = 8.0 Hz, 2H), 3.20 (t, *J* = 8.0 Hz, 2H), 6.71 (s, 1H), 6.77 (s, 1H), 9.80 (s, 1H)

2. Reaction of SGM8 with peptide Ac-MDKVLNRE

2 nmol of peptide Ac-MDKVLNRE (Bachem) were added to 20 nmol of SGM8 in a mixture of DMSO and crystallization buffer (75 mM ammonium sulfate, 50 mM HEPES (pH 7.5) without PEG 12K) at a 1:1 volume ratio. After incubation at room temperature overnight, the reaction was diluted with 0.1% TFA to a final peptide concentration of 1 pmol/μL. 10 μL of the reaction was desalted using C18 zip tips (Millipore), and eluted to a 384 well MALDI plate. Matrix was added (1 μL of 10 mg/mL α-cyano hydroxycinnamic acid in 80% acetonitrile/0.1% TFA) and the mixture allowed to air dry. MS spectra were acquired using a 4800 MALDI-TOF/TOF mass spectrometer (Applied Biosystems) in reflectron mode averaging 1500 laser shots in a random, uniform pattern (30 subspectra, pass or fail, 50 shots/subspectrum) with a laser intensity of ~4000. MS/MS was performed in reflectron mode averaging 5000 laser shots in a random uniform pattern (100 subspectra, pass or fail, 50 shots/subspectrum, laser intensity ~4500) with CID gas on and the precursor mass window set to relative with a value of 200 (fwhm).

Antiviral activity assay

To monitor antiviral activity more rapidly than is possible with plaque or yield reduction assays, we constructed a virus in which a luciferase reporter gene *luc2*, under the control of the HCMV *UL99* promoter (*UL99* encodes the pp28 tegument protein), was inserted into the HCMV genome between the *US9* and *US10* open reading frames. In the context of the viral genome, this promoter is "true late", requiring viral DNA synthesis for its activity.[10] To generate this virus, we used two-step markerless red recombination[11] in GS1783 *E. coli* cells containing bacmid pAD/Cre [12] (kindly provided by Dong Yu and Tom Shenk, Princeton University) as previously described.[11] To generate the cassette for recombination, the *luc2* gene was cut out of vector pGL4.13 (Promega) using restriction enzymes HindIII and BamHI, and cloned into vector pBlueScript SK (+) (Stratagene) to generate plasmid pBlueScript SK (+) luc2. Then, primers ACC ACA ACC GGT GTC CAA GTC CAC CAC CTT AGC CAG TGT TAC AAC CAA TTA ACC and ATT TTA ACC GGT AAG ACA CTG GGT GTG AAC CTA GGG ATA ACA GGG TAA TCG ATT T containing AgeI cut sites were used to amplify a kanamycin (kan) resistance cassette containing SceI cleavage site from plasmid pLAY2-1-SceIAphAI (gift from J Kamil). This PCR product as well as plasmid pBlueScript SK (+) luc2 were digested with AgeI, and ligated together to form plasmid pBlueScript SK (+) luc2 kan SceI. Then, primers TAT CGG TAC CCA CGA TCA CTA TTA ACG CGA GTC and TAT CGG TAC CCC GGC CCA GCA GCT CGG GCG TG were used to amplify the *UL99* promoter from the HCMV genome, and the product was inserted into the KpnI site of plasmid pBlueScript SK (+) luc2 kan SceI. Finally, primers CCG GCC GCA GGA AGC CGC CCG GCG CGT CGT CTG TGT GCG GGA GCC GAA ACT ACC ACA TTT GTA GAG GTT TTA C, and CAT TGT TGT TTA CTG AAA AGG AAT GTG CTT

TCC CGG CAT GGG CCC GAT TCC ACG ATC ACT ATT AAC GCG AGT C, which contain sequences that overlap within the region of the HCMV genome between *US9* and *US10*, were used to amplify the *UL99* promoter-luc2-kan-SceI cassette from the pBlueScript SK (+) luc2 kan SceI plasmid. This PCR product was then used in the two-step markerless red recombination protocol to construct the recombinant virus. The presence of the insertion in the bacmid was confirmed with a HindIII digest. The bacmid DNA was electroporated into HFF cells and viral plaques appeared seven days later. The virus (HCMV-pp28-LUC) was harvested and expression of the *luc2* gene was confirmed using One-Glo™ luciferase assay (Promega), and the dependence of its expression on viral DNA synthesis was confirmed by showing that luciferase activity was drastically reduced by ganciclovir treatment.

To screen for antiviral activity, human foreskin fibroblast (HFF, ATCC) cells in 20 µL medium (Dulbecco's modified Eagle's medium (Gibco) supplemented with 5% fetal bovine serum (FBS, Gibco) and 1% Penicillin Streptomycin (Gibco)) were seeded at 2500 cells per well into a 384 well plate (Corning). After cell attachment, cells were treated with 20 µL medium containing various concentrations of SGM8 in triplicate. After 2 h incubation at 37°C, cells were infected with 10 µL HCMV-pp28-LUC per well at a multiplicity of infection of 0.1. After incubation of the plates for 3 days at 37°C, 20 µL ONE-Glo™ reagents (Promega) were added to each well following the manufacturer's instructions, and the luminescence produced was measured using an EnVision (Perkin Elmer) at ICCB-L. The EC_{50} value was calculated using GraphPad Prism (version 6).

