

# **Immunoinformatic and systems biology approaches to predict and validate peptide vaccine against Epstein–Barr virus (EBV)**

**Arif Ali\*<sup>1</sup>, Abbas Khan<sup>1</sup>, Aman Chandra Kaushik<sup>1</sup>, Yanjie Wang<sup>1</sup>, Syed Shujait Ali<sup>2</sup>,  
Muhammad Junaid<sup>1</sup>, Shoaib Saleem<sup>2</sup>, William CS Cho<sup>3</sup>, Xueying Mao<sup>4</sup>, Dong-Qing Wei\*<sup>1</sup>**

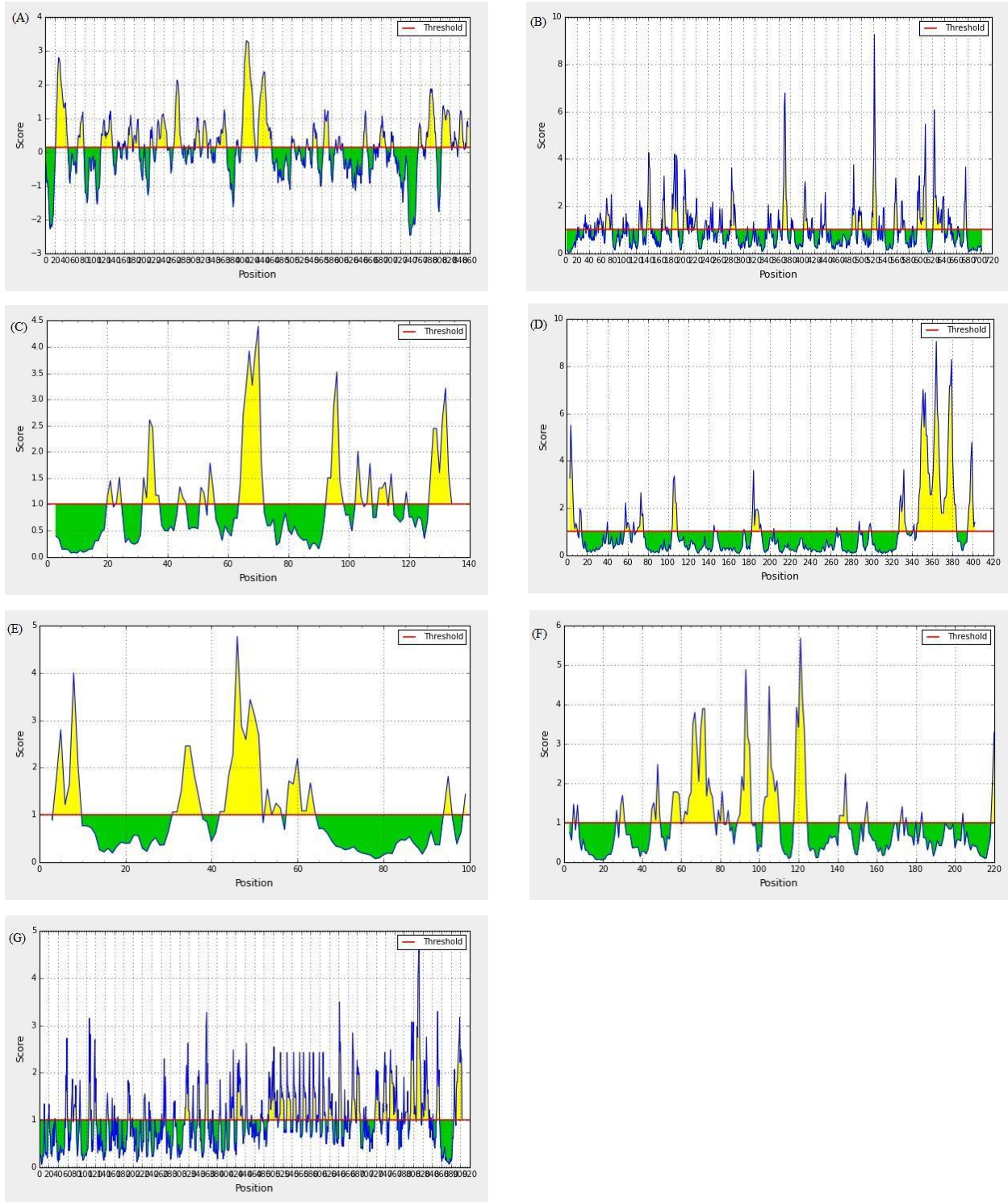
<sup>1</sup>State Key Laboratory of Microbial Metabolism, and College of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

