Supporting Information

Strategic development of a next-generation multiepitope vaccine to prevent Nipah virus zoonotic infection

Rupal Ojha, Aditi Pareek, Rajan K. Pandey, Dhaneswar Prusty, Vijay K. Prajapati*

Rupal Ojha - Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817, Ajmer, Rajasthan, India.

Aditi Pareek- Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817, Ajmer, Rajasthan, India.

Rajan K. Pandey- Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817, Ajmer, Rajasthan, India.

Dhaneswar Prusty- Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817, Ajmer, Rajasthan, India.

Vijay K. Prajapati*- Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817, Ajmer, Rajasthan, India.

Running Title- Multi-epitope immunization for Nipah virus

***Corresponding author address**

Dr. Vijay Kumar Prajapati
Department of Biochemistry,
School of Life Sciences,
Central University of Rajasthan,
NH-8, Bandarsindri, Kishangarh,
305817, Ajmer, Rajasthan, India
Email- vkprajapati@curaj.ac.in

1. ADJUVANT-N LINKER-BCE-CTL-HTL

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKFAGSSSEVIVKKLEFEDEF
AGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPEIDENGSMIKKVGILHYEKLSKKIGP
KVSLIDTAAYVLMGVINSIAAYKTLIRTHIKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAY
MPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYMMASILLTLAAYTLFRRTKKKA
AYAPVENLNKLAAYFMVEILIEVAAYTIKSLMLLYAAYRASAATATLAAYSLMDINPWLAAYRIFFLSI
TKAAYNPWLNRLTWAAYMLSMIILYVAAYKLSKIGLVKAAYTPIKGALEIAAYKLISYTLPVAAYRLSI
GSPSKAAYKPENCRLSMGPGPGLLGSIVIIVMNIMIIGPGPGLEFEDEFAGSSSEVIGPGPGLEFEDEFAGS
SSEVIGPGPGEFEDEFAGSSSEVIVGPGPGLLTLFRRTKKKYRRHGPGPGHAGGIDQNMANRLGLGPGP
GHIKINGVISKRLFAQGPGPGILNKRYYSNLLILIL

2. ADJUVANT-N LINKER-BCE-HTL-CTL

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKFAGSSSEVIVKKLEFEDEF
AGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPEIDENGSMIKKVGILHYEKLSKKIGP
KVSLIDTGPGPGLLGSIVIIVMNIMIIGPGPGLEFEDEFAGSSSEVIGPGPGLEFEDEFAGSSSEVIGPGPGEF
EDEFAGSSSEVIVGPGPGLLTLFRRTKKKYRRHGPGPGHAGGIDQNMANRLGLGPGPGHIKINGVISKR
LFAQGPGPGILNKRYYSNLLILILAAYVLMGVINSIAAYKTLIRTHIKAAYMPKSRGIPIAAYKVLSYAPE
IAAYKSRGIPIKKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYMMASILL
TLAAYTLFRRTKKKAAYAPVENLNKLAAYFMVEILIEVAAYTIKSLMLLYAAYRASAATATLAAYSL
MDINPWLAAYRIFFLSITKAAYNPWLNRLTWAAYMLSMIILYVAAYKLSKIGLVKAAYTPIKGALEIAA
YKLISYTLPVAAYRLSIGSPSKAAYKPENCRLSM

3. ADJUVANT-N LINKER-CTL-BCE-HTL

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKVLMGVINSIAAYKTLIRTH
IKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPI
KKAAYMPKSRGIPIAAYMMASILLTLAAYTLFRRTKKKAAYAPVENLNKLAAYFMVEILIEVAAYTIKS
LMLLYAAYRASAATATLAAYSLMDINPWLAAYRIFFLSITKAAYNPWLNRLTWAAYMLSMIILYVAA
YKLSKIGLVKAAYTPIKGALEIAAYKLISYTLPVAAYRLSIGSPSKAAYKPENCRLSMKKFAGSSSEVIV
KKLEFEDEFAGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPEIDENGSMIKKVGILHY
EKLSKKIGPKVSLIDTGPGPGLLGSIVIIVMNIMIIGPGPGLEFEDEFAGSSSEVIGPGPGLEFEDEFAGSSS
EVIGPGPGEFEDEFAGSSSEVIVGPGPGLLTLFRRTKKKYRRHGPGPGHAGGIDQNMANRLGLGPGPGH
IKINGVISKRLFAQGPGPGILNKRYYSNLLILIL

