The Egyptian Journal of Immunology,
E-ISSN (2090-2506)
Volume 31 (3), July, 2024
Pages: 01–09.
www.Ejimmunology.org
https://doi.org/10.55133/eji.310310

Immunogenicity study of a novel DNA-based HCV vaccine candidate

Eman A. Salem[1], Ashraf Tabll[1,2], Tamer Z. Salem[3], Yasmine S. El-Abd[1], Reem Elshenawy[1], Heba Shawky[4], and Sahar Shoman[5]

[1]Department of Microbial Biotechnology, Biotechnology Research Institute, National Research Centre, Cairo, Egypt.
[2]Egypt Center for Research and Regenerative Medicine (ECRRM), Cairo, Egypt.
Molecular Biology & Virology Lab, Center for X-Ray Determination of the Structure of Matter (CXDS), Zewail City of Science & Technology, Giza, Egypt.
[4]Department of Therapeutic Chemistry, Pharmaceutical & Drug Research Institute, National Research Centre, Cairo, Egypt.
[5]Department of Microbiology, Faculty of Science, Ain Shams University, Cairo, Egypt.
Corresponding author: Yasmine S. El Abd, Department of Microbial Biotechnology, Biotechnology Research Institute, National Research Centre, Cairo, Egypt.
Email: yasminco@yahoo.com

Table of Contents

Supplementary Methods

Physicochemical properties, antigenicity, and allergenicity of the HCV-CE candidate

The complete translated amino acid sequence of the HCV-CE vaccine candidate was used to predict different physicochemical properties, including molecular weight, theoretical isoelectric point (pI), instability index, and grand average of hydropathicity, through the use of the ProtParam online tool.[1] The online server VaxiJen v2.0 (http://www.ddgpharmfac.net/ vaxijen/VaxiJen/VaxiJen.html), which predicts antigenicity depending on the target organism with an accuracy range of 70-89% based on the physicochemical properties of proteins, was used for antigenicity prediction.[2] To evaluate allergenicity, the vaccine construct was analyzed through the AllerTOP v2.0 server (https://www.ddgpharmfac.net/Aller TOP/ index.html), which utilizes an algorithm based on auto/cross-covariance transformation and the *k*-nearest neighbors methods to categorize peptide sequences as allergens or non-allergens with a reported accuracy of 85.3%.[3]

Tertiary structure modeling, refinement, and validation

The secondary structure of the vaccine construct was predicted through the software PSIPRED v4.0 (http://bioinf.cs.ucl.ac.uk/ psipred/), and then the 3D structure of the vaccine construct was generated by a fragment assembly based design approach of a *de novo* protein scaffold through the FoldDesign server (https://zhanggroup.org/ FoldDesign/index.html) using user-defined constraints, including secondary structure and/or contact and distance maps. The modeled structure was then subjected to a two-step refinement procedure through the GalaxyRefine tool (https://galaxy. seoklab.org/cgi-bin/submit.cgi?type=REFINE) followed by loop refinement using the GalaxyLoop server (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=LOOP). The refined model was then validated for structure quality using ProSA (https://prosa.services.came. sbg.ac.at/prosa.php) and the SAVES v6.0 toolkit, which includes a suite of tools (WHAT_CHECK, ERRAT, VERIFY_3D, and PROCHECK) for predicting different stereochemical parameters of the protein structure (https://saves.mbi.ucla.edu/?job= 1311167).

In-silico prediction of B-cell and T-cell epitopes

The translated amino acids of the full-length HCV-CE sequence were assessed for potential binding with T-cell and B-cell receptors required for the induction of specific immune responses. The B-cell epitopes were predicted using the BCPred online server with 75% specificity criteria for epitope prediction,[4] while conformational B-epitopes were predicted by the DiscoTope web server.[5] The peptide sequence was also subjected to helper T- lymphocyte and cytotoxic T-lymphocyte (HTL and CTL, respectively) epitope prediction using the MHC-II epitope prediction module of the Immune Epitope Database (IEDB) database and the NetCTL 1.2 server, respectively.[6,7] The generated HTL epitopes with IC50 values ≤ 50 nM were selected for further investigation. CTL epitopes with a combined score > 0.75 were selected and further submitted to the IEDB MHC-I binder predictor using default parameters. All epitopes were filtered according to their antigenicity and ability to induce interferon-γ (IFN-□) using the VaxiJen[2] and IFNepitope[8] servers, respectively. Both T-cell and B-cell epitopes were analyzed for their conservancy among different HCV genotypes and population coverage based on the MHC allele distribution using the IEDB epitope conservancy analysis (http://tools.iedb.org/conservancy/)[9] and population coverage tool (http://tools.iedb.org/population/),[10] respectively.