Cytotoxicity assay

HFF cells were seeded at 3000 cells in 100 μL medium per well in a 96 well plate. After cell attachment, media were removed and 100 μL fresh medium containing various concentrations of SGM8 were added to each well in triplicate. The plate was incubated at 37 ^{o}C for 3 days. Cell viability was then determined by a WST-1 assay provided by the manufacturer (Roche) according to the manufacturer's instructions, and the absorbance at 450 nm from each well as well as a standard curve generated by serial dilutions of cells was recorded using a SpectraMax M5 (Molecular Device) at ICCB-L. The CC_{50} value was calculated using GraphPad Prism (version 6).

Results

We obtained crystals of unliganded UL44 at neutral pH in both the $C222_1$ and $P2_12_12_1$ space groups. The $C222_1$ crystals did not diffract well enough to use, but we solved the crystal structure of unliganded UL44 in the $P2_12_12_1$ space group to 2.7 Å resolution by molecular replacement using the low-pH crystal form of unliganded UL44 as a starting model[4] (Table S1).

In the neutral pH unliganded UL44 structure, two similar, but crystallographically independent, conformations of the connector loop can be visualized because there are two molecules in the asymmetric unit. Regardless, in each conformation, I135 interacts with the β-sheet, and the connector loop can adopt a conformation compatible with extensive hydrogen bonding with P54 (Fig. S4). These connector loop conformations differ from those seen in the earlier low-pH unliganded structure,[4] in which I135 does not anchor the

connector loop, which assumes a substantially different conformation. The differences in connector loop conformation between the unliganded structures are likely due to an artifact of crystal packing in the earlier structure.

References

(1) Loregian, A., and Coen, D. M. (2006) Selective anti-cytomegalovirus compounds discovered by screening for inhibitors of subunit interactions of the viral polymerase. *Chemistry&Biology 13*, 191-200. DOI: 10.1016/j.chembiol.2005.12.002.

(2) Pilger, B. D., Cui, C., and Coen, D. M. (2004) Identification of a small molecule that inhibits herpes simplex virus DNA polymerase subunit interactions and viral replication. *Chemistry&Biology 11*, 647-654. DOI: 10.1016/j.chembiol.2004.01.08.

(3) Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996) Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell 87*, 297-306. DOI: 10.1016/S0092-8674(00)81347-1.

(4) Appleton, B. A., Loregian, A., Filman, D. J., Coen, D. M., and Hogle, J. M. (2004) The cytomegalovirus DNA polymerase subunit UL44 forms a C clamp-shaped dimer. *Mol. Cell. 15*, 233-244. DOI: 10.1016/j.molcel.2004.06.018.

(5) Bridges, K. G., Chow, C. S., and Coen, D. M. (2001) Identification of crucial hydrogen-bonding residues for the interaction of herpes simplex virus DNA polymerase subunits via peptide display, mutational, and calorimetric approaches. *J. Virol. 75*, 4990-4998. DOI: 10.1128/JVI.75.11.4990-4998.2001.

(6) Almendral, J. M., D., H., Blundell, P. A., Macdonald-Brown, H., and Bravo, R. (1987) Cloning and sequence of the human nuclear protein cyclin: homology with DNA binding proteins. *Proc. Natl. Acad. Sci. USA 84*, 1575-1579.

(7) Loregian, A., Appleton, B. A., Hogle, J. M., and Coen, D. M. (2004) Residues of huamn cytomegalovirus DNA polymerase catalytic subunit UL54 that are necessary and sufficient for interaction with the accessory protein UL44. *J. Virol. 78*, 158-167. DOI: 10.1128/JVI.78.1.158-167.2004.

(8) Chen, H., Beardsley, G. P., and Coen, D. M. (2014) Mechanism of ganciclovir-induced chain termination revealed by resistant viral polymerase mutants with reduced exonuclease activity. *Proc. Natl. Acad. Sci. USA 111*, 17462-17467. DOI: 10.1073/pnas.1405981111.

