NMR STRUCTURE OF THE N-TERMINAL COILED COIL DOMAIN OF THE ANDES HANTAVIRUS NUCLEOCAPSID PROTEIN

Yu Wang,^{1#} Daniel M. Boudreaux,^{2#} D. Fernando Estrada,¹ Chet W. Egan,¹ Stephen C. St. Jeor,² and Roberto N. De Guzman^{1*}

Department of Molecular Biosciences¹, University of Kansas, Lawrence, KS 66045 USA, and Department of Microbiology and Immunology², University of Nevada, Reno, NV 89557 USA

Running head: Andes hantavirus nucleocapsid coiled coil

			~~~	w		MA	~~~	w	MA-	W	w					~~~			~~~	N
		10			20			30		4	0		50			60			70	
AND	MSTLQEL	QEN I	TAH	<u>IQQ</u> I	'VTA <mark>I</mark>	RQK	LKDA	EKA (	/EVD	PDDVN	I <mark>K</mark> ST	LQNR	<b>R</b> AA V	/STL	ETKI	GEL	K RQI	LADL	VAAQI	КL
SNV	MSTLKEV	<mark>2</mark> DN1	TLH <mark>E</mark>	<mark>2</mark> QQI	'VTA <mark>I</mark>	<mark>R</mark> QK	LKDA	ERA 1	/ELD	PDDVN	I <mark>K</mark> ST	LQSR	RAA 1	/SAL	ETKI	GEL	KRE:	LADL	IAAQI	КL
LAN	MSNLQEV	<mark>2</mark> EGI	TLH	<u>IQQ</u> I	'NAV <mark>I</mark>	<mark>R</mark> QK	LKDA	EKA (	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAA V	/SAL	ENKI	AEL	<mark>K</mark> RQI	LADL	VAAQI	КL
MUL	MSNLKEV	<mark>2</mark> eni	TVH <mark>E</mark>	<mark>.</mark> QQI	'VAA <mark>I</mark>	<mark>k</mark> QK	LKDA	ERTN	1EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAAV	/SAL	ETKI	GEL	K RQI	LADL	VAAQI	КL
BAY	MSTLKEV	<mark>q</mark> en i	TVH <mark>E</mark>	<mark>2</mark> QQI	'VTA <mark>I</mark>	<mark>R</mark> QK	LKDA	ERT\	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAA 1	/SAL	<mark>e</mark> tki	GEL	K RQI	LADL	VAAQI	КL
RIO	MSNLQEV	<mark>2</mark> EGI	TLH <mark>E</mark>	<mark>2</mark> QQI	'NAV <mark>I</mark>	R <mark>Q</mark> K	LKDA	EKA1	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAA V	/STL	<mark>E</mark> NKI	AEL	KRQI	LADL	VAAQI	КL
ELM	MSNLKEL	2DN I	TAH <mark>H</mark>	<mark>.</mark> QQI	'VTA <mark>I</mark>	<mark>R</mark> QK	LKDA	EKA (	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAA V	/SAL	ETKI	GEL	KRQI	LADF	VTSQI	ХL
NYV	MSTLKEV	<mark>2</mark> DN1	TLH	ZQQI	'VTA <mark>I</mark>	<mark>R</mark> QK	LKDA	EKA I	/EVD	PDDVN	I <mark>K</mark> ST	LQGR	RAA V	/SAL	ETKI	GEL	K RQI	LADL	IAAQI	ХL
TUL	MSQLKEI	<mark>2</mark> EEI	TRH	2001	VIA <mark>I</mark>	<mark>R</mark> QK	LKDA	EKT\	/EAD	PDDVN	I <mark>K</mark> ST	LQSR	RAA V	/SAL	<mark>e</mark> dki	ADF	K RQI	LADL	vssqi	ĸм
ISL	MSQLREI	<mark>2</mark> EEI	TRH	<mark>.</mark> QQI	VIA <mark>I</mark> AIV	r _{ok}	LKDA	EKT'	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RAAV	/STL	EDKI	LADF	<mark>K</mark> RQI	LADA	ISRQI	ĸм
PHV	MSQLREI	<mark>2</mark> EE I	TRH	<mark>.</mark> QQI	VIA <mark>I</mark> AIV	<mark>R</mark> QK	LKEA	ERT'	/EVD	PDDVN	I <mark>K</mark> ST	LQSR	RSAV	/STL	<mark>E</mark> DKI	AEF	K RQI	LADV	ISRQI	ĸм
PUU	MSDLTDI	<mark>Q</mark> EDI	TRH	<mark>2</mark> QQI	'VVA <mark>I</mark>	<mark>R</mark> QK	LKDA	ERA I	/EVD	PDDVN	I <mark>K</mark> NT	LQAR	7TQ <mark>Q</mark>	/SAL	EDKI	LADY	K RRI	MADA	VSRKI	ХM
