Supplementary Information

Characterization of *Medusavirus* encoded histones reveals nucleosome-like structures and a unique linker histone

Chelsea Marie Toner<sup>1†</sup>, Nicole Marie Hoitsma<sup>1,2†</sup>, Sashi Weerawarana<sup>1</sup>, and Karolin Luger<sup>1,2\*</sup>

<sup>1</sup>Department of Biochemistry; University of Colorado at Boulder, 80309 Boulder, Colorado

<sup>2</sup>Howard Hughes Medical Institute, Chevy Chase, Maryland

† These authors contributed equally to this work

\*Author for Correspondence: karolin.luger@colorado.edu

# Supplementary Figure 1. Complete sequence alignment and secondary structure prediction of *Medusavirus medusae* histones.

- Supplementary Figure 2. MM-NLP preparation for Cryo-EM.
- Supplementary Figure 3. Biochemical analysis of *Medusa medusae* octamers and trinucleosomes.

Supplementary Figure 4. Cryo-EM analysis of Native and GraFix MM-NLP207 bp.

- Supplementary Figure 5. Electrostatic surface representation comparison of eNuc to viral NLPs.
- Supplementary Figure 6. Electrostatic surface representation comparison of host *A*. *castellanii* H1.1 and *X. laevis* H1.0 to *Medusa medusae* linker histone H1.
- Supplementary Figure 7. Biochemical analysis of *Mus musculus* and *Medusavirus medusae* histone H1.

Supplementary Table 1. Summary of cryoEM data collection and refinement.

# Medusavirus medusae H2B

### Secondary Structure

Medusavirus medusae H2B&H2A Medusasvirus stheno H2B&H2A Clandestinovirus H2B-H2A Marseillevirus H2B-H2A Marseillevirus H2B-H2A Melbournevirus H2B-H2A Acanthamoeba castelanii H2B&H2A.1 Acanthamoeba castelanii H2B&H2A.2 Acanthamoeba castelanii H2B&H2A.3 Xenopus laevis H2B&H2A

# Xenopus laevis H2B Secondary Structure

### Consensus

virus medusae H2B & H2A Me Secondary Structure

Medusavirus medusae H2B&H2A Medusavirus stheno H2B&H2A Clandestinovirus H2B-H2A Marseillevirus H2B-H2A Marseillevirus H2B-H2A Melbournevirus H2B-H2A Acanthamoeba castelanii H2B&H2A.1 Acanthamoeba castelanii H2B&H2A.2 Acanthamoeba castelanii H2B&H2A.3 Xenopus laevis H2B&H2A

# Xenopus laevis H2B & H2A Secondary Structure

### Consensus

αN

#### Medusavirus medusae H2A Secondary Structure

Medusavirus medusae H2B&H2A Medusavirus stheno H2B&H2A Clandestinovirus H2B-H2A Marseillevirus H2B-H2A Marseillevirus H2B-H2A Melbournevirus H2B-H2A Acanthamoeba castelanii H2B&H2A.1 Acanthamoeba castelanii H2B&H2A.2 Acanthamoeba castelanii H2B&H2A.3 Xenopus laevis H2B&H2A

#### Xenopus laevis H2A Secondary Structure

Consensus

#### Medusavirus medusae H2A Secondary Structure

Medusavirus medusae H2B&H2A Medusavirus stheno H2A&H2A Clandestinovirus H2B-H2A Marseillevirus H2B-H2A Marseillevirus H2B-H2A Melbournevirus H2B-H2A Acanthamoeba castelanii H2B&H2A.1 Acanthamoeba castelanii H2B&H2A.2 Acanthamoeba castelanii H2B&H2A.3 Xenopus laevis H2B&H2A