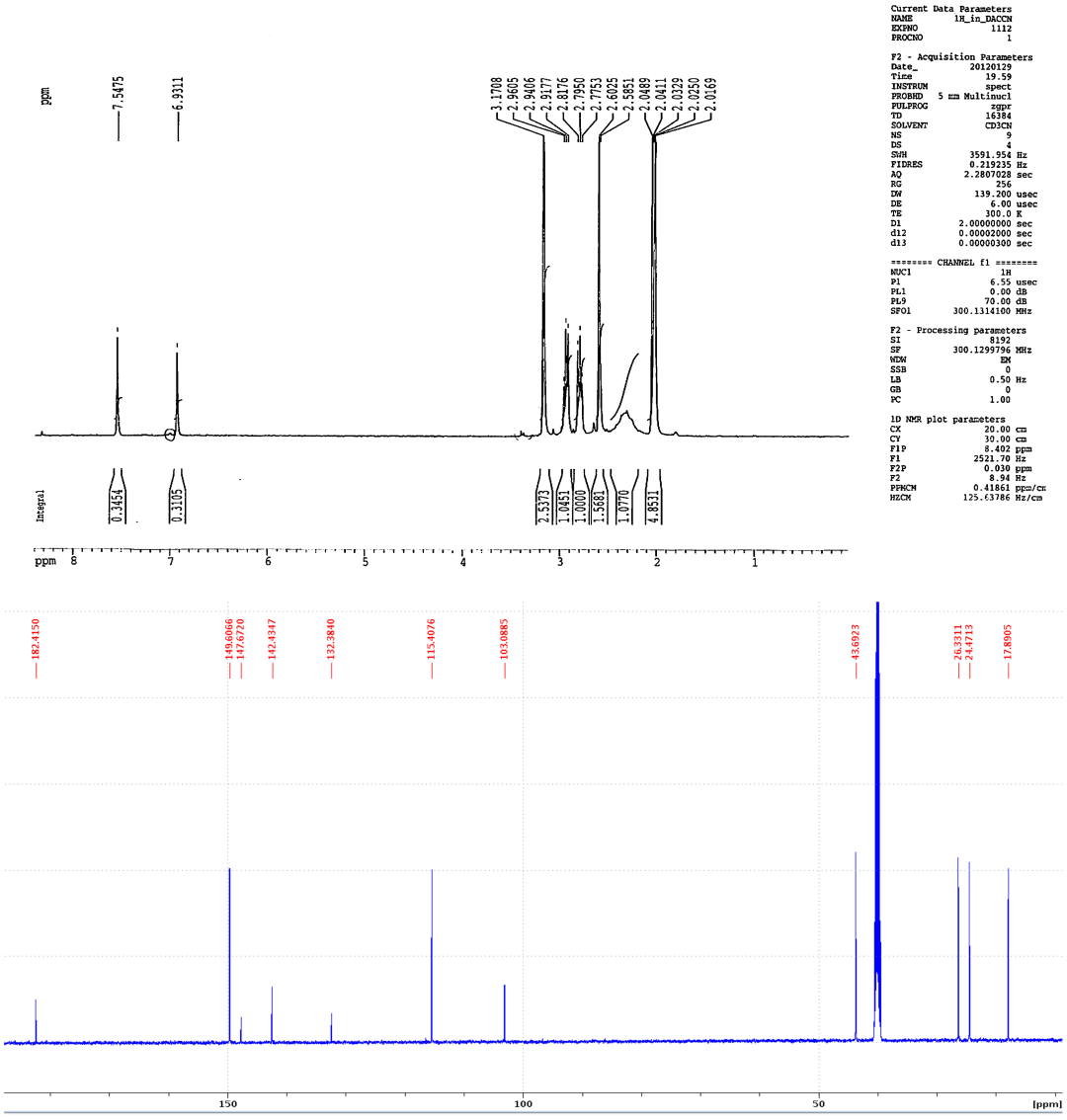
(9) Loregian, A., Rigatti, R., Murphy, M., Schievano, E., Paul, G., and Marsden, H. S. (2003) Inhibition of human cytomegalovirus DNA polymerase by C-terminal peptides from the UL54 subinit. *J. Virol. 77*, 8336-8344. DOI: 10.1128/JVI.77.15.8336-8344.200.

(10) Mocarski, E. S., Shenk, T., Griffiths, P. D., and Pass, R. F. (2013) Cytomegaloviruses. In *Field Virology* (Knipe, D. M., and Howley, P. M., Eds.), 6th ed., pp 1960-2014, Lippincott Williams & Wilkins, Philadelphia.

(11) Tischer, B. K., Smith, G. A., and Osterrieder, N. (2010) En passant mutagenesis: a two step markerless red recombination system. *Methods. Mol. Biol. 634*, 421-430. DOI: 10.1007/978-1-60761-652-8_30.

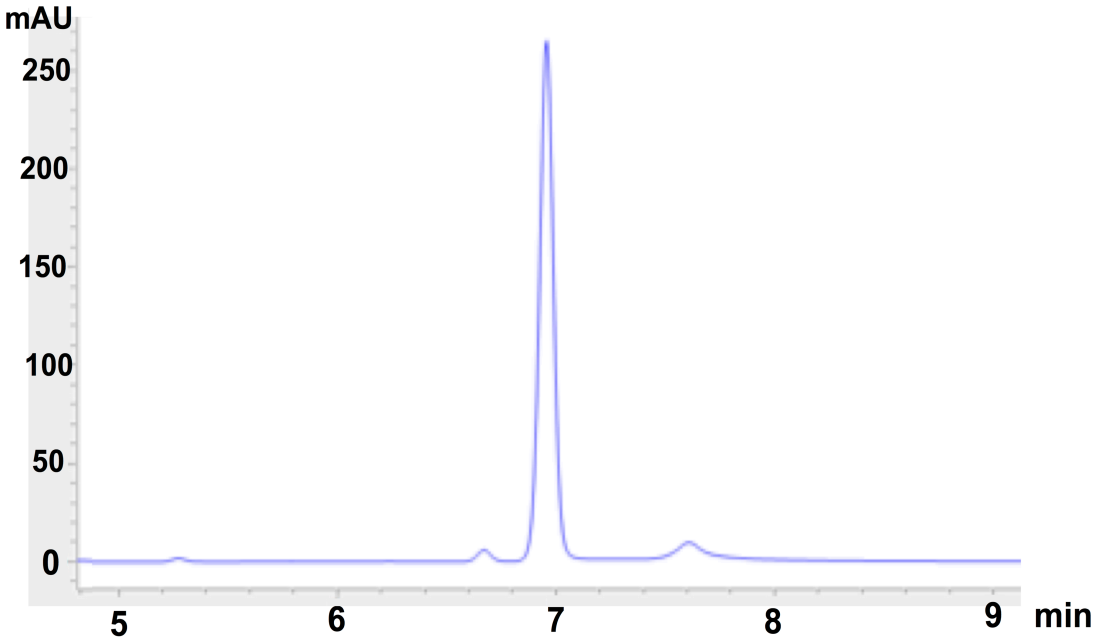
(12) Yu, D., Smith, G. A., Enquist, L. W., and Shenk, T. (2002) Construction of a self-excisable bactrial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL 13 gene. *J. Virol. 76*, 2316-2328. DOI: 10.1128/jvi.76.5.2316-2328.2002.