Molecular docking

To evaluate the potential immunoreactivity of the HCV-CE vaccine candidate to broadly neutralizing antibodies (nAbs), flexible protein–protein molecular docking was performed through the ClusPro 2.0 (https://cluspro.bu.edu/publications.php) server using the HCV-nAbs, AR3C [pdb:4MWF] and HEPC3 [pdb:6MEI], as receptors and the modeled tertiary structure of the HCV-CE vaccine candidate as a ligand in the "Antibody Mode" with masking of the non-complementarity-determining regions (CDR) regions in the receptor (nAb). The crystal structures of the heavy and light chains were retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (https://www.rcsb.org/), and the Molecular Graphics Laboratory (MGL) Tools software suite was used for the analysis of docked structures. The best complex (lowest energy score) was chosen for evaluation, and the powerful molecular graphics program PyMol (http://www.pymol.org) was used for visualization.

Immune simulation

C-ImmSim is an online server agent-based model that predicts peptide interactions with the immune system using position-specific scoring matrices (PSSMs) and machine learning techniques. Therefore, the server was used to simulate and characterize the immunogenicity of the designed vaccine and the anticipated immune responses. Two doses were introduced, with a two-week interval between them, and the simulation was run for 100 and 400 steps (step = 8 h).

Supplementary Tables

Supplementary Table 1. Linear B-cell epitopes predicted from primary HCV-CE peptide sequence.

Structure	Position	Epitope	Antigenic Score	INF-γ Inducer?	Conservation
Primary Peptide Sequence	424-431 (E2)	RTALNCND	**1.5102**	NO	*gp2, 4, 7
	645-670 (E2)	NWTRGEVCGLEHRDRVELSPLLLTTT	1.2433	YES	gp4
	100-120 (Core)	PRGSRPSWGPNDPRGRSRNLG	1.0572	YES	gp4
	42-88 (Core)	PRLGVRATRKTSERSQPRGRRQPIPKARRPEGRS WAQPGYPWPLYGN	0.5747	YES	gp4
3D-Structure	321-335 (E1)	DMMMNWSPTTTLVLA	**0.9456**	NO	gp4
	234-246 (E1)	NQSRCWVALTPTV	0.8678	NO	gp4
	677-746 (E2)	CSFTTLPALSTGLIHLHQNIVDVQYLYGVGSAVVS WALKWEYVVLAFLLLADARVSAYLWMMFMVS QVEA	0.7377	YES	gp4
	182-215 (Core)	LSCLTVPASAVNYRNVSGIYHVTNDCPNSSIVYE	0.6248	YES	gp4
	1-120 (Core)	MSTNPKPQRKTKRNTNRRPMDVKFPGGGQIVG GVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIP KARRPEGRSWAQPGYPWPLYGNEGCGWAGW LLSPRGSRPSWGPNDPRGRSRNLG	0.5962	YES	gp4

***gp: genotype**

Supplementary Table 2. Discontinuous B-cell epitopes predicted from modeled HCV-CE 3D-structure.