### Xenopus laevis H2A

Secondary Structure Consensus

										0				· ·	
1 50 ID EHVT	LO I Emitefean	20	30	40	50 I	60 	7	D	80 I 0 0 0 0	90		110 I		120	130 I
SQIDENVT	EMTEYEAE	IESQYSDHSI	DEEEQELG	ARVSPAKSA:	SEKKK	SASLRKKKS	GSPKKKAG	KKPS	GEER	K R <mark>KKNYDS</mark> F	ATEIAKLVS.	AN TV ED KI	OGKKHTRK		MEVLESV
SA			- EQ	QTPN	TT - EQTEQV	κv · · · · · DI	ETAP-QQTEV	K SDVKT	EKKTKK	RKSNSQQN	IT R FV R KV L K	v 1	NGD	AQV SKES	CQMVNA I
								M	TQKETT	RKRDKSVNF	REGERNMLA	21		ISVQTE	LSELSNI
										- MDRVGK	G L E I K R I S -		PKD	ADITKES	LETVNNN
								M	TQKETT	RK RDK SVN F	R L G L RN MLA	21	<mark>H P D</mark>	ISVQTEA	LSELSNI
s s			• TPKKTG	ЭКАРАККРАІ	DTVEGAEKK	TT GGKB	<b>TAATKKVG</b>	KKTAGEGKI	GT <mark>K</mark> GG <mark>K</mark>	G S <mark>K</mark> KNY Q S Y	STFIYKVLK	2 v	· · · · · H P D	TG I SNKA	MA I MN S F
s s			- T P K K T G G	GKA PAKK PAI	DTVEGAEKK	TT GGKB	CTAATKKVGV	KKTAGEGKI	GTKGGK	I S <mark>KKNYQ</mark> SY	STFIYKVLK	2V	H P D	TGISNKA	MA I MN S I
\$ \$			T P K K T G G	SKAPAKKPAI	DTVEGAEKK	TT GGK8	CTAATKKVGV	KKTAGEGKI	GTKGGK	3 SKKNYQ SY	STEIVKVLK	2V	IIPD	TGISNKA	MA I MN S I
				MSDPAI	KS		· · APAAKKGS	(KKAVI-KI)	KKDGKK	O O	ATYVYKYLK	2	HPD	TOTSSKA	MSTMNST
				1							α1				
										_		_			
	_	_	100	100	1.1	in a shi	الالعاد		- 14						le la
SSIDEHVT	EMTE+EAEI	IES+YS+HSI	DE+PKKTG	367 9766 971	DIVEGAEKK	TTSLRGGKS	CTA+TKKVGV	KKT+GEGKI	GTK+GK	RKKNYQSY	+ TFIYKVLK	QV + V EDKI	OGKK+HPD	TGISNKA	M+1+NS+
			- 2						196 1						
αΖ		<u> </u>	α 3		0	ι.	<b></b>								
140	150	160	17	70 1	180	190	200	210	22	0 2	30 3	240	250	26	50
1	150	11 1	1/	i	1	150	200	1	. 1	· 2	1	1	1	1	
ATEMTIVA	N E LAKHQG	RQTLGAGD	RTALAVRO	G S L I A R E P A	TVKALTEM	EKAVLKYQS	SLGRPAKTA	PKKKKATKI	A S A - ME	I D S H V Q P A E	VLAAASESM	2 L E E Q T Q	LPASAAGE	SALEELQT	EGKKTSF
ATEMITVA															
N R R R R R R R R R	N DEAKHQO	RQTLDSRD	RTALAVRO	SGILAREPG	TVKTLIAEG	EKAVLKYNA	N S LG R T A K K A	A P K K K + - K)	AAH ME	IESHVQPAE	VLAAASESM	Q L E E Q T Q	LPAFVAGE	ALEELQT	EGKKISI
VEELGNIC	A LLL RD S TI	RQTLDSRD	RTALAVRO	3G I LA REPG FKD SK <b>E</b>	TVK TLIAEG ILKRALAAG	EKAN LKYN / A LANQKII S /	A SLGRTAKKA AYVPVKEEG-	A PKKK K	алан   ме	I E SHVQ PA E	VLAAASESM	2 L E E Q T Q	L PA FVAGE	ALEELQT	
V E E I GN I C. GKK I SHG <b>A</b> A EK L TKO <b>X</b>	A L <b>L</b> LRD S TI V T <b>L</b> LP - EG	- RQ I LD S RD IKK I LK LN T - TK I IK S S A' K I LRHDA	<sup>2</sup> R TALAVRO . EAALS III / LL <mark>A</mark> AG I ? L - WLL	3G I LA REPG FKD SK E D L Y GKD TD I 9G E	TVK TLIA EG ILKRALAAC LGRHAVGEN LGKH SQDFA	E K A V L K Y N / A L A V Q K II S / ITK A V T R Y Q S .N S	A <b>S L G R T A K K A</b> A Y V P V K E E G - A K E <b>S</b> K E G <b>S</b> F	1A PKKK K) 18	алн ме	IESHVQPAE	V LAAA S E S M	2 L E E Q T Q	LPAFVAGE	ALEELQT	EGKK TS1
V E E I GN I C GKK I SHO <b>N</b> A EK L TKQ <b>N</b> GKK I SHG <b>N</b>	A L L L R D S TI V T L L P - EG N I I I D Q V T L P - EG	- RQ L L D S R D I K K T L K L N T - T K T I K S S A' K T L R H D A - T K T I K S S A'	* RTALAV R : EAALS III / LLAAG I * L - WLL * / LLAAG I	36 I LAREPG FKD SK E D L Y GKD T D I Q G E D L Y GKD	TVK TLIAEG ILKRALAAC LGRHAVGEN LGRHAVGEN LGRHAVGEN	EKAV LKYN / A LAVQKII S / ITKAV TRYG S N S	S L G R T A K K A A Y V P V K E E G S A K E S K E G S F A K E S K E G S F	арккк к. ts	алн ме	I E SHVQ PA E	V LAAASESM	2 L E E Q T Q	LPAFVAGE	ALEELQT	EGKKTS1
V E E I GN I C GKK I SHO A A EK L TKO A GKK I SHG A F E R I A L E A	A L L L R D S TI V T L P - EG N I I I D Q V T L P - EG G R L A RY N K	- RQ ILD S RD IKK ILK LN T - TK IK S SA' K ILRHDA - TK IK S SA' - RN IS S RE	2 RTALAVRO JEANLSTII / LLNAG I ? L - WLL * / LLNAG I IQTAVR I	GGILAREPG FKD SK E DLY GKD TDI QGE DLY GKD LLL PGE	TVK TLIAEG I LKRALAAC LGRHAVGEN LGRHAVGEN LGRHAVGEN LAKHAVSEG	E K AV L K Y N / A L AV Q K II S / IT K AV T R Y G S IT K AV T R Y G S T K AV T R Y Q S	NS LG RTAKKA NYV P VK EEG - SAKESKEGSF SAKESKEGSF SS LG STAASC	A P K K K K - K - K - K - K - K - K -	MAAH ME	I E S H V Q P A E	V LAAA S E SM	2 L E E Q T Q	L PA FVAG E	A L E E L Q T	EGAK
VEELGNIC GKKISHON AEKLTKON GKKISHON FERIALEN	ALLERDSTI VTLLP-EG NIIIDQ VTLLP-EG GRUARYNK GRUARYNK	- RQ I LD S RD IKK I LK LN T - TK I IK S S A' K I LRHDA - TK I IK S S A' - RN I IS S R E - RN I IS S R E	PRTALAVRO JEANLSTII / LLNAG I / LLNAG I / LLNAG I IQTAVR I IQTAVR I	GGILAREPG FKD SK E DLY GKD TDI QGE DLY GKD LLL PGE LLL PGE	TVK TLIAEG I LK RALAAG LG RHAVG EM LG RHAVG EM LG RHAVS EG LAKHAVS EG	E K AV L K Y N / A L AV Q K II S / ITK AV T R Y Q S ITK AV T R Y Q S TK AV T R Y Q S T K AV T K Y Q S	AS LG RTAKKA AYV PVK EEG AK ESK EG S AK ESK EG S S LG S TAA S S LG S TAA S S LG S TAA S	A PKKK K t8 t8 PD E PD E	MAAH ME	I E S H V Q P A E	V LAAA S E SM	Q L E E Q T Q	L P A F V A G E	SALEELQT 	EGAK
VEEIGNIC GKKISHGA AERLTKQA GKKISHGA FERIALEA FERIALEA FERIALEA	A LEL RD STI VTEL P - EG N I I I DQ VTEL P - EG GREARYNK GREARYNK GREARYNK	RQ I LD S RD IKK I LK LN T - TK I IK S SA' - K I LRHDA - TK I IK S SA' RN I IS SRE RN I IS SRE RN I IS SRE	= RTA LAV RO : EAALS III / LLAAG L / LLAAG L / LLAAG L IQTAVR L IQTAVR L IQTAVR L	GG I LA RE PG FKD SK E D LY GKD TD I QGE D LY	TVK TLIAEG ILKRALAAG EGRHAVG EM EGRHAVG EM EGRHAVG EM EAKHAVS EG EAKHAVS EG EAKHAVS EG	EKAVLEYNA ALAVORIISA ITKAVTRYOS ITKAVTRYOS ITKAVTRYOS ITKAVTRYOS	N S L G R T A K K A NY V P V K E E G - G A K E S K E G S F S L G S T A A S G S L G S T A A S G S L G S T A A S G	APKKK K. 18	MAAH ME MS MS MA	S	VLAAASESM	2LEEQTQ	L P A F V A G E K D A G K K E G A K K S G D A	A L E E L Q T 	EGAK