A.




```
Current Data Parameters
NAME          1H_in_DACCN
EXPNO             1112
PROCNO               1

F2 - Acquisition Parameters
Date_          20120129
Time              19.59
INSTRUM           spect
PROBHD   5 mm Multinucl
PULPROG           zgpr
TD               16384
SOLVENT           CD3CN
NS                   9
DS                   4
SWH           3591.954 Hz
FIDRES        0.219235 Hz
AQ            2.2807028 sec
RG                 256
DW              139.200 usec
DE                6.00 usec
TE               300.0 K
D1            2.00000000 sec
d12           0.00002000 sec
d13           0.00000300 sec

======== CHANNEL f1 ========
NUC1                1H
P1                6.55 usec
PL1               0.00 dB
PL9              70.00 dB
SFO1        300.1314100 MHz

F2 - Processing parameters
SI                 8192
SF          300.1299796 MHz
WDW                  EM
SSB                   0
LB                 0.50 Hz
GB                    0
PC                 1.00

1D NMR plot parameters
CX                20.00 cm
CY                30.00 cm
F1P               8.402 ppm
F1              2521.70 Hz
F2P               0.030 ppm
F2                8.94 Hz
PPMCM          0.41861 ppm/cm
HZCM          125.63786 Hz/cm
```

B.



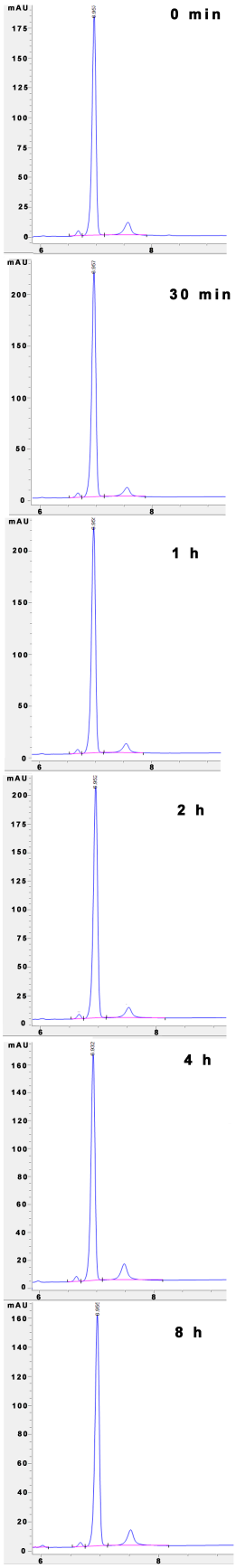

0 min
30 min
1 h
2 h
4 h
8 h
mAU

Figure S1. Characterization of SGM8. A) ^{1}H NMR (top) and ^{13}C NMR (bottom) spectra of SGM8; B) LC-MS analysis indicating the purity of SGM8 (eluting at 6.9 min); C) LC-MS analysis of the stability of SGM8 in HEPES buffer (pH 7.5) containing 1mg/mL BSA following incubation at room temperature for the times indicated in the figure.

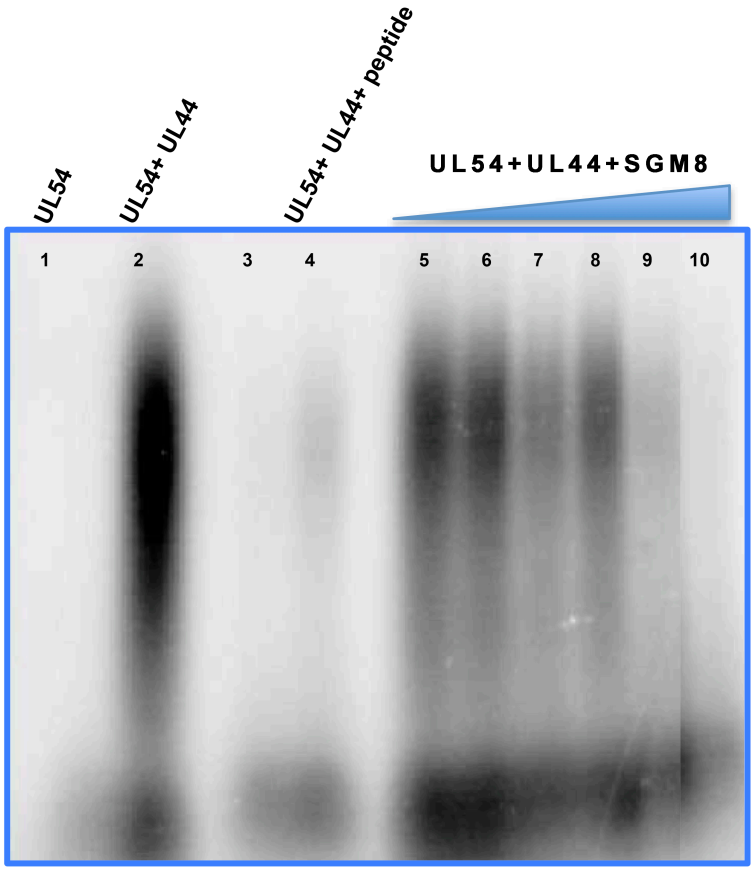


Figure S2. Inhibition of long-chain DNA synthesis by UL54 plus UL44. Primer-template DNA and radiolabeled dTTP were incubated with either UL54 alone (lane 1) or by UL54 plus UL44 (lanes 2-10), either in the absence of inhibitors (lanes 1 and 2), in the presence of 100 μM (lane 3) or 10 μM (lane 4) peptide P54 (as a positive control) or varying concentrations of SGM8 (lanes 5-10) ranging from 0.5 μM to 50 μM. The products were visualized by autoradiography following electrophoresis on a 4% alkaline agarose gel.