Residues	Epitope	Score	Antigenic Score	Conservation
665-746 (E2)	L665, C677, S678, F679, T680, T681, L682, P683, A684, L685, S686, T687, G688, L689, I690, H691, L692, H693, Q694, N695, I696, V697, D698, V699, Q700, Y701, L702, Y703, G704, V705, G706, S707, A708, V709, V710, S711, W712, A713, L714, K715, E717, Y718, V719, V720, L721, A722, F723, L724, L725, L726, A727, D728, A729, R730, V731, S732, A733, Y734, L735, W736, M737, M738, F739, M740, V741, S742, Q743, V744, E745, A746	0.888	0.6050	**gp4**
1-335 (Core-E1)	M1, S2, T3, N4, P5, K6, P7, Q8, R9, K10, T11, K12, R13, N14, T15, N16, R17, R18, P19, M20, V22, K23, F24, P25, G26, G27, G28, Q29, I30, V31, G32, G33, V34, Y35, L36, L37, P38, R39, R40, G41, P42, R43, L44, G45, V46, R47, A48, T49, R50, K51, T52, S53, E54, R55, S56, Q57, P58, R59, G60, R61, R62, Q63, P64, I65, P66, K67, A68, R69, R70, P71, E72, G73, R74, S75, W76, A77, Q78, P79, G80, Y81, P82, W83, P84, L85, Y86, G87, N88, E89, G90, C91, G92, W93, A94, G95, W96, L97, L98, S99, P100, R101, G102, S103, R104, P105, S106, W107, G108, P109, N110, D111, P112, R113, G114, R115, S116, R117, N118, L119, G120, I123, D124, L126, T127, C128, G129, F130, A131, D132, L133, M134, G135, Y136, I137, P138, L139, V140, G141, A142, P143, V144, G145, S146, V147, A148, L151, A152, R156, D160, I162, N163, Y164, A165, N168, L169, P170, G171, C172, L182, S183, L185, T186, V187, P188, A189, S190, A191, V192, N193, R195, N196, V197, S198, G199, I200, Y201, H202, V203, T204, N205, D206, C207, P208, N209, S210, S211, I212, V213, Y214, E215, N234, Q235, S236, C238, W239, V240, A241, L242, T243, P244, T245, V246, A247, A248, Y250, L264, M265, V266, G267, A268, T270, V271, G274, L275, Y276, I277, G278, D279, L280, C281, G282, G283, L284, F285, L286, V287, G288, Q289, M290, F291, S292, F293, R294, P295, R296, R297, H298, W299, T300, T301, Q302, D303, C304, N305, C306, S307, I308, Y309, T310, G311, H312, I313, T314, G315, H316, R317, M318, A319, W320, D321, M322, M323, M324, N325, W326, S327, P328, T329, T330, T331, L332, L334, A335	0.692	0.5679	**gp4**
512-639 (E2)	P512, R630, T631, F632, V633, G634, G635, I636, E637, H638, R639	0.662	0.6801	**gp4**
423-571 (E2)	N423, R424, R493, P494, C495, G496, I497, G518, T519, T520, D521, H522, V523, G524, V525, P526, T527, Y528, T529, G531, E532, N533, E534, T535, D536, V537, L539, N541, S542, T543, R544, P545, P546, H547, G548, A549, P569, C570, E571	0.634	0.6413	**gp4**

Supplementary Table 3. CTL epitopes predicted from primary HCV-CE peptide sequence.