<sup>2</sup>Center for Biotechnology and Microbiology, University of Swat, Khyber Pakhtunkhwa, Pakistan

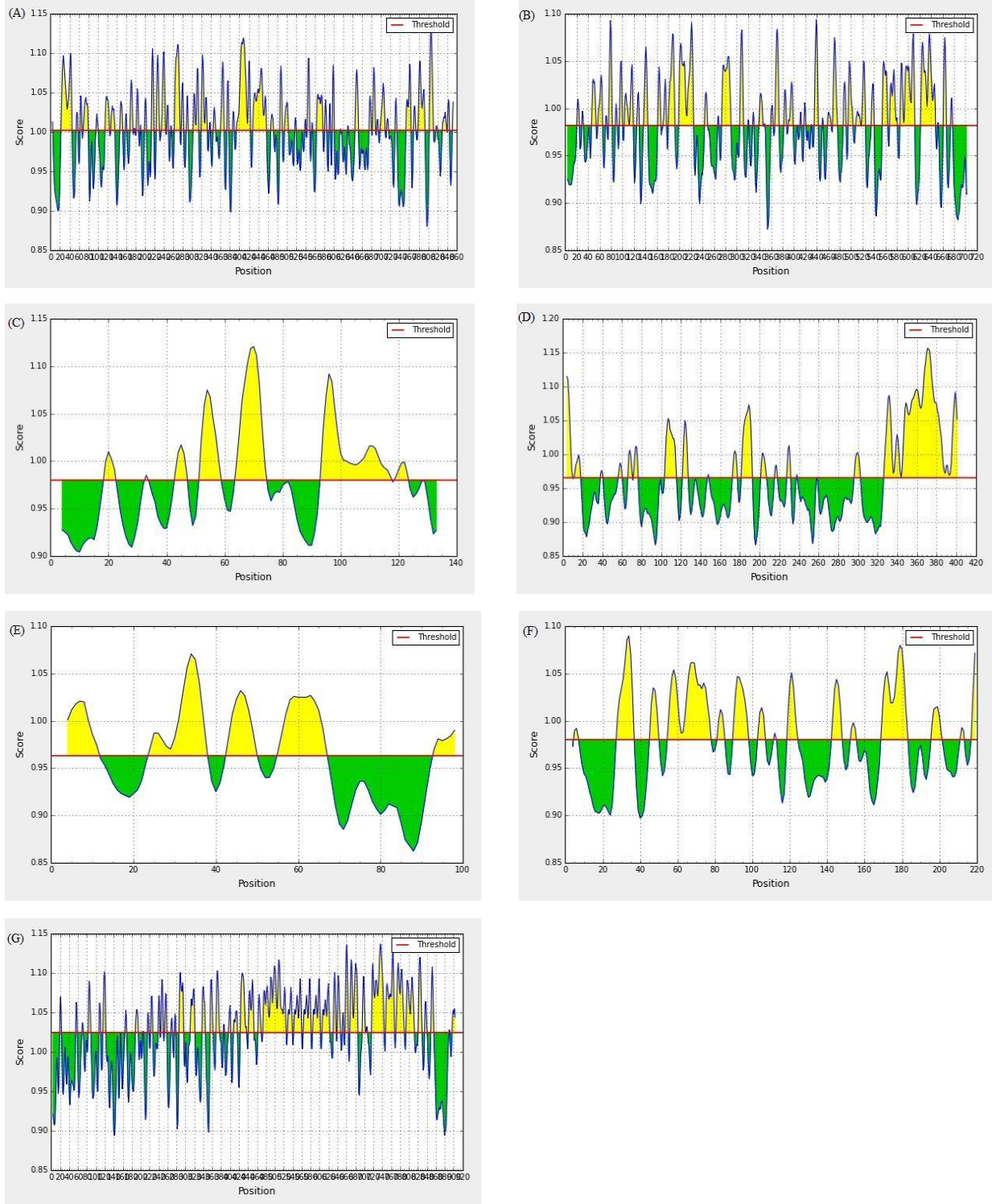
<sup>3</sup>Department of Clinical Oncology, Queen Elizabeth Hospital, Kowloon, Hong Kong