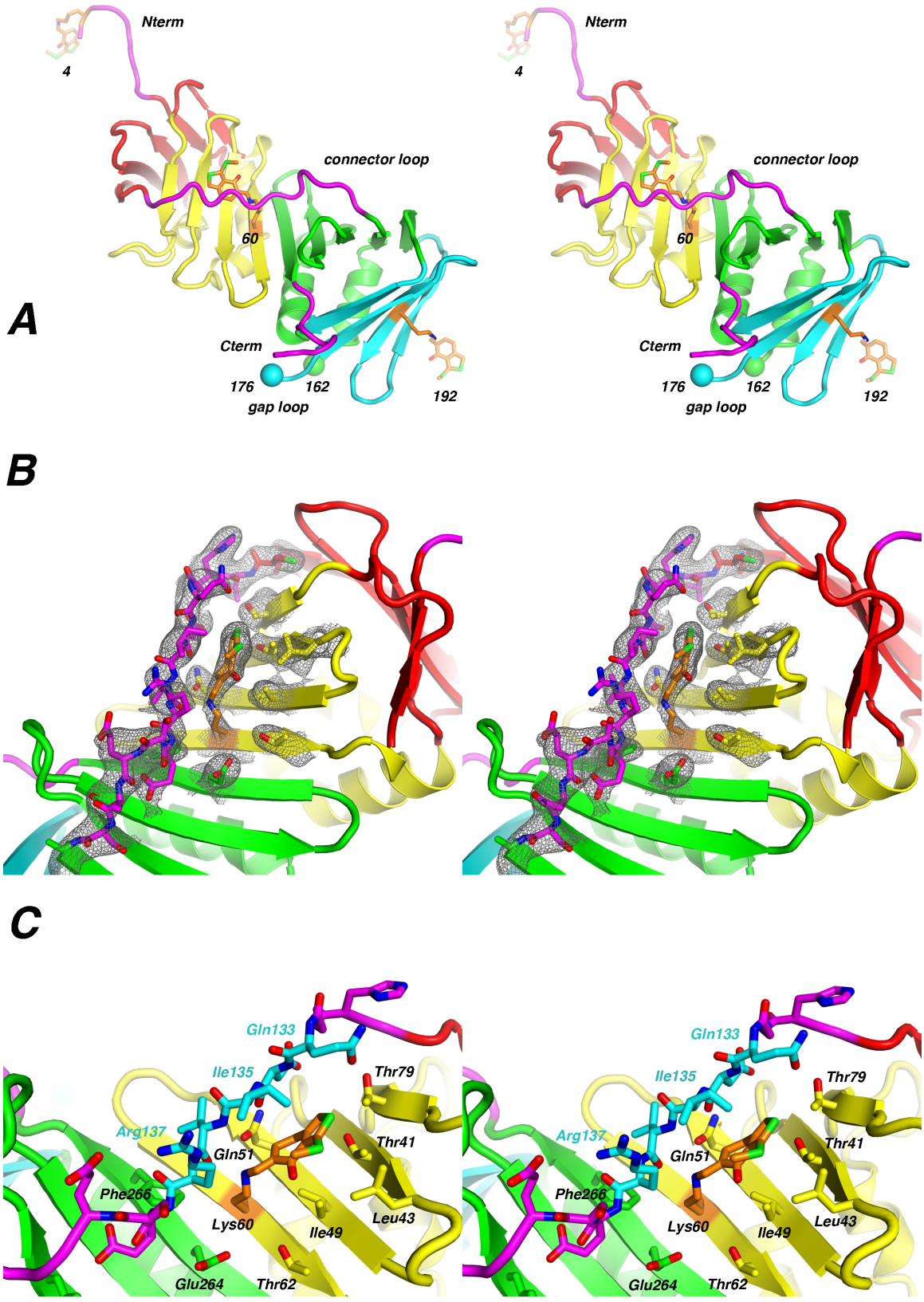

A
Nterm
4
connector loop
60
Cterm
176
162
192
gap loop
B
C
Gln133
Ile135
Thr79
Arg137
Thr41
Gln51
Phe266
Lys60
Ile49
Leu43
Glu264
Thr62

Figure S3. SGM8-binding to UL44 in the crystal structure. (A) Stereo ribbon diagram of the SGM8-modified UL44 monomer. The location of the major SGM8 substitution site, at Lys60, and of partially occupied sites, at Lys192 and Lys4, are shown in the context of UL44 monomer architecture. UL44 has a central 9-stranded beta sheet (yellow and green), flanked by 5-stranded and 4-stranded sheets at the amino and carboxyl ends. The red-yellow amino sub-domain connects to the structurally similar green-cyan carboxyl sub-domain via the UL54-binding connector loop (horizontal magenta trace). The "gap loop" that plays an important role in DNA binding is disordered between residues 162 and 176 (which are indicated by labeled spheres). The “C-clamp” dimer is formed by the stable interaction of two copies of the red β-sheet. (B) The 2Fo-Fc electron density map (gray mesh) in the vicinity of Lys60 (orange). Density features are well accounted for by the atomic model of the SGM8-UL44 complex, including the complex-specific conformation of the connector loop (magenta) and the position of the SGM8 modification of Lys60 (orange). (C) The SGM8 binding site is shown in stereo. SGM8-modified Lys60 binds to UL44 in a hydrophobic slot, making contacts with residues 133-137 of the connector loop (cyan), and with side chains from the central beta sheet (yellow and green). In this ribbon representation of a portion of the SGM8-UL44 complex, important contacts are shown as stick models, and are labeled.

