



## Supplementary Material

# Single Nucleotide Polymorphism Induces a Positive Selection Pressure at Gag-Pol Sites in Human Immunodeficiency Virus Favoring Drug Resistance Mutations

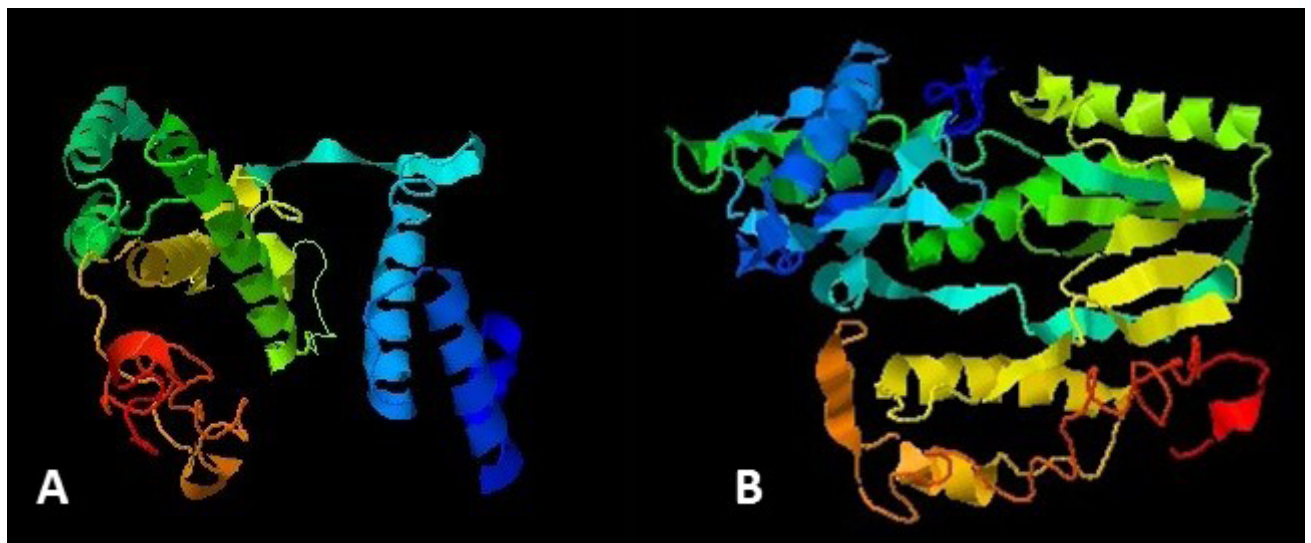
Amreen Zahra<sup>1</sup>, Mushtaq A. Saleem<sup>1\*</sup>, Hasnain Javed<sup>2</sup>, Muhammad Azmat Ullah Khan<sup>3</sup>, Muhammad Naveed<sup>1</sup> and Abdul Rauf Shakoori<sup>4</sup>

<sup>1</sup>Department of Biochemistry and Biotechnology, University of Central Punjab, Lahore 54782, Pakistan

<sup>2</sup>Advanced Diagnostic Laboratory, Punjab AIDS Control Programme, Government of Punjab, Lahore 54000, Pakistan

<sup>3</sup>Department of Biochemistry and Molecular biology, University of Gujrat, Gujrat 50700, Pakistan

<sup>4</sup>School of Biological Sciences, University of the Punjab, Quaid-I-Azam Campus, Lahore 54590, Pakistan



Supplementary Fig S1 A: I-TASSER protein predicted models, A:3D view of Gag protein, B: 3D view of Pol protein

\* Corresponding author: [dr.mushtaq@ucp.edu.pk](mailto:dr.mushtaq@ucp.edu.pk)  
0030-9923/2021/0005-1703 \$ 9.00/0  
Copyright 2021 Zoological Society of Pakistan



# University of Central Punjab

(Incorporated by Ordinance No. XXIV of 2002 promulgated by Government of the Punjab)

April 24, 2018

**The Human Research Ethics Committee**  
Faculty of Life Sciences  
University of Central Punjab  
Lahore

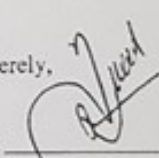
Subject: **REQUEST FOR THE APPROVAL OF BLOOD SAMPLE COLLECTION FOR PHD PROJECT ON MOLECULAR AND COMPUTATIONAL ANALYSIS OF HIV-1 IN PAKISTAN**

Respected Sir/Madam,

It is stated that Miss Amreen Zahra is a PhD Research scholar (Biochemistry) at Faculty of Life Sciences, University of Central Punjab under my Supervision. Her PhD research topic is **MOLECULAR AND COMPUTATIONAL ANALYSIS OF HIV-1 IN PAKISTAN**. She has successfully defended the PhD synopsis and thus will do sample collection of HIV-1 with research biotechnologist expert under complete biosafety conditions. Kindly review the project for approval in the above mentioned topic of research and your approval is highly appreciated and I express my deep gratitude and thanks for it.