VEEIGNIC GKKISHGA AENLTKQA GKKISHGA FERIALEA FERIALEA FERIALEA FERIALEA	ALLERDSTI VTELP-EG NIIIDQ VTELP-EG GREARYNK GREARYNK GREARYNK SREARYNK	RQ L L S RD IKK I LK L L L T - K L L RH DA - K L L RH DA - K I L S S A <sup>1</sup> - K I I S S RE RN I S S RE RN I S S RE - K S I T S RE	= R T A LAV RO : E A A LS III / L L A G I / L L A G I / L L A G I I Q T A Y R I I Q T A Y R I I Q T A Y R I	GGILAREPG FKD SKE DUY GKD TDI QGE DLY	TVK TLIAEG ILK RALAAC LG RHAVG EN LG RHAVG EN LG RHAVG EN LAKHAVS EG LAKHAVS EG LAKHAVS EG	E K AV L KYN A A L AV Q KII S A ITK AV T R Y Q ITK AV T R Y Q ITK AV T R Y Q ITK AV T KY Q ITK AV T KY Q ITK AV T KY Q ITK AV T KY Q	N S L G R T A K K A XY V P V K E E G - A K E Š K E G S F S L G S T A A S C S L G S T A A S C S L G S T A A S C A K -	K P K K K K K K K K K K K K K K K K K K	MAAH ME MS MS MA MA	S		QLEEQTQ	KDAGK 	ALEELQT 	EGAK
VEEIGNIC GKKISHGA AEKLTKQA GKKISHGA FERIALEA FERIALEA FERIALEA FERIALEA FERIAGEA	ALLERDSTI VTLP-EG NIIIDQ VTLP-EG GRIARYNK GRIARYNK SRIAHYNK	RQ LLD S RD IKK LLK LN T - TK LIK S SA' - K LRHDA TK IK S SA' RN LS SRE RN LS SRE RN LS SRE RS IT SRE O	RTALAVRO JEARLSTIT /LLRAG1 /LLRAG1 IQTAVR1 IQTAVR1 IQTAVR1 IQTAGR1 α3	GGILAREPG FKD SKE DUY GKD TDI QGE DUY	TVE TLIA EG ILE RALAAG LE RHAVE EN LE RHAVE EG LA RHAVS EG LA RHAVS EG LA RHAVS EG LA RHAVS EG LA RHAVS EG	EKAN LEYNA A LANOKIISA ITKAN TRYO ITKAN TRYO ITKAN TRYO ITKAN TRYO ITKAN TRYO ITKAN TRYO ITKAN TRYO	NIGRTAKKA NYVPVKEEG AKEŠKEGŠF SSLGSTAAŠG SSLGSTAAŠG SSLGSTAAŠG SLGSTAAŠG	KAPKKK K. KS	MAAH ME MS MS MS MA MS	S		2 L E E Q T Q	- KDAGK - KGAK - KSGDA	ALEELQT TTK KA TTK KA	EGAK
VEEIGNIC GKKISHOA AEKLTKOA GKKISHGA FERIALEA FERIALEA FERIALEA FERIAGEA	ALLERDSTI VTLEP-EG NIIIDQ VTLEP-EG GREARYNK GREARYNK	RQ         L D S RD           IKK         L K L N T           T K         I K S S A'           - K         L RH DA           T K         I K S S A'           RN         I S S RE           RN         I S S RE           RN         I S S RE           RS         I T S RE	RIALAVRO JEARLSTII /LLRAGL /LLRAGL /LLRAGL IQTAVR1 IQTAVR1 IQTAVR1 IQTAGR1 α3	GGILAREPG FKD SK E DLY GKD TD I	TVE TLIA EG ILE RALAAG EG RHAVG EA EG RHAVG EA EG RHAVG EA EA KHAVS EG EA KHAVS EG EA KHAVS EG EA KHAVS EG	EKAN LEYNZ A LANOKIISZ ITKAN TRYOS ITKAN TRYOS ITKAN TRYOS ITKAN TRYOS ITKAN TRYOS	AND	K P K K K - K / K S	MAAH ME 	S	VLAAASESM	2LEEQTQ	L PA FVAGE KDAGK KEGAK K SGDA QGGK	а L E E L Q T 	EGAK
VEEIGNIC GKKISHGA AENLTKOA GKKISHGA FERIALEA FERIALEA FERIALEA PERIAGEA α2	ALL LRDSTI VTLLP-EG NIIIDQ VTLLP-EG GRIARYNK GRIARYNK SRIARYNK SRIAHYNK	RQ         LDS RD           IKK         LK LN T           TK         IKS SA'          K         LRHDA           TK         IKS SA'           RN         ISS RE           RN         ISS RE           RN         ISS RE           RS         ITS RE           O         O	- RTA LA V RG - RTA LA V RG - LA AG I - L - WEL  	GGILAREPG FKD SK E DLY GKD TDI	TYE TELIA HO IEERA LAAG IGRHAVG EN IGRHAVG EN IGRHAVG EN IAEHAVS EC IAEHAVS EC IAEHAVS EC IAEHAVS EC IAEHAVS EC	EKAYLEYNA A LAYOKIISA ITKAYTRYG NS NS ITKAYTRYG ITKAYTRYG ITKAYTKYG ITKAYTKYG ITKAYTKYG ITKAYTKYG ITKAYTKYG	AKESKEGS AKESKEGS AKESKEGS SLGSTAASC SLGSTAASC SLGSTAASC AK	<ul> <li>A P K K K K.</li> <li>C S</li></ul>	MAAH ME MS MS MA MS 1	S	V LAAA S E SM	2 LEEQ TQ	L PA FVAGE	K L E E L Q T	EGAK
verignic Garishor Alentrop Garishor Ferialer Ferialer Ferialer α 2	A L ARNO	KQLLDRKD IKKÖLKLNT - KLIKSA' - KLIRHDA - KLIKSA' KNISSKE KNISSKE KNISSKE 	- RTA LAV RG           - EA & LS HIII           / LL & AG I           / LL & AG I           / LL & AG I           I Q TA V R - I           I Q TA V R - I           I Q TA V R - I           I Q TA G R - I           X 3	3GILAREPG FKD SKE DIY GKD TD I QGE DIY GKD U PGE LI PGE LI PGE	TV KTLLA RO TI KR LAAG GORHAVG EA GORHAVG EA GORHAVG EA GORHAVG EA LA KHAVS EC LA KHAVS EC LA KHAVS EC QA KHAVS EC QA KHAVS EC	IE KAN LEYNA AL EVORTISA ITKAN TRYG ITKAN TRYG ITKAN TRYG ITKAN TRYG ITKAN TRYG ITKAN TRYG ITKAN TRYG ITKAN TRYG I	NEGRTAKKA NYVPVKEEG AKESKEGS AKESKEGS SEGSTAASC SEGSTAASC AK 126	(S	АААН МЕ МS MS MS MS MS 	5	V LAAASESM	2LEEQ TQ	L PA FVAGE	а L E E L Q T КТТК КА КТТК КА КОТО КТ	EGAK
VEE GNIC GKKISHOR AEKLTKOR GKKISHOR FERIALER FERIALER FERIALER α2 FERIALEA	A L LEDSTI VILP-EG NILIDO VILP-EG GRARYNK GRARYNK GRARYNK SRAHYNK	KQLDSKD IKKLLKIT - TKLKISAA - KLKISAA - KLKISAA - CKLKISAA - CKLKISAA - CKLKISAA - CKLKISAA - CKLSSKA - CKLSSKA	- RTA LAV RC - RTA LAV RC - LAV RC - LAV RC 	GGILAREPG FKD SKE DTY GKD TDI QGE DTY GKD DTY GKD LTL PGE LTL PGE LTL PGE	TV KTLLA ER TV KTLLA ER LORHAVG EA LORHAVG EA LORHAVG EA LORHAVS EC LA KHAVS EC LA KHAVS EC LA KHAVS EC	EK AV L KYNA A L EVORTISA I K AV TRYO I K AV TRYO	SLOSTAASC	(S	AAAH ME	1 = SHVQPAE	V LAAA S E SM	2LEEQTQ	L PAFVAGE 	КТТК КА КТТК КА КТТК КА (GVD КТ 	EGAK EGAK EG+KTSI
VEE GNIC GKK ISHOR AEKLTKOR GKK ISHOR FERTALER FERTALER FERTALER FERTALER	AL LEDSTI VILP-EG NILIDO VILP-EG GRARYNK GRARYNK GRARYNK SRARYNK + RLARYNH α N	KQLDSKD IKKLLKLNT TKLKSAA TKLKSAA KNLSSKE KNLSSKE KNLSSKE KSLTSKE O	RTA LAV RC - RTA LAV RC - EA & LS THT //LLAAG I - (L-VEL /LOTAVR 1 IOTAVR 1	GGILAREPG FKD SKE DTY GKD TDI QGE DTY GKD DTY GKD LIL PGE LIL PGE LIL PGE	TV KTLLA EC TV KTLLA EC LO RHAVG EA LO KH SQO FA LG RHAVG EA LG RHAVS EC LA KHAVS EC CC (	E K AV L KYNZ	SLORTAKKA VYVVVK EG AKESKEGS AKESKEGS AKESKEGS SLOSTAASC AK 126 SSLOSTAASC	(S	ΔΔΑΗ ΜΕ 	S	V LAAAS ESM	2LEEQTQ	■ PAFVAGE • • KDAGK • • KDAGK • • KEGAK • • K SGDA • • • QGGK ■ • • • QGGK ■ • • • • • • • • • • • • • • • • • • •	ALEELQT STIK - KA STIK - KA GVD - KT	EGAK EGAK EGAK
