Supplementary Information

Protein Interactions in the Murine Cytomegalovirus Capsid Revealed by CryoEM

Wong H. Hui^{1,2}, Qiyi Tang³, Hongrong Liu^{1,2,6}, Ivo Atanasov^{1,2}, Fenyong Liu⁴, Hua Zhu⁵, Z. Hong Zhou^{1,2,*}

¹The California NanoSystems Institute (CNSI), University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594, USA
²Department of Microbiology, Immunology & Molecular Genetics, UCLA, Los Angeles, CA 90095-736422, USA
³Department of Microbiology/AIDS program, Ponce School of Medicine
395 Zona Industrial Reparadara-2, Ponce, PR 00716-2348, Puerto Rico
⁴ School of Public Health, University of California at Berkeley, Berkeley, CA 94720, USA
⁵Department of Microbiology and Molecular Genetics, International Center for Public Health (or ICPH), Room E350D, 225 Warren Street, UMDNJ-New Jersey Medical School, Newark, NJ 07103-3506, USA
⁶Current address: College of Physics and Information Science, Hunan Normal University, 36 Lushan Road, Changsha, Hunan 410081, China

Running title: Murine cytomegalovirus capsid structure

Key words: cytomegalovirus; herpes simplex virus type 1; cryo electron microscopy; three-dimensional; major capsid protein.

^{*}To whom correspondence should be addressed.

E-mail: Hong.Zhou@ucla.edu.

Tel: 310-983-1033

Legends of Supplemental Figures:



Figure S1. Fourier shell correlation (FSC) curve. The estimated resolution of the MCMV capsid is 9.1 Å at FSC = 0.5.

HSV-1:	484	ANPYGAYVAAP-AGPAADMQQLFLNawgQRLAHGRVRWVAEGQMTPEQFMQPDNANLALE	542
HMM :		aellgleverepeanlarr1refyeakaeapvtdnaakaelttedflkpenelLr1E	504
MCMV :	478	RALCRREFVREHMAHATRRLVHFYQARIDPPRTANEAKHDFSTKEFAKVDNYLLFTE	534
HSV-1:	543	LHPAFDFFVGVADVELPggdvppagpgEIQATWRVVNGNLPLALCPAAFRDARGLELGV-	601
HMM :		lhPlFDffvaqedvelpavratpRiivGniPlpLapadFreargkqleea	