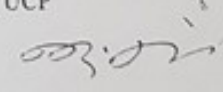
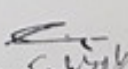
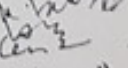
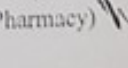

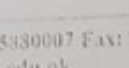
Thanking you very much in anticipation.

Your's sincerely,

Supervisor: 

**Dr. Muhammad Naveed**  
Assistant Professor  
FLS, UCP

**Approval by the Ethics Committee:**

- 1- Member 01 Dr. Muhammad Saqib Shahzad (UCP-FLS) 
- 2- Member 02 Dr. Sumayya Nizamuddin (SKMCHRC) 
- 3- Member 03 Dr. Muhammad Faheem (AIMC) 
- 4- Member 04 Mr. Nadeem Alvi (UCP-Pharmacy) 
- 5- Member 05 Mr. Muhammad Shafeeq ur Rahman (UCP-Pharmacy) 
- 6- Member 06 Dr. Muhammad Naveed (UCP-FLS) 

1-Khayaban-e-Jinnah, Johar Town, Lahore, Pakistan. Tel: 92-42-35330007 Fax: 92-42-35954392  
E-mail: info@ucp.edu.pk Website: www.ucp.edu.pk  
E-mail: admissions@ucp.edu.pk

**A**

<b>PUNJAB AIDS CONTROL PROGRAM (PACP) PACP COMPLEX ADVANCED DIAGNOSTIC LAB</b>	Document #	PACP_ADL-R&D-SOP-LDOC
	Version	PACP/V02
	Issue per	10-09-2018
	Effective per	15-09-2018
	Pages	12/22
	Author	AH/QAO
	Reviewed By	IRC
	Approved By	PD PACP
<b>PACP-IRC Legal Document For Conducting R&amp;D</b>		

**PACP INSTITUTIONAL REVIEW APPLICATION FORM**

University/ Institution Name UNIVERSITY OF CENTRAL PUNJAB, LAHORE

Ph No 042-35880007 Email: info@ucp.edu.pk

**1. SECTION A**

Project Title:				<u>MOLECULAR AND COMPUTATIONAL ANALYSIS OF HIV-1 IN PAKISTAN - PHD SCHOLAR - AMREEN ZAHRA</u>			
Researchers		Names/Designation		Institutions		Signature	
Principal Investigator/Supervisor (from Univ)		<u>DR. MUHAMMAD NAVEED, ASSISTANT PROFESSOR</u>		<u>FACULTY OF LIFE SCIENCES, UNIVERSITY OF CENTRAL PUNJAB LAHORE</u>			
Co-Investigator/Supervisor, if any (from Univ)		<u>DR. IRFAN AHMAD ASSISTANT PROFESSOR</u>		<u>ALLIED HEALTH SCIENCES, UNIVERSITY OF HEALTH SCIENCES</u>			
Co-Investigator/Supervisor (from PACP)		<u>DR. HASNAIN JAVED MOLECULAR BIOLOGIST</u>		<u>PUNJAB AIDS CONTROL PROGRAM (PACP) PRIMARY AND SECONDARY HEALTH CARE</u>			
Dates		Submission:		Review:		Protocol No.	
Project involves the use of: (✓ Check all pertinent ones)							
Therapeutic drug(s)		Epidemiological Research		Diagnostic research		Experimental Translational procedures	
		✓		✓			
Behavioral research		Community research		Radioactive agents		Other (please specify):	
Full Review		Expedite Review		Exempt Review		Renewal or Modifications	

**B** **RESTRICTED** 2

Supplementary Fig. S2. Approval of blood sample collection.

**Supplementary Table S1. Conserved domains prediction by Pfam: The domains predicted by Pfam of Gag-Pol region, are Gag\_p17, Gag\_p24, zf-CCHC RVT-1, RVT\_connect, RNase-H, RVP, RVT-thumb, MLVINC and Integrase-Zn.**