V EE GN IC GKK ISHOR A EKLTKOR GKK ISHOR FERIALER FERIALER FERIALER FERIALER FERIALER	X THE ARGONAL STREET	KQLDS KD IKK LK LN T. T. KK IK SA KN IS SKE KN IS SKE KN IS SKE KN IS SKE O O	RTA LAV RC 26 A LS 111 /LLAGG - 1 /LLAGG - 1 /LTAGG -	GGLLAREFG KKDSK DTYGKD DTYGKD DTYGKD LTLFGE LTLFGE LTLFGE LTLFGE LTLFGE	ΤΥ ΚΤΕ Ι Α Κ ΤΙ Κ ΚΗΑ ΥΟ ΕΑ LO RHA YO ΕΑ LO RHA YO ΕΑ LO RHA YO ΕΑ LA KHA YO ΕΟ LA KHA YO ΕΟ Δ ΚΗΑ YO ΕΟ Λ ΚΑ ΥΛΟΥ Λ ΚΗΑ YO ΕΟ Λ ΚΑ ΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥ	ie κ τν τ κ τν τ κ το τ κ το τ κ το τ κ το τ κ το τ κ το τ κ το	SLORTAKKA VYVVVKEG AKESKEGS AKESKEGS SLOSTAASC SLOSTAASC AK SLOSTAASC	14 FKKK K/ 15	маан ме ма ма ма 1  а 3	1 = SHVQPAE - SHVQPAE - SHVQPAE - C	V LAAS ESM	2LEEQTQ		KTTK KA KTTK KA KTTK KA GVD KT 	EGAK EGAK EGAK
V EE GN IC GKK ISHOR A EKLTKOR GKK ISHOR FERIALER FERIALER α 2 FERIALER	x = 1 = 1 = 1 = 1 = 1 = 1 = 1 = 1 = 1 =	KORLOSKO IKR LKLNT - K LKNT - K LRHDA - K LRHDA TK K K SA - K SSKE K SSKE K SSKE K SSKE - C	RTA LAV RC           CA LAV RC           CA LAV RC           VLL AAO - 1           VLL AAO - 1           VLL AAO - 1           IQTAVR - 1	GULAREFG FKDSKE DEVOKD DEVOKD DEVOKD DEVFOR LTLFOR LTLFOR LTLFOR LTLREFG DEV	TV KTLLAR LARKLAAR LORHAVG EA LORHAVG EA LARHAVG EA LARHAVS EG LARHAVS EG LARHAVS EG LARHAVS EG	Lek XV L K Y N / Lek XV L K Y N / ITK AV TH Y O ITK AV TH Y O	SLORTANS	IS	ΔΑΑΗ ΜΕ 	S	VLAAASESM VLAAASESM 370	2LEEQTQ 3LEEQTQ 388		XTIK KA (TTK KA (TTK KA (OVD KT 	EGAK
V HE CONTC OKK SHOR A KALTKOR GKK SHOR F HEALER F HEALER	A LARNO A LARNO A LARNO A LA LARNO TI LP-EG NI LI LP-EG OR ARVNE GRARVNE GRARVNE GRARVNE SRANNE + RLARVNE + RLARVNE - RANNE -	KQ LLOS KD IKK LLK LNT TK LK KS A' KS LRHDA' TK LK KS A' KS LS KE KS LS KE KS LS KE KS LS KE KS LS KE KS LS KE C C 290 C	RTA LAV RC           CEA LAV RC           CEA LAV RC           VILLAGO - I           VILLAGO - I           IQTAVR - I           I	301LARFG FKDSKT DEYOKD DEYOKD DEYOKD DEYGKD LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTCPGE LTC	12 κτι LA Ν 12 κπ ΔΑΑ 12 κπ 30 μ 10 κπ	LEK XV LEX YA / LA LEX OF US / LA LEX OF US / UK X Y TH X OF US / UK X OF US / U	IN LEGRT ANKAN NYY PYK REG AK ESK RES IN LESK RES IN L	APEKK - K, (S	ΔΔΑΗ ΜΕ 	1 = SHVOPAE 3	VLAAASESM VLAAASESM 370	31660 TQ		ALEELQT TTK KA TTK KA GVD KT 	EGAK
V HE IGN I C GKK ISHOK A HK L TKOK GKK ISHOK F I KIALEK F I KIALEK	α Ν         α           φ         κ           α Ν         α           φ         κ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           α Ν         φ           κ         κ           α Ν         φ           κ         κ	K ( L L D K L) (K K L K L K T T T K L K K K K T T K L K K K K K K K L K K K K K K K K K K K K K K K K K K K		GOLLAREFG FKD SK F FV OKAD TD I OGR DI OGR	11 Κ RA LA A 11 Κ RA LA A 10 K H 30 C A	LE X Y LE X Y A LE X Y LE X Y A X Y LE X Y LE X Y Y K X Y T Y Y A Y K X Y K Y A Y K Y K Y A Y K Y K Y K Y K Y A Y K Y K Y K Y K Y K Y A Y K Y K Y K Y K Y K Y K Y K Y K Y K Y K	NUTREAR AND	арккк - к. (8	ΔΑΑΗ ΜΕ 	1 = SH VQ PA E 	V LAAAS ESM V LAAAS ESM 370 Lõda PV SIIG			ALEELQT TTK - KA TTK - KA GVD - KT GVD - KT STT - KA GVD - KT ST - KA GVD - KT ST - KA ST -	EGAK
VIEIGNIC GKKISHOR AIKLIKOR GKKISHOR FERALE FERALER FERALER σ2	A LARNAGE AND A	KOLDSKO IKK LIKUNT K LIKUSA K LIKUSA K LIKUSA 		301LAREFG FKD SK F FKD SK F FKD GR FKD GR FKD		LEX XX LEX XX A LEX XX LEX XX LEX XX XX XX LEX XX LEX XX	NEGRARKA E E G	APEKK - K. (S. (S. )D E. )D E. )D E. )D E. (S. )D E. )D E. (S. )D E. )D E. (S. )D E. )D E. )D E. (S. )D E. )D E. )D E. (S. )D E. )D E	AAAH ME MS MS MS MS MS MS MS MS MS MS MS MS MS		V LAAAS E SM V LAAAS E SM 370 LODA PV SHOU LODA PV SHOU LODA PV SHOU			390 I G J K K R K R	EGAK IST EGAK EGAK EGAK EGAK EGAK EGAK EGAK EGAK
V ΠΕ [GN I C GKK I SHO Ř A HL I KNO GKK I SHO Ř P I R A LE K P I R	α Ν         α         N         1         LP - EQ         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         1         N         N         1         N         N         1         N         N         N         1         N         N         1         N	KOLDSKID         KOLDSKID           TK         KKRIJKKINT           TK         KKRIJKKINT           TK         KKRIJKKINT           KKRIJKKINT         KKRIJKKINT           KKRIJ         KKRIJKKINT           KKRIJ         KKRIJ           KKRIJ         KKRIJKINT           KKRIJ<	2         C	GULLARFOG FKD SKE FKD SKE TD I OUE TD I OUE TD I OUE TD FRE LE FRE LE FRE LE FRE LE FRE LE FRE LE FRE TD OUE LE FRE TD	11 k R. L. LA KE 11 k R. L. LA KE 10 K H SQD F A 10 K H SQD F A 10 K H SQD F A 10 K H SQ F	LE LAN LE XIN A LAN DELLA LE XIN LA LAN DELLA LE LA LA LA LE LA LA VIENE LE LA VIENE LE	NEGRATAKKA NYY PYKEEG AKESK IGSI SEUSTAAS SEUSTAAS SEUSTAAS IIII SEUSTAAS SEUSTAAS AK IIII SEUSTAAS AK IIIII SEUSTAAS AK IIIII SEUSTAAS AK	харккк - к. (8	ΔΔΑΗ ΜΕ Δ		370 100 AFV SIG 100 AFV SIG 100 AFV SIG 100 AFV SIG 100 AFV SIG 100 AFV SIG	316 E E Q T Q 21 E E Q T Q 21 E E Q T Q 38 5 V E Q E A E 3 V E A E		TTIK KA TTIK KA TTK	EGAK