4. ADJUVANT-N LINKER-CTL-HTL-BCE

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKVLMGVINSIAAYKTLIRTH
IKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPI
KKAAYMPKSRGIPIAAYMMASILLTLAAYTLFRRTKKKAAYAPVENLNKLAAYFMVEILIEVAAYTIKS
LMLLYAAYRASAATATLAAYSLMDINPWLAAYRIFFLSITKAAYNPWLNRLTWAAYMLSMIILYVAA
YKLSKIGLVKAAYTPIKGALEIAAYKLISYTLPVAAYRLSIGSPSKAAYKPENCRLSMKKFAGSSSEVIV
KKLEFEDEFAGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPEIDENGSMIKKVGILHY
EKLSKKIGPKVSLIDTGPGPGLLGSIVIIVMNIMIIGPGPGLEFEDEFAGSSSEVIGPGPGLEFEDEFAGSSS
EVIGPGPGEFEDEFAGSSSEVIVGPGPGLLTLFRRTKKKYRRHGPGPGHAGGIDQNMANRLGLGPGPGH
IKINGVISKRLFAQGPGPGILNKRYYSNLLILIL

5. ADJUVANT-N LINKER-HTL-BCE-CTL

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKLLGSIVIIVMNIMIIGPGPG
LEFEDEFAGSSSEVIGPGPGLEFEDEFAGSSSEVIGPGPGEFEDEFAGSSSEVIVGPGPGLLTLFRRTKKKY
RRHGPGPGHAGGIDQNMANRLGLGPGPGHIKINGVISKRLFAQGPGPGILNKRYYSNLLILILKKFAGSS
SEVIVKKLEFEDEFAGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPEIDENGSMIKKV

GILHYEKLSKKIGPKVSLIDTAAYVLMGVINSIAAYKTLIRTHIKAAYMPKSRGIPIAAYKVLSYAPEIAA
YKSRGIPIKKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYMMASILLTLA
AYTLFRRTKKKAAYAPVENLNKLAAYFMVEILIEVAAYTIKSLMLLYAAYRASAATATLAAYSLMDIN
PWLAAYRIFFLSITKAAYNPWLNRLTWAAYMLSMIILYVAAYKLSKIGLVKAAYTPIKGALEIAAYKLI
SYTLPVAAYRLSIGSPSKAAYKPENCRLSM

6. ADJUVANT-N LINKER-HTL-CTL-BCE

GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKLLGSIVIIVMNIMIIGPGPG
LEFEDEFAGSSSEVIGPGPGLEFEDEFAGSSSEVIGPGPGEFEDEFAGSSSEVIVGPGPGLLTLFRRTKKKY
RRHGPGPGHAGGIDQNMANRLGLGPGPGHIKINGVISKRLFAQGPGPGILNKRYYSNLLILILAAYVLM
GVINSIAAYKTLIRTHIKAAYMPKSRGIPIAAYKVLSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYKV
LSYAPEIAAYKSRGIPIKKAAYMPKSRGIPIAAYMMASILLTLAAYTLFRRTKKKAAYAPVENLNKLAA
YFMVEILIEVAAYTIKSLMLLYAAYRASAATATLAAYSLMDINPWLAAYRIFFLSITKAAYNPWLNRLT
WAAYMLSMIILYVAAYKLSKIGLVKAAYTPIKGALEIAAYKLISYTLPVAAYRLSIGSPSKAAYKPENC
RLSMKKFAGSSSEVIVKKLEFEDEFAGSKKFEDEFAGSSSKKMEVLKEEAWRKKFAPGGYPLLWKKPE
IDENGSMIKKVGILHYEKLSKKIGPKVSLIDT

Scheme S1 (1-6)- Designed multi-epitope subunit vaccine constructs. The vaccine construct is consisting of 45 amino acid residues of adjuvant sequence (β-defensin), at N terminal end EAAAK linker was added which link the epitopes and adjuvant together. The B-cell epitope were united with KK linker whereas, GPGPG and AAY linkers were used for combining the, HTL and CTL epitopes, respectively.