<sup>4</sup>Qianweichang College, Shanghai University, Shanghai China

Dong-Qing Wei\*<sup>1</sup>, (Corresponding Author) [dqwei@sjtu.edu.cn](mailto:dqwei@sjtu.edu.cn)

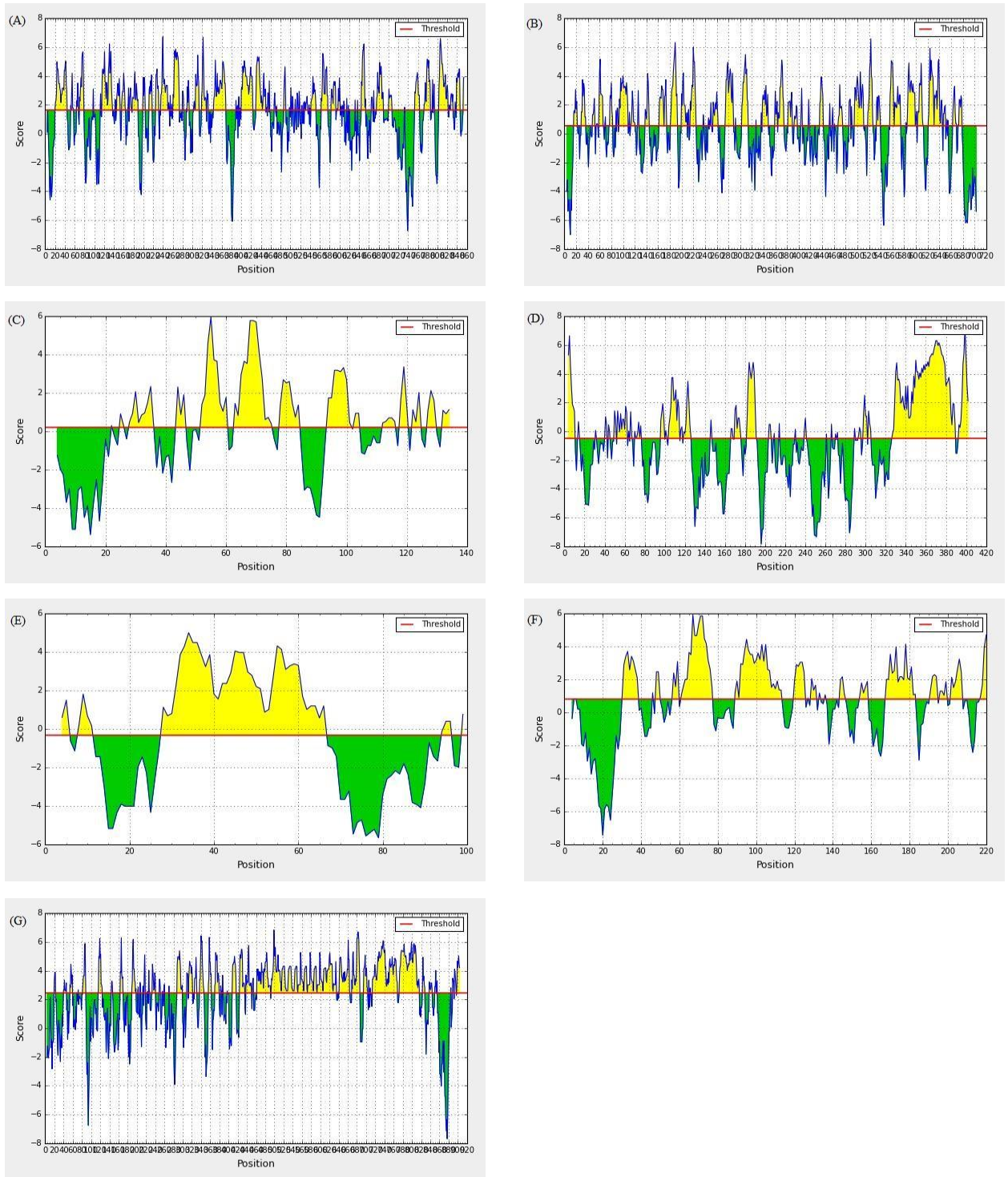


**Figure S1:** Graphical representation of predicted **Surface accessibility** Epstein-Barr Virus.



**Figure S2:** Graphical representation of predicted **Surface flexibility** of Epstein-Barr virus.





**Figure S3:** Graphical representation of predicted **Parker Hydrophilicity Prediction** for Epstein-Barr virus.

**Table S1:** The table is showing the predicted minimum and maximum Surface Accessibility score of selected proteins.

Species	Position	Residue	Start	End	Peptide	Score
<b>GB</b>	450	<b>R</b>	428	433	RRRRRD	9.415
	747	<b>I</b>	745	750	VVILVI	0.032
<b>GH</b>	522	<b>E</b>	520	525	EREDRD	9.265
	7	<b>F</b>	5	10	CVFCLV	0.033
<b>GL</b>	70	<b>G</b>	68	73	KNGSNQ	4.395
	10	<b>I</b>	8	13	LAICLV	0.071
<b>GM</b>	364	<b>R</b>	362	367	RNRRRS	9.054
	311	<b>P</b>	309	314	IIPILC	0.067
<b>GN</b>	46	<b>A</b>	44	49	TEAQDQ	4.777
	78	<b>L</b>	76	81	LVLVII	0.07
<b>GP42</b>	121	<b>K</b>	119	124	TKKKHT	5.691
	20	<b>V</b>	18	23	VIVLLL	0.054
<b>GP350</b>	812	<b>P</b>	810	815	PRPRYN	4.859
	5	<b>L</b>	3	8	AALLVC	0.058

**Table S2:** The table is showing the predicted minimum and maximum Schulz Flexibility Prediction score of selected proteins.

Species	Position	Residue	Start	End	Peptide	Score