TOP	MSNLKDI	<mark>Q</mark> DE I	TRY	<mark>2</mark> QQI	'IVA <mark>I</mark>	R <mark>Q</mark> K	LRDA	EKT'	/EVD	PDDVN	I <mark>K</mark> NT	LQAR	RQT V	/SAL	<mark>E</mark> DKI	LADF	KRQI	LADH	VSRQI	ХM
KHA	MSNLKDI	<mark>Q</mark> DE I	TRY	<mark>.</mark> QQI	'IVY <mark>I</mark>	<mark>R</mark> QK	LRDT	EKA (	/EED	PDDVN	I <mark>K</mark> NT	LQAR	RQT V	/SAL	EDKI	LADF	KRQI	LADH	VSRQI	ХM
SEO	MALCEEI	<mark>2</mark> RE I	SAH	GQI	VIA	R <mark>Q</mark> K	VKDA	EKQ3	EKD	PDDFN	I <mark>K</mark> RA	LHDR	ESV/	ASI	<mark>Q</mark> SK1	DEL	K RQI	LADR	IAAGI	ĸN
THA	MATMEEL	<mark>2</mark> RE I	NAH	GQI	'VVA <mark>I</mark>	<mark>R</mark> QK	VKDA	EKQ3	EKD	PDDFN	I <mark>K</mark> RA	LHDR	ESV/	ASI	<mark>Q</mark> SK1	DEL	K RQI	LADR	IASG	КT
HTN	MATMEEL	<mark>2</mark> RE I	NAH	GQI	VIA <mark>I</mark> AIV.	<mark>R</mark> QK	VRDA	EKQ3	EKD	PDELN	I <mark>K</mark> RT	LTDR	EGV!	AVSI	<mark>Q</mark> AK I	DEL	KRQ [	LADR	IATG	ΧN
DOB	MATLEEL	<mark>0</mark> KE I	NSH	GQI	VIA <mark>I</mark>	R <u>Q</u> K	VKDA	EKQ3	EKD	PDDLN	I <mark>K</mark> RT	LSDR	ESV/	<u>\Q</u> SI	<mark>0</mark> SK1	DEL	<mark>r</mark> rqi	LADR	VAAGI	ΧN

**Fig. S1.** Sequence alignment of hantavirus nucleocapsid N¹⁻⁷⁴ coiled coil domain with the conserved hydrophobic (gray) and polar (yellow) heptads highlighted. The sequences are arranged according to hantaviral species that cause: **(top)** HCPS (Hantavirus CardioPulmonary Syndrome), **(middle)** non-pathogenic or mild form of HFRS (Hemorrhagic Fever with Renal Syndrome) and **(bottom)** severe form of HFRS. The hantavirus species are: AND, Andes; SNV, Sin Nombre; LAN, Laguna Negra; MUL, Muleshoe; BAY, Bayou; RIO, Rio Mamore; ELM, El Moro Canyon; NYV, New York; TUL, Tula; ISL, Isla Vista; PHV, Prospect Hill; PUU, Puumala; TOP, Topografov; KHA, Khabarovsk; SEO, Seoul; THA, Thailand; HTN, Hantaan; DOB, Dobrava.



**Fig. S2.** Secondary  $C^{\alpha}$ ,  $H^{\alpha}$ , C', and  $C^{\beta}$  chemical shifts show that the first 33 residues (shaded), which are the His-tag from pET151, lack secondary structure whereas the Andes virus N¹⁻⁷⁴ region consists of two  $\alpha$ -helices.

Table S1. Structural statistics for 20 NMR structures of Andes virus	5
N ¹⁻⁷⁴ coiled coil domain.	

1432	
92	
361	
700	
279	
38	
135	
73	
62	
0.21	
0.53	
0.27	
3.2	
-3861	
117	
91.4%	
8.2%	
0.4%	
0.1%	
	$     \begin{array}{r}       1432 \\       92 \\       361 \\       700 \\       279 \\       38 \\       135 \\       73 \\       62 \\       0.21 \\       0.53 \\       0.27 \\       3.2 \\       -3861 \\       117 \\       91.4\% \\       8.2\% \\       0.4\% \\       0.1\% \\     \end{array} $

## **Supplemental Data**



**Fig. S3**. Immunocytochemistry of full length N protein with mutations in the  $N^{1.74}$  coiled coil domain. Cos-7 cells were transfected with will type and mutant N protein, and doubly stained with (A) monoclonal and polyclonal anti-N antibodies, and (B) monoclonal anti-N and Golgi-specific antibodies. The anti-N monoclonal antibody was AB34757 from Abcam (Cambridge, Mass.).