V ΠΕ GN I C GKK ISING & AIKLIKON GKK ISING & PERIALEN PERIALEN PERIALEN PERIALEN FERIALEN C 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	A LARNO A A L LARNS HALL N T LP - EG (N	KOLLOS KO           YKOLLOS KO           YKOLLOS KO           YKOLLOS KO           YKOLLOS KO           YKOLLOS KO           KIKANAN           YKOLLOS KO           KIKANAN           YKOLLOS KO           KIKANAN           YKOLLOS KO           KIKANAN           YKOLLOS KO           YKOLOS KO           YKOLOS KO		301LAREFG FKD	11 k fill lake 11 k fill lake 10 k fill og fill dake 10 k fill og fill og fill dake 10 k fill og fill og fill og fill dake 10 k fill og fill og fill dake 10 k fill og f	1 E K K 1 E K M A	NUT AT ALL ALL ALL ALL ALL ALL ALL ALL ALL	APEKK - K. (S			V LAAAS ESM 370 100 A F V Stor 100 A L C C 100 A L C C 100 A L C C	31660 TQ 31660 TQ 3160 TQ 3		390 I GTKK RKR STK - KA STK - STK - STK STK - STK - STK - STK STK - STK	400 400 400 1 400 1 400 1 400 1 400 1 400 1 400 1 400 1 400 1 400 400
V ΠΕΙ ΟΝΙ Ο GKK ISHO & A EKL TKO & GKK ISHO & A EKL TKO & GKK ISHO & FIELALE & FIELALE & FIELALE & FIELALE & FIELALE & 22 	A BARNON A BARN	KOLLOS KO           + KOLLOS KO           - TK	RTALLAV RC           RTALLAV RC           VELTAGO 1           *L. WEL *           *. *           *. *           *. *           *. *           *. *           *. *           *. *           *. *		11 K RILLA KE 11 K RILLA KE 10 K RIN OD F A 10 K RIN O	1 = K × K × K × K × K × K × K × K × K × K	NEGRATAKKA NYY PYKEEG AKESKIGS SLOSTAAS SLOSTAAS SLOSTAAS AK 126 121 320 320 320 320 320 320 320 320 320 320				370 100 A PV SHOU 100 A PV SHO	384 394		390 380 380 380 380 380 380 380 38	EGAK
V ΠΕ ΟΝΙ C GKK ISHO R A HKLTKOR GKK ISHO R F HKLALER F HKLALER F HKLALER C 22 T - SSGKNV SAASGNV 		KOLLOS KO           + KE         KK           - TK         TK	RTALLAV RC           REAS LES INT           LELAVES           LELES           LE		17 & T. L. I. A RE 10 R HEN VG EN 10 R HEN V		VERT REAL AND A CONTRACT AND A CONTR				VIAAASESM 370 100 A FV SIG 100			11 11 11 11 11 11 11 11 11 11 11 11 11	EGAK
ν πεισητίς ακκ τη παια ματά τη τη του	A LEAR STATUS N I LID - IG N I LID - IG N I LID - IG N I LID - G G REARYNE G REARYNE S REARYN	Koll Lökkö           Koll Lökkö           TKE ÜkE Kön           TKE ÜkE Kön           TKE ÜkE Kön           TKE ÜkE Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön Kön           Kön Kön           Kön Kön <td>2         2           2         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3</td> <td></td> <td></td> <td>1 A KAN ( KAN ) A KAN ) A KAN ( KAN ) A KAN ( KAN ) A KAN ) A</td> <td>S LEG FLAKEA AK E SK HOS I AK E SK HOS I S LOS TAA SC S L</td> <td></td> <td>AAAH AE           MS           MS           MAA           MS           MS           MAA           MS           MAA           MS           MAA           MS           MS           MAA           MS           MAA           MS           MAA           MS           MS     <!--</td--><td></td><td>370 100 a PV SIG 100 a PV SIG 100 a I V SIG 100 a I V SIG 100 a I V SIG 100 a I LG G 100 a I</td><td>321 E E O TO 32 E E O TO 33 E O TO 33 VOC V A E A 34 O TO 34 O</td><td></td><td></td><td>EGAK EGAK EGAK EGAK EGAK EGAK EGAK EGAK</td></td>	2         2           2         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         2           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3           3         3			1 A KAN ( KAN ) A KAN ) A KAN ( KAN ) A KAN ( KAN ) A KAN ) A	S LEG FLAKEA AK E SK HOS I AK E SK HOS I S LOS TAA SC S L		AAAH AE           MS           MS           MAA           MS           MS           MAA           MS           MAA           MS           MAA           MS           MS           MAA           MS           MAA           MS           MAA           MS           MS </td <td></td> <td>370 100 a PV SIG 100 a PV SIG 100 a I V SIG 100 a I V SIG 100 a I V SIG 100 a I LG G 100 a I</td> <td>321 E E O TO 32 E E O TO 33 E O TO 33 VOC V A E A 34 O TO 34 O</td> <td></td> <td></td> <td>EGAK EGAK EGAK EGAK EGAK EGAK EGAK EGAK</td>		370 100 a PV SIG 100 a PV SIG 100 a I V SIG 100 a I V SIG 100 a I V SIG 100 a I LG G 100 a I	321 E E O TO 32 E E O TO 33 E O TO 33 VOC V A E A 34 O TO 34 O			EGAK EGAK EGAK EGAK EGAK EGAK EGAK EGAK
v με ση τε	A BLARDY I LIP - EG IN I LIP - EG	KOLLOSKO         KOLLOSKO           - TRE         KRESA           - TRE         KRESA           - TRE         KRESA           - RE         KRE           - RE         KRE <td>FIG. 14. WE THE THE THE THE THE THE THE THE THE TH</td> <td></td> <td>11 K K K L A K K K K K K K K K K K K K K K K K K K</td> <td></td> <td>SECRETERES AND SECRETERES AND SECRETERES SECRETERES SECRETERES SECRETERES AND SECRETERES AND SECRETERES SECRET</td> <td></td> <td>κλλή με           κλλή με           κλή με           κα           κα           α           α           α           3550           1           1           1           α           3550           1</td> <td></td> <td>370 100 A 17 V LAAAS E SM 370 100 A 17 V SHO 100 A 12 C V SHO 100 A 10 C V</td> <td></td> <td>раруасы Краска Краска Кола</td> <td></td> <td>EGAR</td>	FIG. 14. WE THE THE THE THE THE THE THE THE THE TH		11 K K K L A K K K K K K K K K K K K K K K K K K K		SECRETERES AND SECRETERES AND SECRETERES SECRETERES SECRETERES SECRETERES AND SECRETERES AND SECRETERES SECRET		κλλή με           κλλή με           κλή με           κα           κα           α           α           α           3550           1           1           1           α           3550           1		370 100 A 17 V LAAAS E SM 370 100 A 17 V SHO 100 A 12 C V SHO 100 A 10 C V		раруасы Краска Краска Кола		EGAR