MCMV :	535	LHPFFDFCFHTENGQVRPLCTPRIMVGNLPEALAPADFHDLRAKQALEL	583
HSV-1:	602	-GRHAMAPATIAAVRGAFDDRNYPAVFYLLQAAIHGSEHVFCALARLVVQCITSYWNN	658
HMM :		aklakle.seatvevvqetledpnYPelfYliealiHGseeaFlalarlvaqcinsywen	
MCMV :	584	${\tt TKVRAPEgHEATLQVLRASLTDHQYPELFYLIESLIHGDPAAFETGIELVTRCVNNYWRQ$	643
HSV-1:	659	TRCAAFVNDYSLVSYVVTYLG-GDLPEECMAVYRDLVAHVEALAQLVDDFTLtGPELGGQ	717
HMM :		sgrlaFvnsfemvklIathlgdgelpeevlavYrkllsevralkrlvskltl.neqlgee	
MCMV :	644	${\tt RGLLAFANSYDMVRLIATRLGDGAVVPAAYTHYRNLLSITRFVARTCELTGL-{\tt NGRLCDE}$	702
HSV-1:	718	AQAELNHLMRDPALLPPLVWDCDALMRRAAldrhRDCRVSAGGHD-PVYAAACNVATADF	776
HMM :		$\verb sleelvnallDpallpPlvydcdallkreernvkvkaggeelkaaeaeernaevsf $	
MCMV :	703	PLLAYVSALHDPRLWPPFVQALPRNANLVRVVADDVPLDAAHIEERNPGTSD	754
HSV-1:	777	N-RNDGQLLHNTQARAADAADDRPHRGADWTVHHKIYYYVMVPAFSRGRCCTAGVRFD	833
HMM :	:	verlenllaheravvddrreadedephrdeeelvleKifYyvvlPaltngrvCgmGvdlk	
MCMV :	755	VARMIAMDQAEPLFVDARRTSDEEMVAQKVYYLCLVPAVLNNHACGAGLNLK	806
HSV-1:	834	RVYATLQNMVVPEIAPGEECPSDpvtdpahplhpaNLVANTVNAMFHNGRVV-V	886
HMM :		nvlltlfyneavvvpdevaeeeaildnllaetlndllhnseva.v	
MCMV :	807	HLLVKLFYTKFFLTADpdsLTAGEEALTNnPLLAALVRDVATDENVTaN	855
HSV-1:	887	DGPAMLTLQVLAHNMAERTTALLCSAAPDAGANTASTT-NMRIFDGALHAGILLMAPQHL	945
HMM :		$\label{eq:constraint} dadalre L qelv l n va ert kalevea al Daa qrt aat e.n frvldgv L yn Glllmaip kr$	
MCMV :	856	QAAEELFHLVAHVPENAQMLEIRAALDPAQRHGAPSaGFESLQHVLYNGFCMTTVPKL	913
HSV-1:	946	DHTIQnGDYFYPLPVHALFAGAdhvANAPNFPPALRDLSRQVPLVPPALGA	997
HMM :		${\tt drtva.eey} fypv {\tt Pvhklyadpavaatlnaeirelleelps.qRndgg {\tt Fpvppalaa}$	
MCMV :	914	LQEYLTVIPFHRFYSDPGLAATANHDIRVFLNDFPQyQRCDGGFPLSPIFAH	965
HSV-1:	998	NYFSSIRQPVVQHVRESAAGENALTYALMAGYFKISPVALHHQLKTGLH 1045	
HMM :		eyyewhrsPvakYaaecaatlnsls.aLlamyfKlsPvalilqlktklH	
MCMV :	966	EYHHWHRTPFSCYSAACAHTLESVL-TLAIMHHKMSPVSIAALSRMGLH 1013	