Source	Domain	Start	End	Gathering threshold (bits)		Score (bits)		E-value	
				Sequence	Domain	Sequence	Domain	Sequence	Domain
Pfam	Gag_p17	2	132	20.80	10.00	218.50	217.40	7.6e-62	1.6e-61
coiled_coil	n/a	87	121	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	98	131	n/a	n/a	n/a	n/a	n/a	n/a
low_complexity	n/a	104	119	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	Gag_p24	149	363	29.00	29.00	277.70	276.10	1.6e-79	4.8e-79
disorder	n/a	175	176	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	183	192	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	197	200	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	202	203	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	211	249	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	361	362	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	374	378	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	zf-CCHC	390	407	20.80	16.70	37.90	25.10	2.9e-06	0.032
Pfam	zf-CCHC	411	428	20.80	16.70	37.90	22.20	2.9e-06	0.26
coiled_coil	n/a	424	444	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	447	481	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	RVT 1	650	821	29.60	29.60	119.60	118.40	3.4e-31	8.1e-31
Pfam	RVT thumb	828	891	20.70	20.70	92.80	90.80	1.7e-18	5.6e-18
low_complexity	n/a	883	891	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	RVT connect	905	1006	21.00	21.00	197.00	196.10	1.1e-55	2.2e-55
low_complexity	n/a	982	990	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	RNase H	1023	1144	25.80	25.80	100.10	98.20	3.1e-25	1.1e-24
Pfam	Integrase Zn	1155	1192	21.70	21.70	55.70	53.80	8.6e-12	3.3e-11
disorder	n/a	1294	1296	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1298	1310	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1340	1341	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	Rve	1443	1558	32.50	32.50	92.40	90.30	5.5e-23	2.4e-22
disorder	n/a	1549	1551	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1590	1591	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1607	1611	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1617	1618	n/a	n/a	n/a	n/a	n/a	n/a
low_complexity	n/a	1619	1631	n/a	n/a	n/a	n/a	n/a	n/a
Pfam	MLVIN C	1652	1730	29.20	29.20	116.20	115.00	1.3e-30	3.1e-30
disorder	n/a	1709	1719	n/a	n/a	n/a	n/a	n/a	n/a
disorder	n/a	1726	1731	n/a	n/a	n/a	n/a	n/a	n/a

**Supplementary Table S2. Scanprosite results for Gag-Pol with domain descriptions.**

Domain	Description	Confidence level	Range	Sequence	Score
2FE2S_FER_12Fe-2S ( <i>ferredoxin-type iron-sulfur binding region signature</i> )	PS00197	(-1)	24-32	CTGATCATC	
EGF_1 ( <i>EGF-like domain signature 1</i> )	PS00022	(-1)	317-328		
ANAPHYLATOXIN_1 ( <i>Anaphylatoxin domain signature</i> )	PS01177	(-1)	530-562	CCtaGgtgacat.GCTctgatg...TacCat-ttgc.CC	
ANAPHYLATOXIN_1 ( <i>Anaphylatoxin domain signature</i> )	PS01177	(-1)	825-855	CCagGgttaatt.GCGaacc....TtC-cagctc.CC	
ASP_PROT_RETROV ( <i>Aspartyl protease</i> )	PS50878	(-1)	16 - 85	KEALLDTGADDTVLEDINLP-GKWKPK--MIGGIGFIKVKQY-DHILIEICGKKAIGTVLV GPTPVNIIGRNM	32.439
RT_POL ( <i>Reverse transcriptase (RT) catalytic domain profile</i> )	PS50878	(-1)	139 - 329	EGKISRIGPENPYNTPIFAIKKKD-STkWRKLVDFRELNKRTQDFW-EVQLGIPHPAGLKKKKSVTVLD-VGDAYFSVPLYEDFRKYTAFT-IPstnnetpGVRVQYNVLPQGWKG-SPAIFQSSMTKILEPFRKQNpdivI-YQYMDDLYVGSDDLEIGqHRTK-IEELRGHLLKWGFTTPDKKHQK EPPFLWMGYEL	48.849
ASP_PROTEASE ( <i>Eukaryotic and viral aspartyl proteases active site</i> )	PS00141	(-1)	18 - 29	ALLDTGADDTVL	0

**Supplementary Table S3. Molecular docking was done to predict bound conformations and free energies of binding of small-molecule ligands to macromolecular targets. A: Shows the binding energies of Gag with Dolutegravir and the binding energies predicted were -5.7kcal/mol with distance from rmsd lower binding energy as 0.00 and best mode rmsd upper binding energies as 0.00. B: Shows the binding energies of pol with Dolutegravir and the binding energies predicted were -7.9kcal/mol with distance from rmsd lower binding energy as 0.00 and best mode rmsd upper binding energies as 0.00.**

Mode	affinity (kcal/mol)	distance from rmsd l.b.	best mode rmsd u.b.
1	-7.9	0.000	0.000
2	-7.8	3.061	7.458
3	-7.7	9.667	13.758
4	-7.6	4.018	7.656
5	-7.1	3.347	4.479
6	-7.1	9.281	12.743
7	-7.1	5.381	7.396
8	-7.0	4.021	6.435
9	-6.9	5.399	6.658

**Supplementary Table S3B. Binding affinity energies with Pol and Dolutegravir.**

Mode	affinity (kcal/mol)	distance from rmsd l.b.	best mode rmsd u.b.
1	-5.7	0.000	0.000
2	-5.7	14.813	18.458
3	-5.7	13.564	17.660
4	-5.7	19.777	24.082
5	-5.6	19.700	24.163
6	-5.6	20.369	24.682
7	-5.4	19.507	23.677
8	-5.4	13.887	17.966
9	-5.4	18.641	20.808