13

419

432 321

268 268 308

				23	3		
410	) 420	430	440	450	460		
73 KKIAKI	I AAAKKATGAB	KVVKKKSGSTF	 skttgksvtki	L CASSRKVASA			
88 KAKK - /	TOVEPERTIN	KVVKKKSGSA)	SKTAGKSVTKI	CTSTRKASA -			
76 DEIPAI	LVAQEK P	RVQKLKESKKI		<ul> <li>кокулел</li> </ul>	· · · · · · · · · · · · · · · · · · ·		
					PR BOCMEBON		
22	S P K K B	AAPKKKKAASI	QK - KSLSDKELAI	CLINKBUARI	EKEQUMATOT		
	S РКК В		OK - KSESDKELA				
22	S РКК »		QK - KSLSDKELA	CLTKKELAKY	EK EQGMS PGY		
22		ллрккккллs) 	QK - KSLSDKELAI	CLTKKELAKY	EKEQGMSPGY		
22	- S P K K 8 S P K K 8 S Q E Y	лл ркккклл s)	QK - KSLSDKELA	CLTKKELAKY	EKEQGMSPGY		
22	S PKK	ллрккккллs)	QK - KSLSDKELA	(LTKKELAKY)	EK EQGMS PGY		

K + 1 + + + + S + KK + + + K + V + KKK + + + S KQK SK + + G + + + + AK L + K + AK + EK EQGMS FGY

а.

# Medusavirus medusae H4

### Secondary Structure

Medusavirus medusae H4&H3 Clandestinovirus H4&H3 Marseillevirus H4-H3 Melbournevirus H4-H3 Acanthamoeba castelanii H4&H3.1 Acanthamoeba castelanii H4&H3.3 Acanthamoeba castelanii H4&H3.3 Xenopus Jaevis H4&H3

#### Xenopus laevis H4

### Secondary Structure

Consensus

#### Medusavirus medusae H3 Secondary Structure

Medusavirus medusae H4&H3 Clandestinovirus H4&H3 Marseillevirus H4-H3 Melbournevirus H4-H3 Acanthamoeba castelanii H4&H3.2 Acanthamoeba castelanii H4&H3.3 Xenapus Jaevis H4&H3

#### Xenopus laevis H3 Secondary Structure

Consensus

#### Medusavirus medusae H4 Secondary Structure

Medusavirus medusae H4&H3 Clandestinovirus H4&H3 Marseillevirus H4-H3 Melbournevirus H4-H3 Acanthamoeba castelanii H4&H3.1 Acanthamoeba castelanii H4&H3.3 Xenopus Jaevis H4&H3

#### Xenopus laevis H3 Secondary Structure

Consensus



								α	N	α :				
								0					0	
	140	150	100	170	100	100	200	210	220	220	240	250	200	
	140	120	160	170	180	190	200	A 210 A	220	230	240	250	260	
	1	1	<u>1</u>	1	<u>L</u>	· _	1	<u> </u>		r		I Ľ		
93		P P	KE R A I <mark>K</mark> K	S	K SAVK <mark>K</mark> VAEF	KKAAIK S <mark>K</mark> akk.	AAT GVKK PH	R FR PGT T AKRL S	KKEQKL S ST	KTTVRRA PEGR	IVRTIA - SL	SSADSMRFS	AN <mark>AVDL LQ</mark> QG I <mark>E</mark>	LY 195
96					K <mark>Q</mark> VAR <mark>K</mark> ETTI	PKKVEKAAAET	KK PE I <mark>KK PH</mark>	R Y R PGT VAL R I	н <mark>к</mark> е <mark>0к</mark> кт	DSIV <mark>R</mark> T S <mark>P</mark> LR	L M <mark>R E</mark> R L N T DI	AEGQKGFHL S	P SAVAAMSE I VE	os⊤ 215
103	s		GA <mark>K</mark> G SV -	FL S C	R <mark>Q</mark> <mark>K</mark> G	AG SA	б <mark>т б</mark> 5	E T N SQ <mark>E</mark> V	RSQ MKST	CLIIPKER ERT	MAKEISK	- KE GHDVH I A	E A <mark>A</mark> L DM <mark>L Q</mark> V I V <mark>E</mark>	sc 180
103	s		GA <mark>K</mark> G SV -	FL S C	R <mark>Q</mark> <mark>K</mark> G	AG <mark>S</mark> A	б <mark>т б</mark> ѕ	E T N SQ <mark>E</mark> V	R SQ MK ST	CLIIPKERERT	MAK <mark>EI</mark> SK	- KE GHDVH I A	E A <mark>A</mark> L DM <mark>L Q</mark> V I V <mark>E</mark>	sc 180
98	5 5 5 5 5 A A A 5 <mark>6</mark>	G P P SGE T T S PI	RRGR <mark>K</mark> SSTP	T KNA SV S S PR S	R QMAR K ST G/	AAR P <mark>A</mark> T SSS <mark>S</mark> S	S <mark>T</mark> P PE <mark>KK</mark> KRI	K F <mark>R PG</mark> AK <mark>AL RE I</mark>	RKYQR T	DLLI PKMPFAR	LIKEVA - SN	/ SYGLE PF <mark>R</mark> WQ	AQ <mark>A</mark> LLALQE SAE	uy 248
98	QT AR K ST G	KAPRKMAS		••••к•••••			L A <mark>ggvkk ph</mark> i	R FR PGT VAL RE I	RKYQK ST	DLLIRKA PEQR	LVRELA-QD	КТG · · · G · · ·	· · · · · · · · · C B	egi 196
98	QT ARK ST G	K A PR K MA S		• • • • • • • • • • • •		S <mark>A K</mark> K <mark>S</mark> A	L A <mark>ggvkk ph</mark> i	R FR PGT VALRE I	RKYQK ST	DLLIRKAPEQR	LVRELA - QD	KTD L R FQ	A SAVAALQEAGE	A¥ 203
98	QT AR K ST G G	KA PRKQLA		· · · · · T · · · · ·		KAARK SAP	AT <mark>GGVKK PH</mark>	RYRPGTVALREI	RRYQK ST	ELLIRKL PEQR	LVREIA - QD	KTD LRFQ	S S <mark>AVMALQE</mark> A SE	AY 209
				<b>.</b>				α	N		α1		0	-

والمالية ويرجعه والطرا العثران المتلافة والمتلج والمرجع والمرجع ومحمد والمحم
+TARK STGGKA PRKMA SSP++RGAKG SVPTKNA S+L SPR+RQ+ARKGT++K+SAK SK+SA+TGGVKK PHR FR PGTVAL RE I RKYQKMK STDLL I RK+PF+RL VRE I AKQD FKT GGHD+R FQA SAVDAL QE I VE SY