Peptide*	Position	Serotype	Antigenicity Score	MHC-I Binding Allele**	INF-γ Inducer?	Conservation
RLGVRATRK	43-51	A3	2.3975	A*03:01,A*30:01,A*11:01	YES	gp1,2,3,4,6
PRLGVRATR	42-50	B27	2.3225	NONE	YES	gp4
FSFRPRRHW	291-299	B58	2.1015	B*57:01,B*58:01,A*32:01,B*53:01,B*44:02	YES	gp1,2,3,4,6,7
RRGPRLGVR	39-47	B27	1.8838	NONE	YES	gp1,2,3,4,6,7
QMFSFRPRR	289-297	A3	1.8009	A*31:01,A*33:01,A*03:01,A*68:01,A*30:01,A*11:01,A*32:01	YES	gp4
RVELSPLLL	659-667	B7	1.6869	A*32:01,A*01:01,B*07:02,A*30:01	YES	gp4
GPRLGVRAT	41-49	B7	1.5251	B*07:02	YES	gp1,2,3,4,6
DPRGRSRNL	111-119	B7,B8	1.5088	B*08:01,B*07:02,B*51:01	YES	gp4
LEHRDRVEL	654-662	B39,B44	1.4938	B*40:01,B*44:02,B*44:03	YES	gp4
HQNIVDVQY	693-701	A1,B62	1.4915	B*15:01, A*30:02,B*35:01,A*01:01,B*44:03,B*44:02,A*32:01,A*26:01,B*53:01,B*58:01,B*40:01	YES	gp1,3,4,6
CWVALTPTV	238-246	A24	1.4908	NONE	YES	gp1,3,4,6
GQMFSFRPR	288-296	B27	1.4655	A*31:01,A*11:01,A*30:01,A*03:01,A*33:01	YES	gp4
AYFSMQANW	360-368	A24,B58	1.3713	A*23:01,A*24:02,A*32:01,B*58:01,B*57:01,B*53:01,B*44:02,B*44:03,A*30:02	YES	gp4
NRRPMDVKF	16-24	B27	1.3678	NONE	YES	gp4,6
LPRRGPRLG	37-45	B7	1.2399	B*07:02	YES	gp1,2,3,4,6,7
RSRNLGKVI	115-223	B7	1.2197	A*30:01,B*57:01,B*07:02,A*32:01	YES	gp1,3,4,6
AVVSWALKW	708-716	B58	1.1546	A*32:01,B*58:01,B*57:01,B*53:01,A*26:01,B*44:02,A*23:01,B*44:03,A*24:02,A*30:02	YES	gp4
SAVVSWALK	707-715	A3	1.1417	A*11:01,A*30:01,A*68:01,A*03:01	YES	gp4
SPRGSRPSW	99-107	B7	0.8571	B*07:02,B*53:01,B*35:01,B*08:01,B*44:02,B*58:01,B*57:01,B*51:01,B*44:03,A*26:01,A*32:01	YES	gp1,2,3,4,6
LADARVSAY	726-734	A1,B62	0.8549	A*01:01,B*35:01,A*30:02,B*15:01,B*53:01,A*26:01,B*58:01	YES	gp4
KTSERSQPR	51-60	A3	0.8428	A*31:01,A*11:01,A*30:01,A*03:01,A*68:01,A*33:01	YES	gp1,2,3,4,6
YVVLAFLLL	718-726	A2,A26,B7,B39	0.8036	A*26:01,A*02:06,A*68:02	YES	gp4
ALSTGLIHL	684-692	A2	0.8008	A*02:03,A*02:01,A*02:06,A*32:01,B*15:01	YES	gp1,3,4,5,6
VVLAFLLLA	719-727	A2	0.7884	A*02:06	YES	gp4
FLLLADARV	723-731	A2	0.7163	A*02:01,A*02:06,A*02:03	YES	1,2,3,4
LLADARVSA	725-733	A2	0.7101	A*02:03,A*02:01,A*02:06	YES	gp4
RSTAGLANL	394-402	B58	0.7071	A*32:01,B*58:01,A*30:01,B*57:01	YES	gp4
QPRGRRQPI	57-65	B7,B8	0.6960	B*07:02,B*08:01,B*51:01	YES	gp1,2,3,4,
QPRGRRQPI	57-65	B7, B8	0.6960	B*07:02, B*08:01, B*51:01	YES	gp1, 2, 3, 4, 5, 6
SWHINRTAL	419-427	A24, B8, B39	0.6607	B*08:01, A*24:02, A*23:01, B*07:02	YES	gp2, 4, 6, 7
LTPTVAAPY	242-250	A1, A26, B62	0.6511	A*26:01, A*01:01, A*30:02, B*15:01, B*35:01	YES	gp4
GVGSAVVSW	704-712	B58	0.6490	A*32:01, B*58:01, B*57:01, B*53:01	YES	gp4
GSAVVSWAL	706-714	B39, B58	0.5648	A*32:01, B*58:01	YES	gp4
TQDCNCSIY	301-309	A1, B62	0.5321	A*01:01, A*30:02, B*15:01, B*35:01	YES	gp1, 4
GRSWAQPGY	73-81	B27	0.5047	A*30:02	YES	gp3, 4
ARALAHGVR	148-156	B27	0.4900	NONE	YES	gp1, 2, 3, 4
FSIFLLALL	174	B27	0.4729	A*68:02	YES	gp1, 2, 4
TVPASAVNY	186	B7	0.4727	A*26:01,A*30:02,B*35:01,B*15:01,A*01:01,A*68:01,A*11:01,B*53:01	YES	gp4
ALAHGVRAL	150-158	A2, B7, B39, B62	0.4296	A*02:03, A*02:01, A*02:06, A*32:01, B*15:01 , B*07:02, B*08:01	YES	gp3, 4
AVGRSTAGL	391	B7	0.4232	NONE	YES	gp4
STRPPHGAW	542-551	B58, B62	0.4133	B*57:01, A*32:01, B*58:01, A*26:01, A*30:01, B*53:01, A*30:02,B*07:02, B*15:01, B*44:02, B*35:01, A*01:01, B*44:03, A*23:01	YES	gp4
LVGQMFSFR	286-294	A3	0.4086	A*33:01, A*31:01, A*68:01, A*11:01	YES	gp4
NWAKVILVL	367-375	A24, B8, B39	0.4053	A*23:01,A*24:02, B*08:01	YES	gp4, 5

*Threshold >0.75

**Percentile rank <2. Binding alleles are listed in descending order.

Supplementary Table 4. HTL epitopes predicted from primary HCV-CE peptide sequence.