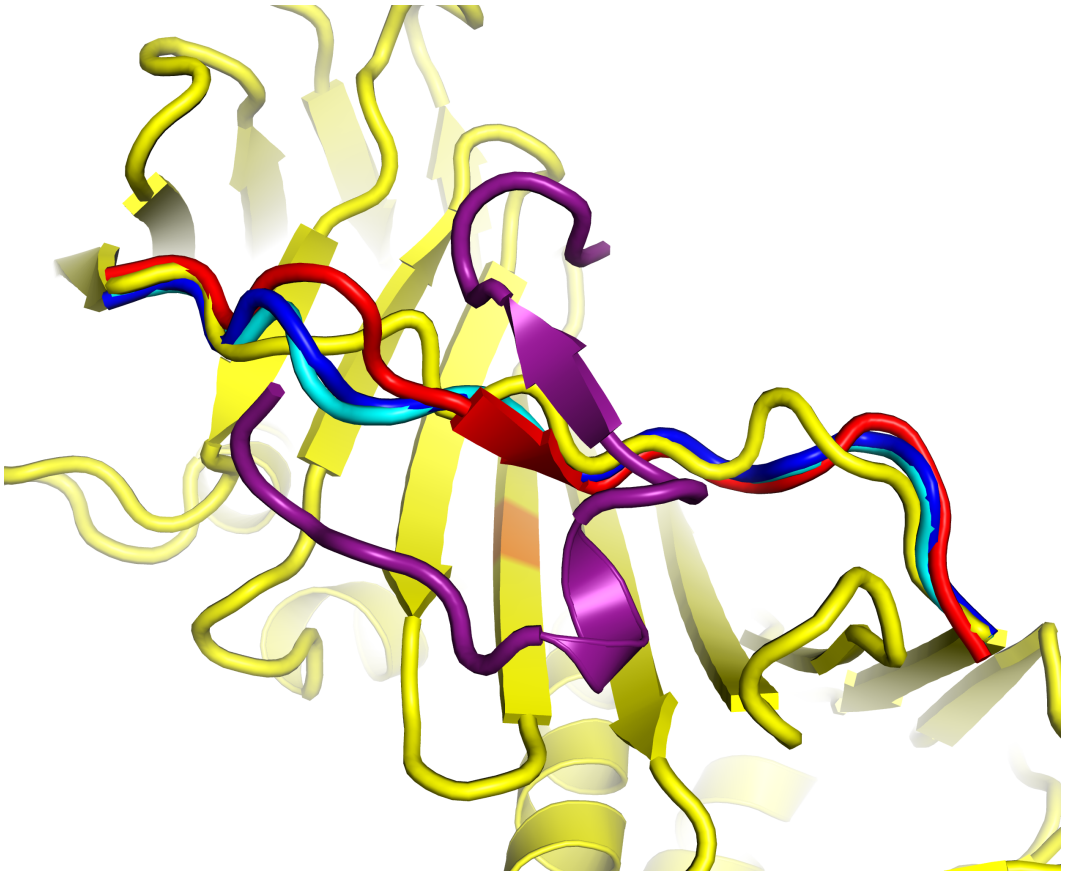


Figure S4. The ligand-free conformation of the HCMV UL44 connector loop is altered in different ways by SGM8 and by binding of the C-terminal tail of HCMV Pol. The ribbon diagram in yellow represents a portion of the UL44 complex with SGM8, in space group $C222_1$ (PDB entry 5IWD), but with SGM8 and side chains omitted. The main chain position of Lys60, which is modified by SGM8, is indicated in orange. For comparison, its connector loop (residues 128-143; yellow) is shown superimposed on connector loops from other UL44 crystal structures. The blue and cyan connector loops represent two crystallographically independent copies of ligand-free UL44 (PDB entry 5IXA), crystallized at neutral pH in space group $P2_12_12_1$. The binding of P54 peptide (shown in purple) causes the connector loop to adopt the red conformation, when crystallized from high salt at low pH in space group $C222_1$ (PDB entry 1YYP, Appleton et al.). None of

these connector loops are involved in crystal packing contacts, except for minor contacts in the cyan-colored copy. Clearly, the connector loop is flexible, and its conformation readily adapts, in different ways, depending upon which of its ligands is available to bind.