255 253 216

216 286

231 244

239

	α 2	0	)	α3	)		163
		280	290	300	310	320	330
196	ML D <mark>L</mark> MK	- NAALAAK	QAKRMTL MGKI	I DIDLIQTAN	I H <mark>e</mark> mi de Aha.	I AKLANT S SAG FAF	I RRVTKKE
216	LLKLHO	Q SQKVA SH	NGD <mark>r</mark> I T I F P <mark>k</mark> i	VKLAADLT	GVKKYL · · ·		
181	TYRLLE	- KALVITY	SG <mark>KR</mark> TRVT S <mark>KI</mark>	ET <mark>a</mark> fmle	HG PL		
181	T VR LLE	- K <mark>a</mark> l V I T Y	SG <mark>KR</mark> TRVT S <mark>KI</mark>	ET <mark>A</mark> FMLE	HG PL · · · ·		
249		- DAN <mark>L</mark> CC I	HA <mark>KR</mark> VTIMT <mark>KI</mark>	DIQL <mark>A</mark> RRIE	GPREALF		
197	PRG <mark>LFE</mark>	DTNLCAI	HA <mark>KR</mark> VTIMP <mark>KI</mark>	<mark>d I QL A</mark> R R I F	GERA		
204	L VG <mark>L FE</mark>	DTNLCAI	HA <mark>KR</mark> VT I MP <mark>KI</mark>	<mark>DI QL A</mark> RR I F	GERA		
210	L VG <mark>L FE</mark>	- DT NLCG I	R-D Clamp	<mark>di</mark> ql <mark>a</mark> rrif	GERA		
	α2	)	-	α3	136		
	ala.	-	المعاد	L L			

LVGL FE QDANL CA I HAKR VT I MPK DI QLARR I RGERA + L+HAAKLANT SSAG FARR VT KKE

b.

#### c.

#### Medusavirus medusae H3 Secondary Structure

Medusavirus medusae H3&H4 Medusavirus stheno H3-H4 Acanthamoeba castelanii H3.1&H4 Acanthamoeba castelanii H3.2&H4 Acanthamoeba castelanii H3.3&H4 Xenopus laevis H3&H4

#### Xenopus laevis H3 Secondary Structure

#### Consensus

#### Medusavirus medusae H3 & H4 Secondary Structure

Medusavirus medusae H3&H4 Medusavirus stheno H3-H4 Acanthamoeba castelanii H3.1&H4 Acanthamoeba castelanii H3.2&H4 Acanthamoeba castelanii H3.3&H4 Xenopus laevis H3&H4

# Xenopus laevis H3 & H4 Secondary Structure

#### Consensus

#### Medusavirus medusae H4 Secondary Structure

Medusavirus medusae H3&H4 Medusavirus stheno H3-H4 Acanthamoeba castelanii H3.1&H4 Acanthamoeba castelanii H3.2&H4 Acanthamoeba castelanii H3.3&H4 Xenopus laevis H3&H4

#### Xenopus laevis H4 Secondary Structure

Consensus









Medusavirus medusae H1 Medusavirus stheno H1 Clandestinovirus H1/H5 Acanthamoeha castelanii H1 Xenopus laevis H1 Gallus gallus H5

> Gallus gallus H5 Secondary Structure



10

1 MEG

1 мте-

20

30

40

50

- MFRN STYG PLV I ANLKHLNE PTOSTADA. I E CVNRL FIDK I SA SYKRAVVRA TRKOLEN OD AR SG - DGNN SV PI E TKKNVH PSTKRMVLL I OSKDEK - I ASLKTING FLEERWS - - TOKTGRRFVR SALIKOVEKK I FTORK-

α2

60

MP - KN I YA PLIAR ALKHLDD PAG STAE ALADC YNRL FHDE IT AGYR RGYLR ALK SGLEK GDL YR SG

80

· MP· KVTHASMYKDALKDENE PT BS VIA I IERVAKLYEGKET STYERE I IKA IR DOVLAGT VREG..... SRIR FPTKD.... DEDRELTVEEELIMIACREGDMY PORLAK 103

ß 1

···· SA PAAK PRESKASKKSTDIIPKU SDMILDU VQAEKSES GUSSAGS IOKY IKNNUT ·· VUENAD SQIKLSI UKU KUTKU KUTKOVGASOS FULAKADEVKK IAKKPKKE ··· I KKAV SPKKAAK UKKAK

α3

1 MTE SLVL SPARKKKVKA SRR SA SHPT Y SEMIAAN I RAEK SRG SS ROS LOKY I K SHYK - - VGHNADLQIKL SURRLAAN V KOTKOVGA SG SP<mark>RLAK SD</mark>KKKKR S DGKKKKA - - - VRR ST SPKKAAR PRKAR S 130

100

- AT YR I AKCKIKEK PK

110

120

KRYRL ST SGDAISPPRELT PEEELIMTGCHGDGMLPDHLSR 107

- RRYRLL PPGGAID PPPRELT FEEELIMIGCHGDGMF PDHLAR 109

130

- RKKVRS 98

usavirus medusae H1	$\beta 3 \qquad \beta 4 \qquad 171$													
Secondary Structure	140	150	160	170	180	190	200	210	220	230	240	250	260	
usavirus medusae H1	108 g evde	I MSDDELEE	 	I MEQ <mark>GYL KL</mark> G <mark>S</mark> C	I PEGFEV <mark>Y</mark> K <mark>WT</mark>	I A <mark>ka</mark> rkt	I Y S <mark>Y</mark> V <mark>PGD</mark> E Y SD	 A - S <b>I</b> ACY <mark>L</mark> L						17
edusavirus stheno H1	109 DEVDY	· · MDDDELEE	•••••	MDQ <mark>GYL</mark> KLG <mark>S</mark> C	PDG FE VYK WT	SKARKT	YC YV PGDE Y SE	A - SIDC FLL						1
andestinovirus H1/H5	99 SALNPKPAK	PVIAKR	SK P <mark>K</mark> P <mark>K</mark> PK PI	K STNAT VT T ST	PQ T SL VWT	WQ FHDNG W	CNYH <mark>P</mark> DA SKVV	E AT Y QE Y <mark>L</mark> R N I	PGITDVR SVQ:	SGE WHYL VD F	RQMTQQNIQHI	ENHTVRRIRRY	/NIPL SDTE FRKR FQ	22
amoeba castelanii H1	104 AFADD	MDDEEYEE	LQ SL I	GR <mark>GY</mark> I <mark>KL</mark> V SD	GAYR WT	DR <mark>A</mark> G <mark>K</mark> I	FIYMPGDVHGA	S - FIERL PA						16
Xenopus laevis H1	128 SPAKAK	PKVAEKKVKK	PKK <mark>K</mark> PA P <mark>S P</mark> I	Каккт <mark>к</mark> т	VR AK PV <mark>W</mark> A	SKAKKA	K P SK PKAKA S P	· · · · · · · · · · · · · · · · · · ·	SGRKK					19
Gallus gallus H5	131 рак	PKATAR F	CARK <mark>K</mark> SRA <mark>SP</mark> I	ККАКК Р <mark>К</mark> Т	VK AK SRKA	SKAKKV	KR SK <mark>P</mark> RAK SGA	RK 5	S PKKK					1
Gallus gallus H5									190					
Secondary Structure														
Consensus	i di se	a se		a de la de	ألاسها	11. J	a dha an	o de la ferra						
	SE+D+KPKK	PKM+D+E1EE	+ K K K PK P S +	+V +OV +VI O ST	$\mathbf{P} + \mathbf{C} = \mathbf{F} + \mathbf{V} + \mathbf{W}\mathbf{T}$	SKA+KTGW	V + + V + PG + + V S A	AASI +CVL+K	C + V V V P SVO	SCEWEYLVDE	ROMTOONLOW	ENHTVERIER	INT PLADTE ERVREC	<i>.</i>

Medusavirus Medusavin Clandestind Acanthamoeba Xeno

Medusavirus

Gall Secondo



# Supplementary Figure 1. Complete sequence alignment and secondary structure prediction of Medusavirus medusae histones.