Fig. S2

Figure S2. Sequence Alignment of MCPud of HSV-1 and MCMV

Alignment of sequence between the MCP upper domain of MCMV with HSV-1, obtained from the Pfam Sequence Search server (Finn et al., 2010) (<u>http://pfam.sanger.ac.uk/search</u>). Only the results corresponding to residues of HSV-1 MCPud (aa 484 to 1045) are shown. The section which is discussed in Figure 2 is highlighted in red.



Fig. S3

Figure S3. Secondary structures of MCMV MCP. The prediction was carried out by three different prediction servers (Jpred Secondary Structure Prediction Server (Cole et al., 2008), PsiPred Protein Structure Prediction Server (Buchan et al., 2010) and Phyre² Protein Fold Recognition Server) and only the results from Phyre² Protein Fold Recognition Server (<u>http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index</u>) are shown. Green helical ribbon: helix; blue arrow: β strand; grey line: loop.

MCMV ·	4	10 20 30 40 50 60	62
Pred.:	-		02
HSV-1:	7	7 ETDIAIPSGISRPDAAALQRCEGRVVFLPTIRRQLTLADVAHESFVSGGVSPDTLGLLLA	66
KSHV : HMM :	6	SIVVNLTSRLFADELAALQSKIGSVLPLGDCHRLQNIQALGLGCVCSRETSPDYIQIMQY evevtlpsrLsaddvakLgkleGavvplptlrrlinlgdvglksvvvkgsepDvlsllaa	65
		erererpersonaaransynteed opproximityd grhe of hygeopsytesiaa	
MCMV ·	63	70 80 90 100 110 MREMTLTIMRVEGNOMILGVPTHGOCYTIRNT-GP-VSWEKGDVLTTLPPVESGE	116
Pred.:			110
HSV-1:	67	Y YRRRFPAVITRVLPTRIVACPLDVGlthaGTVNLRNT-SP-VDLCNGDPISLVPPVFEGQ	124
KSHV :	66	LSKCTLAVLEEVRPDSLRLTRMDPSDNLQIKNVyAPfFQWDSNTQLAVLPPLFSRK	121
	6	iiiitiaiiievephqivitkidvgqgyaikht.gr.idwengdaitiirrvigie	
NOR	117	120 130 140 150 160 170	171
Pred :	117	VTG-LVSVSDWDLVLPWIVPMALATEINQRMMMLALLSLDRSHEEVRAATAQLRVV	1/1
HSV-1:	125	ATDVRLDSLDLTLRFPVPLPSPLAREIVARLVARGIRDLNPSPR-NPGGLPDLNVL	179
KSHV :	122	DSTIVLESNGFDIVFPMVVPQQLGHAILQQLLVYHIYSKISAGAPGDVNMAELdlyTTNV	181
HMM :		earvrLesndlelvfPlvvPtelAreilqkllaralyslarraeeveaavaellrvv	
		<i>180 190 200 210 220</i>	
MCMV :	172	RYRDATLTLP-EITIDDTVLIDMRNVCISLSMIANLSSEVTLAYVRKL-ALEDSNMLLMK	229
HSV-1:	180	YYNGSRLSLLADVQQLGPVNAELRSLVLNMVYSITEGTTIILTLIPR1FALSAQDGYVNA	239
KSHV :	182	SFMGRTYRLDVDNTDPRTALRVLDDLSMYLCILSALVPRGCLRLLTA-LVRHDRHPLTEV	240
HMM :		tyrgrtytlpldltnddallaalrtlvlslsflinlgpelllrlipk.Laledqdellna	
		230 240 250 260 270 280	
MCMV :	230	CQEILGRRMpqvgvgagssgdrndPPARSRTNYNITPTEELNKLTALFVMIRQ	282
HSV-1:	240	LLQMQSVTREAAQLIHPEAPALMQDGERrlplyeaLVAWLTHAGQ	284
KSHV :	241	FEGVVPDEVMRVMFSYLQS	275
HMM :		llqllsaerlsalltmlsq	
		290 300	
MCMV :	283	3 ITDVISEQPAFLVCDVSPDDKSALCI 308	
HSV-1:	285	LGDTLALAPVVRVCTFDGAAVVRSGD 310 Fig S4	
KSHV :	276	S LSSIFNLGPRLHVYAYSAETLAASCW 301	
HMM :		Lgdilnlkpllrvcdfsgdnkvatce	

Figure S4. Multi-sequence alignment of TRI-2 protein homologs across three herpesvirus subfamilies: MCMV M85, HSV-1 VP23, KSHV ORF26 and secondary structure prediction of MCMV TRI-2. Secondary structure prediction of MCMV TRI-2 by the PsiPred Protein Structure Prediction Server (Buchan et al., 2010) (http://bioinf.cs.ucl.ac.uk/psipred/) and are depicted in the second row. Red box: helix; blue arrow: β strand; black line: loop. The Hidden Markov Model (HMM) of the alignment generated by the Pfam server is shown in the last row. Numbers indicate amino-acid residual numbers in each sequence. Dashed lines in the sequences (dotted lines in the HMM model) represent deletions.