Table S1- Predicted B-cell epitopes for the selected Nipah Virus proteins with their respective peptide sequences and scores

S.No.	Protein Name	Peptide sequence	Score
1	Phosphoprotein, Protein P	FAGSSSEVIV	0.8
2	Protein W	LEFEDEFAGS	0.79
3	Non-Structural Protein V	FEDEFAGSSS	0.82
4	Protein C	MEVLKEEAWR	0.82
5	Nucleoprotein N	FAPGGYPLLW	0.78
6	Matrix Protein M	PEIDENGSMI	0.9
7	Fusion Glycoprotein F	VGILHYEKLS	0.82
8	Glycoprotein G	IGPKVSLIDT	0.82

Table S2-Predicted HTL epitopes for the selected Nipah Virus proteins with their respective peptide sequences, percentile rank, IC50 value, allele and geographical distribution

S. No	Protein Name	Peptide Sequence	Percentile Rank	IC50	Allele	Geographical Distribution
1	Phosphoprotein, Protein P	LEFEDEFAGSSSEVI	0.45	41	HLA-DRB1*07:01	Australia, China, India, Indonesia, Philippines, Taiwan, Vietnam
2	Protein W	EFEDEFAGSSSEVIV	0.56	39	HLA-DRB1*07:01	Australia, China, India, Indonesia, Philippines, Taiwan, Vietnam
3	Non-Structural Protein V	LEFEDEFAGSSSEVI	0.45	41	HLA-DRB1*07:01	Australia, China, India, Indonesia, Philippines, Taiwan, Vietnam
4	Protein C	LLTLFRRTKKKYRRH	0.17	6	HLA-DRB5*01:01	China, India, Indonesia, Philippines, Taiwan, Thailand
5	Nucleoprotein N	HAGGIDQNMANRLGL	0.15	17	HLA-DRB1*13:02	China, India, Indonesia, Philippines, Taiwan, Vietnam
6	Matrix Protein M	HIKINGVISKRLFAQ	0.66	27	HLA-DRB1*07:01	Australia, China, India, Indonesia, Philippines, Taiwan, Vietnam
7	Fusion Glycoprotein F	ILNKRYYSNLLILIL	0.31	41	HLA-DRB1*15:01	Australia, Bangladesh, China, India, Indonesia, Taiwan, Thailand, Vietnam
8	Glycoprotein G	LLGSIVIIVMNIMII	0.30	40	HLA-DRB1*15:01	Australia, Bangladesh, China, India, Indonesia, Taiwan, Thailand, Vietnam

Table S3 -Predicted CTL epitopes by considering 3 supertypes along with their respective scores

S. No	Protein Name	A2 Supertype	Score	A3 Supertype	Score	B7 Supertype	Score
1	Phosphoprotein, Protein P	VLMGVINSI	1.3173	KTLIRTHIK	1.4482	MPKSRGIPI	1.7638
2	Protein W	KVLSYAPEI	1.2246	KSRGIPIKK	1.4246	MPKSRGIPI	1.7537
3	Non-Structural Protein V	KVLSYAPEI	1.2246	KSRGIPIKK	1.4349	MPKSRGIPI	1.7667
4	Protein C	MMASILLTL	1.4762	TLFRRTKKK	1.6469	APVENLNKL	1.5393
5	Nucleoprotein N	FMVEILIEV	1.5545	TIKSLMLLY	1.5024	RASAATATL	1.5713
6	Matrix Protein M	SLMDINPWL	1.5699	RIFFLSITK	1.7228	NPWLNRLTW	1.3532
7	Fusion Glycoprotein F	MLSMIILYV	1.4056	KLSKIGLVK	1.5671	TPIKGALEI	1.0884
8	Glycoprotein G	KLISYTLPV	1.5522	RLSIGSPSK	1.5554	KPENCRLSM	1.6985

Table S4- Physiochemical properties along with Antigenicity of designed vaccine constructs

Parameters	Construct 1	Construct 2	Construct 3	Construct 4	Construct 5	Construct 6
Number of Amino Acid	592	592	591	591	589	589
Molecular Weight	64464.05	64464.05	64415.06	64415.06	64355.01	64355.01
Isoelectric point	9.77	9.77	9.8	9.8	9.79	9.79
Instability index	38.6	38.46	38.36	38.36	38.46	38.61
Aliphatic Index	95.73	95.73	95.55	95.55	96.21	96.21
GRAVY	0.042	0.042	0.025	0.025	0.037	0.037
Antigenicity	0.5249	0.5301	0.5197	0.5168	0.5114	0.5124
Estimated Half Life **(in vitro)** **(in vivo)** **(E.coli)**	 30 h >20h >10h	 30 h >20h >10h	 30 h >20h >10h	 30 h >20h >10h	 30 h >20h >10h	 30 h >20h >10h