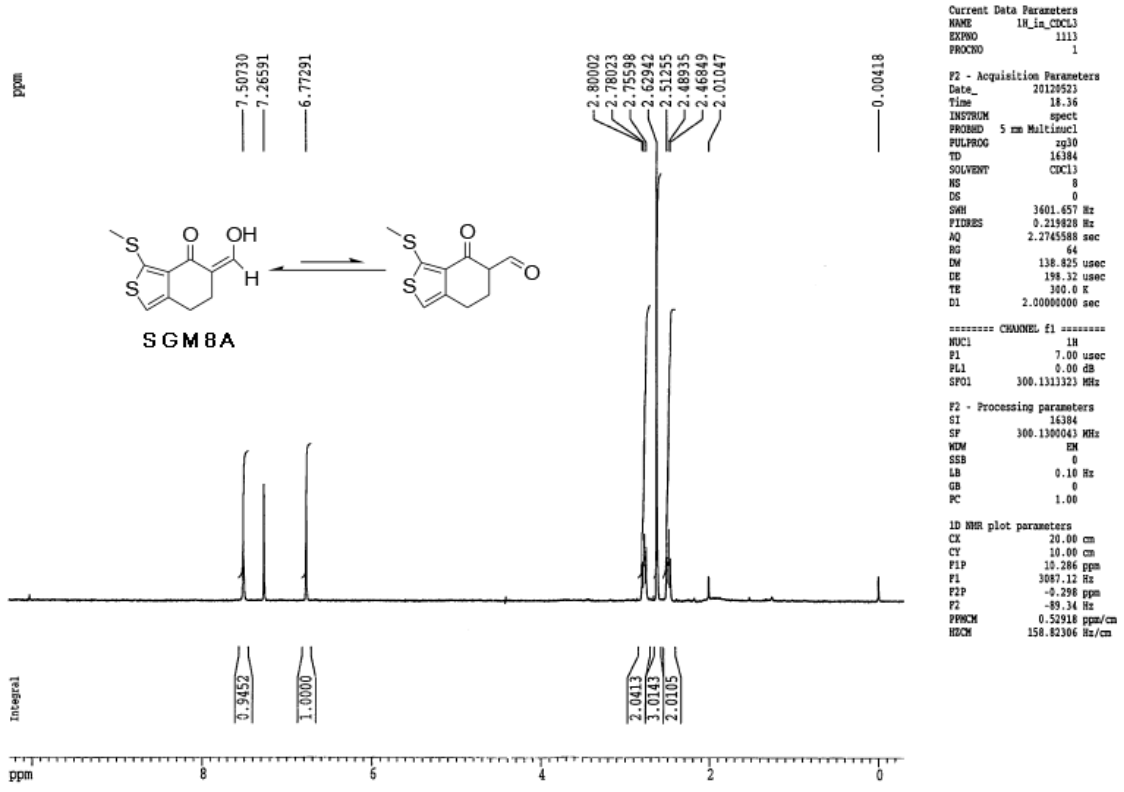


Figure S5. ^{1}H NMR spectrum of SGM8A

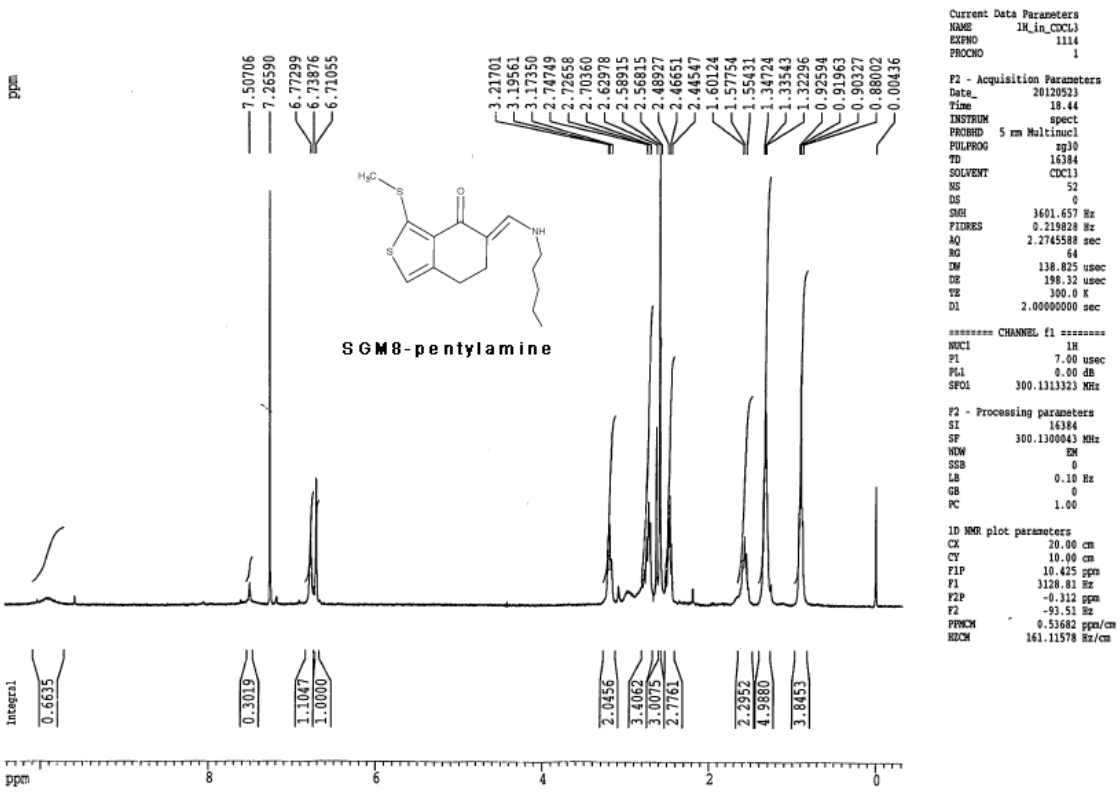


Figure S6. ^{1}H NMR spectrum of SGM8-pentylamine

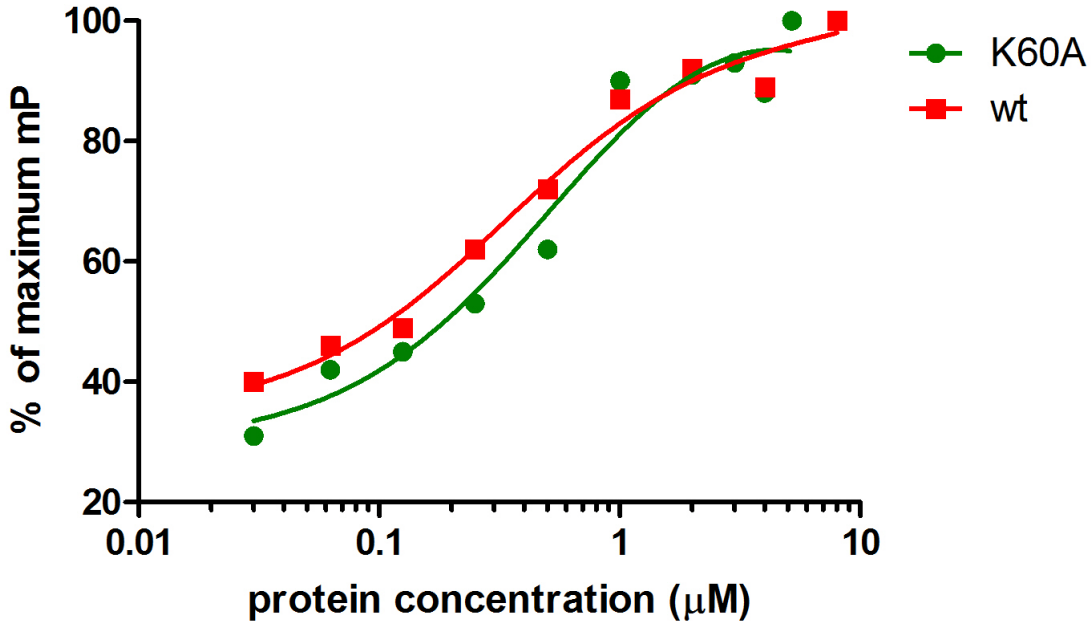


Figure S7. Comparison of binding of peptide P54 to wild type UL44 (red squares) and UL44 mutant K60A (green circles). Increasing amounts of wild type or mutant protein were assayed for binding to fluorescently labeled peptide in an FP assay, and the % maximal binding was plotted vs. protein concentration.

Table S1. Crystallographic data collection and refinement statistics for the UL44-SGM8 complex and apo-UL44

	UL44ΔC290 + SGM8	apo-UL44ΔC290
Data collection		
Space group	$C222_1$	$P2_12_12_1$
Cell dimensions		
a,b,c (Å)	91.80, 127.04, 66.00	75.03, 100.83, 109.64
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50 - 2.55 (2.64-2.55)	50 – 2.7 (2.8 – 2.7)
R_{sym}	0.038 (0.478)	0.109 (0.748)
I/σI	27.1 (2.1)	10.6 (1.4)
Completeness (%)	98.0 (100.0)	98.0 (99.1)
Redundancy	3.0 (3.1)	3.1 (3.1)
Refinement		
Resolution (Å)	27.12 - 2.56 (2.82 – 2.56)	37.1 – 2.7 (2.8-2.7)
No. reflections	12465	23307
R_{work}	0.2125 (0.2764)	0.2103 (0.3038)
R_{free}	0.2670 (0.3271)	0.2587 (0.3642)
No. atoms		
Protein chain A	2055	1956
chain B		1969
Water	1	98
R.m.s deviations		
Bond lengths (Å)	0.008	0.011