Viral histone dimer pairs (or doublets) H2B-H2A and H4-H3 were aligned against A. castellanii and X. laevis histones using HHPRED's multiple sequence alignment tool, ClustalΩ. Conservation

d.

of each residue within the alignment is represented by blue shading, where darker blue signifies a greater conservation. Known  $\alpha$  helices of *X. laevis* are shown in dark colored tubes (H2B-red, H2A-yellow, H4-green, H3-blue). Predicted  $\alpha$  helices of MM (light colored tubes) were generated using HHPRED's Quick 2D prediction web server.

- (a) Complete sequence alignment of H2B and H2A viral histones (*Mamonoviridae* and *Marseilleviridae* families) against viral host histones *A. castellanii*, and *X.laevis*.
- (b) Complete sequence alignment of H4 and H3 viral histones (excluding *M. stheno* doublet) against viral host *A. castellanii*, and *X, laevis*.
- (c) Complete sequence alignment of H3 and H4 viral histones (including *Medusavirus stheno* H3-H4 doublet) against viral host *A. castellanii*, and *X. laevis*. This differs from previous alignment in order of histone pairs (H3-H4 instead of H4-H3).
- (d) Viral putative linker histone H1 was aligned against *A. castellanii, X. laevis*, and *Gallus gallus* H5. Known α helices of *X. laevis* H1 are shown in dark grey colored tubes. Predicted α helices of MM (pink) were generated using HHPRED's Quick 2D prediction web server.
- (e) Isoelectric point (pI) of each predicted and known linker histone H1.
- (f) Heat map comparing percent identity of predicted viral linker histone H1/H5 and eukaryotic H1/H5 sequences. MM-putative H1 is outlined in black.

Related to Figure 1. Source data are provided as a Source Data file.



Supplementary Figure 2. MM-NLP preparation for Cryo-EM.

Sucrose gradient sedimentation and GraFix of MM-NLP with 207 bp DNA (MM-NLP<sub>207W</sub>). Fractions of each were analyzed by 4-12% SDS-PAGE stained with BlazinBlue (protein visualization) and 5% Native-PAGE stained with SYBRGold (DNA visualization) to determine composition of particles. Experiment was repeated independently more than three times with similar results. Related to Figure 2 and 3. Source data are provided as a Source Data file.



Supplementary Figure 3. Biochemical analysis of *Medusa medusae* octamers and trinucleosomes.

- (a) Thermal shift stability of MM and eukaryotic octamer utilized in formation of NLP. The raw relative fluorescence units were normalized for plotting (n=1).
- (b) MM tri-nucleosomes (MM-tri-NLP) and eukaryotic tri-nucleosomes (eNuc-tri) on LE DNA. Experiment was repeated independently more than three times with similar results.
- (c) (d) Representative AFM topography images of (c) eNuc-tri and (d) MM-tri-NLP, white squares represent particles shown in Figure 2G. Scale bar = 500 nm.

Related to Figure 2. Source data are provided as a Source Data file.





Supplementary Figure 4. Cryo-EM analysis of Native and GraFix MM-NLP<sub>207 bp</sub>.

(a) Raw micrograph of GraFix MM-NLP<sub>207</sub>, 2D class averages generated from represented dataset (scale bar = 140 Å), 3D structure of MM-NLP with local resolution map, FSC curve, and CryoSPARC data processing flow chart. An FSC cutoff of 0.143 was used to determine resolution. Bump in FSC curve at ~5 Å is due to the flexible free-DNA arm.

-n/4 0 n/4 Azimuth π/2

(b) Raw micrograph of native MM-NLP<sub>207</sub>, 2D class averages generated from represented dataset (scale bar = 80 Å), 3D structure of MM-NLP with local resolution map, FSC curve, and CryoSPARC data processing flow chart.

Related to Figures 3, 4 and 5.



b. <u>eNuc</u>

MM-NLP

MV-NLP











Disk-view 1

Disk view -

Bottom view

Supplementary Figure 5. Electrostatic surface representation comparison of eNuc to viral NLPs.

- (a) Charged surface representation of histones from the eNuc (PDB ID: 3LZ0), MM-NLP and Melbournevirus-NLP (MV-NLP; PDB ID: 7N8N) in different orientations.
- (b) Oblique view of three nucleosomes to highlight contributions of the MM H2B L1 loop and the MM H3  $\alpha$ C helix
- (c) Charged surface and (d) cartoon representation of MM (H3-H4)<sub>2</sub> tetramer with 60 bp of DNA.

Related to Figure 5.



Supplementary Figure 6. Electrostatic surface representation comparison of host *A*. *castellanii* H1.1 and *X. laevis* H1.0 to *Medusa medusae* linker histone H1.

Charged surface representation of *A. castellanii* H1.1, *X. laevis* H1.0, and MM-H1 with rotational views. Coordinates for *X. laevis* H1 were acquired from 5NL0. Coordinates for *A. castellanii* H1.1 and MM-H1 were determined through AlphaFold. Related to Figure 6.



Tris Acetate Native-Page

Supplementary Figure 7. Biochemical analysis of *Mus musculus* and *Medusavirus medusae* histone H1.

- (a) Purified MM-putative linker histone H1 (MM-H1) and *Mus musculus* H1.0 (eH1.0).Experiment was repeated independently more than three times with similar results.
- (b) CD spectra of purified eH1.0 (blue) and MM-H1 (pink). For each H1 protein, five replicate CD spectra were averaged, baseline-corrected for signal contributions by the buffer.
- (c) Secondary structure estimation based on experimental CD data (shown in b), using DichroIDP.
- (d) Fluorescence polarization of MM-H1, eH1.0 and Amoeba H1.1 with fluorescently labeled25-mer DNA. Data points shown are the mean (SD as error bars, n=3).
- (e) Gel shift of MM-H1 or eH1.0 with MM-tri or eNuc-Tri, analyzed on a Tris-Acetate gel (n=1), stained with EtBr (to visualized DNA, left) and Instant Protein Stain (to visualize protein, right).

Related to Figure 7. Source data are provided as a Source Data file.

# Supplementary Table 1. Summary of cryoEM data collection and refinement.

	MM-NLP <sub>207</sub> Crosslinked (EMDB- 42053) (PDB- 8UA7)	MM-NLP <sub>207</sub> Native (EMDB-45981)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Electron exposure (e-/Å <sup>2</sup> )	50	46.29
Defocus range (µm)	0.6-1.7	0.8-2.2
Pixel size (Å)	0.97	1.017
Symmetry imposed	C1	C1
Initial particle images (no.)	2,270,549	545,861
Final particle images (no.)	159,734	72,786
Map resolution (Å)	3.3	4.9
FSC threshold	0.143	0.143
Map resolution range (Å)	2.8-6.0	4.5-8.5
Refinement		
Initial model used (PDB code)	1AOI	
Model resolution (Å)	3.25	
0.143 FSC threshold		
Model resolution (Å)	3.77	
0.5 FSC threshold		
Map versus model cross-correlation	0.77	
Model resolution range (Å)	2.8-6.0	
Map sharpening B factor (Å <sup>2</sup> )	-51.7	
Model composition		
Non-hydrogen atoms	11265	
Protein residues	778	
Nucleotide	260	
Ligands	0	
<i>B</i> factors (min/max/mean)		
Protein	28.44/229.15/71.75	
Nucleotide	59.20/319.86/127.66	
R.m.s. deviations		
Bond lengths (Å)	0.004	
Bond angles (°)	0.676	
Validation		
MolProbity score	2.76	
Clashscore	12	
Poor rotamers (%)	5.54	
Ramachandran plot		
Favored (%)	88.98	
Allowed (%)	10.63	
Outliers (%)	0.39	