		50	60	70	80	90		
MCMV : Pred.:	46	QAITARFF	VRESLGEVEQK	NLGVLMFRLDI	GI-EMPST	LVSLFFLSMV	AENVS	98
HSV-1:	151	RQ-TERLG	EAWGQLMEATA	LGSGRA	ES-GCTRAG	SLVSFNFLVAA	CAASYDARDAA	203
KSHV :	67	QPTYGDFL	VYSQTFSPQEP	LGTFLFSFKQE	DNgSSMDMI	LTPTSLFMLS	GMEAAKA	122
HMM :		rsvyrdfl	vaaktlseeep	lggvLflfkqt	dg.ssvral	lvslsflaaa	rageyaaanaa	
		100	110	120		120	140	
MCMV ·	99	AATKNTLA	ATYGRE-GEAT	RTWLEDGAWRT	HF	VVHPLG-CTN	SITPGATCLIT	149
Pred.:								
HSV-1:	204	DAVRAHVT	ANYRGTRVGAR	LDRFSECLRAM	VHTHVFpHE	VMRFFGGLVS	WVTQDELASVT	263
KSHV :	123	PQTHKVAG	VWYGSGSG	LADFIPNLSEI	MDTGEF-H1	LLTPVGPMVQ	SVHSTFVTKVT	178
HMM :		eavrtivk	alYgssrlear	ldrlvenlral	vesrvF.Hi	fLtpvGplve	nisstelakiT	
MONET .	150	150 COMPOURY	160	170 DI LUDKETVI	DID	180	190 DUVDVI UVDVI	107
MCMV :	150	CSMRGHSI	NMLKTEII	-PLLVPKEIYI	יחדה.	GESTDEIR	EVIEVITIDIN	197
HSV-1.	264	AVCAGPOE	AAHTCHPGror	SAVILPACAEV			FLYLVFTYROR	318
KSHV :	179	SAMKGVGL	ARDEPRAH	VGLTLPCDMLV	DLDESCom	aRREPAGLNV	TIYASLVYLRV	235
HMM :	2,2	avmkGpve	sarkeeas	vklvvPadaFl	DlDael	.grepaegls	vvYlvfvYtqr	200
							-	
	272.54	200	210	220	230			
MCMV :	198	SdRQGRPS	AFVVVSRITHR	RHTLINVLRYRF	RVSRFHFLN	IN		237
Pred.:	210	D DOELCO	UN UTROLDE	DCIEDALEDI		Ustadetees	nn unn dfn l na	276
HSV-1:	236	K-DÖFTCC	VI-VIKSQLPP	RGLEPALERLE	GRERITNII	.ндсеамсрра	phinpdiplag	272
HMM ·	250	r neesvr	lv vfksslae	sellellrelf	sreratnli	r		212
		1.1100011	19.01200190		ororachiri			
					240	250	260	
MCMV :	238			S	ISGYGPST-	GCLGTLQRLG	WF-CSRDSR	264
Pred.:	077	1		C 1 1				120
HSV-1:	3//	laanpqtp	rcsagqvtnpq	[fadrlyrwqpl	LRGRPTART	CTYAAFAELG	MMPEDSPRCIN	436
HMM ·	212				lrarnting	ltygayorlG	tweesedit	501
					rigipeine	reygaverio	LV5555011	
		270	280	290				
MCMV :	265	SGIVASRA	GQLSVVKLEKF	YVDVGPLVEF	293			
Pred.:						Lia SE		
HSV-1:	437	RTERF-GA	VSVPVVILEGV	VWRPGEWRAC	464	iy. 55		
KSHV :	302	HQSLNVKG	TSLPVLVFANF	EAACGPWTVF	330			
HMM :		rseltvkg	vsipvvelegi	vvalgawtec				

Figure S5. Multi-sequence alignment of TRI-1 protein homologs in three herpesvirus subfamilies: MCMV M46, HSV-1 VP19c, KSHV ORF62 and secondary structure prediction of MCMV TRI-1. The secondary structures of MCMV TRI-1 was predicted by PsiPred server (Buchan et al., 2010) (http://bioinf.cs.ucl.ac.uk/psipred/) and are depicted in the second row. Red box: helix; blue arrow: β strand; black line: loop. The HMM of the alignment generated by the Pfam server (Finn et al., 2010) is shown in the last row. Numbers indicate the residual numbers in each protein. Dashed lines in the sequences (dotted lines in the HMM model) represent deletions.



Figure S6. Estimation of molecular boundaries by varying density threshold for display. The averaged triplex density map was shown at three different density threshold values of 3.0 (A), 3.4 (B) and 4.5 (C). The densities corresponding to TRI-2a (purple) begin to be resolved from the rest of densities when the threshold density is above 3.4. This threshold-based estimation was used in conjunction with other available constraints to establish the final boundaries (see text).

Legends for Supplementary Movies:

Movie 1: MCMV Capsid Map from 3-fold View.

Radially color-shaded surface representation of the MCMV capsid reconstruction at ~9.1 Å resolution. One half of the full map is rotated, revealing hexons, pentons, triplexes arranged according to T=16 icosahedral symmetry on the capsid.

Movie 2: Fitting of cryoEM density of MCMV MCP with atomic models of HSV-1 MCPud and with the pseudo-atomic model of MCMV MCPud.

The cryoEM density map is shown in semitransparent gray, then the X-ray model of HSV-1 MCPud (ribbon with red helices) is superimposed, followed by the pseudo-atomic model of the MCMV MCPud (ribbon with purple helices).

Movie 3: Structure of triplex heterotrimer. Related to Fig. 4.

A triplex heterotrimer, is shown first, followed by separation of its three subunits, TRI-1 (magenta), TRI-2a (green), TRI-2b (yellow) individually, to reveal interactions among these subunits.

Movie 4: Interaction between triplex and its adjacent hexons. Related to Fig 5.

A triplex and its three associated hexons are shown together to reveal interactions between triplex heterotrimer and MCP.