Bond angles (°)	1.102	1.587
PDB code	5IWD	5IXA

Table S2. Anti-HCMV and cytotoxicity activities of SGM8.

	EC_{50}	CC_{50}
SGM8	6 μM	7 μM

Table S3. Screened libraries for HTS

Library	**Number of compounds**
Biomol ICCB Known bioactive 2-High Conc.	480
Ninds Custom Collection 2	1,040
Prestwick 1 Collection	1,120
ICBG 9 – Fungal Extracts	2,816
ICBG 11 – Fungal Extracts	2,112
MMV 1	2,464
NCDDG 2	2,816
Maybridge 5	3,212
ChemDiv 4	14,677
Asinex1	12,378
Enamine 2	13,376
Enamine 2, cont.	13,200
ICBG 2-Fungal Extracts	460
STARR Foundation 2 - Extracts	1,000
Life Chemicals 1	3,893
NCDDG 1	380
ICBG 4 - Fungal Extracts	704
ICBG 6 – Sherman Lab Extracts	1,276
ICBG 5 – Fungal Extracts	2,112

ICBG 7 – Sherman Lab Extracts	1,316
ICBG 8 – Sherman Lab Extracts	1,120
NCDDG 3	1,056
ChemDiv 3	16,544
MixCommercial 5	268
ActiMol TimTech 1	8,518
Bionet 2	1,700
Peakdale 2	352
ChemDiv 2	8,560
I.F. Lab 2	292
Maybridge 4	4,576
Enamine 1	6,004
Maybridge 3	7,639
MixCommericial 2	320
MixCommericial 3	251
Maybridge 2	704
ChemDiv 6	22,176
Organic Fractions – NCI plant and Fungal Extract	1,408
ChemDiv Antimitotic Collection	1,254
Chembridge 3	6,336
ChemDiv 6, cont.	21,120
CEREP	4,800
MixCommericial 1	352
Chembridge 3, Cont.	4,224
ChemDiv 6, cont.	704
NCDDG 4	4,576
ICBG 12 - Fungal Extracts	2,816

MMV 2	2,464
ICBG 10 - Fungal Extracts	1787