<b>GB</b>	806	<b>Q</b>	803	809	EQNQEQK	1.13
	798	<b>M</b>	795	801	QAIMLAL	0.88
<b>GH</b>	439	<b>Q</b>	346	442	VITQGPN	1.094
	354	<b>M</b>	351	357	LAAMLMA	0.872
<b>GL</b>	70	<b>G</b>	67	73	PKNGSNQ	1.121
	10	<b>I</b>	7	13	FLAICLV	0.904
<b>GM</b>	371	<b>S</b>	368	374	STSSSSS	1.157
	196	<b>W</b>	193	199	FLWWVVF	0.866
<b>GN</b>	34	<b>S</b>	31	37	GASSPTN	1.071
	88	<b>M</b>	85	91	IYLMYVC	0.862
<b>GP42</b>	34	<b>G</b>	31	37	VRGGGRV	1.09
	40	<b>A</b>	37	43	VAAAAIT	0.897
<b>GP350</b>	736	<b>S</b>	733	739	PGNSSTS	1.137
	880	<b>M</b>	877	883	LLVMADC	0.894

**Table S3:** The table is showing the predicted minimum and maximum Surface Hydrophilicity score of selected proteins.

<b>Species</b>	<b>Position</b>	<b>Residue</b>	<b>Start</b>	<b>End</b>	<b>Peptide</b>	<b>Score</b>
<b>GB</b>	240	<b>N</b>	237	243	DGKNKET	6.729
	739	<b>L</b>	736	742	MLILVLV	-6.743
<b>GH</b>	523	<b>D</b>	520	526	EREDRDA	6.586
	10	<b>V</b>	7	13	FCLVLLW	-7.014
<b>GL</b>	55	<b>T</b>	52	58	SNQTCDG	5.971
	15	<b>I</b>	12	18	LVTIFVL	-5.40
<b>GM</b>	399	<b>S</b>	396	402	TIDSEEE	7.186
	196	<b>W</b>	193	199	FLWWVVF	-7.857
<b>GN</b>	34	<b>S</b>	31	37	GASSPTN	5.014
	79	<b>V</b>	76	82	LVLVIA	-5.671
<b>GP42</b>	67	<b>A</b>	64	70	NKTEAQE	5.943
	20	<b>V</b>	17	23	LVIVLLL	-7.457
<b>GP350</b>	499	<b>T</b>	496	502	DNGTESK	6.843
	877	<b>L</b>	874	880	LLLLLVM	-7.70