Peptide*	Position	Binding Allele	Antigenicity Score	INF-γ Inducer?	Conservation
ALKWEYVVLAFLLLA	**713-727**	**DPA1*02:01/DPB1*01:01**	**1.2602**	**YES**	**gp4**
LKWEYVVLAFLLLAD	714-728	DPA1*01:03/DPB1*02:01	1.1707	YES	**gp4**
WEYVVLAFLLLADAR	716-730	DPA1*01:03/DPB1*02:01	0.9845	YES	**gp4**
YVVLAFLLLADARVS	718-732	DRB1*01:01	0.8790	YES	**gp4**
LAFLLLADARVSAYL	721-735	DRB1*01:01	0.8137	YES	**gp4**
VVLAFLLLADARVSA	719-733	DRB1*01:01	0.7812	YES	**gp4**
EYVVLAFLLLADARV	717-731	DRB1*01:01	0.7289	YES	**gp4**
VLAFLLLADARVSAY	720-734	DRB1*01:01	0.6458	YES	**gp4**
THVSGAAVGRSTAGL	385-399	DQA1*05:01/DQB1*03:01	0.5288	YES	**gp4**
AFLLLADARVSAYLW	722-736	DRB1*01:01	0.4230	YES	**gp4**

***IC50 < 10 nM, percentile rank< 1.5**

Supplementary Table 5. Residues of HCV-CE with polar contacts with the nAb residues

nAb	Residues in 'ligand'-> 'receptor' residues in close contact	H-bonds (donor residue->acceptor residue(s))
AR3C	1**. HCV-CE**:VAL626->**AR3C**:PRO182, 2. **HCV-CE**:LYS64->**AR3C**:LYS64, 3. **HCV-CE**:CYS644->**AR3C**:HID179, 4. **HCV-CE**:ARG639->**AR3C**:GLN120, 5**. HCV-CE**:CYS607->**AR3C**:HID179, 6**. HCV-CE**:PHE627->**AR3C**:THR180, HID179, PHE181, 7. **HCV-CE**:ARG606->**AR3C**:HID179, 8. **HCV-CE**:SER625->**AR3C**:PHE181,	**1. HCV-CE**:ARG606->**AR3C**:HID179 **2. HCV-CE**:ARG639->**AR3C**:GLN120
HEPC3	1. **HCV-CE**:THR444->**HEPC3**:CYS108, 2. **HCV-CE**:LYS588->**HEPC3**:GLY202, THR203, 3. **HCV-CE**:SER450->**HEPC3**:GLY174, 4. **HCV-CE**:GLY451->**HEPC3**:THR172, LEU171, 5. **HCV-CE**:GLU591->**HEPC3**:SER199, 6. **HCV-CE**:CYS452->**HEPC3**:LEU171, 7. **HCV-CE**:TYR443->**HEPC3**:ARG107, 8. **HCV-CE**:ASN448->**HEPC3**:VAL175, 9. **HCV-CE**:ARG587->**HEPC3**:GLY169, THR203,	**1. HCV-CE**:GLY451->**HEPC3**:THR172 **2.HCV-CE**:ARG587->**HEPC3**:GLY169 **3.HEPC3**:SER199->**HCV-CE**:GLU591

Supplementary References

1. Dimitrov I, Bangov I, Flower DR, et al. (2014). AllerTOP v.2--a server for in silico prediction of allergens. *J Mol Model.* 20(6):2278. https://doi.org/10.1007/s00894-014-2278-5

2. El-Manzalawy Y, Dobbs D, Honavar V (2008). Predicting linear B-cell epitopes using string kernels. *J Mol Recognit.* 21(4):243-55. https://doi.org/10.1002/jmr.893

3. Kringelum JV, Lundegaard C, Lund O, et al. (2012). Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol.* 8(12):e1002829. https://doi.org/10.1371/journal.pcbi.1002829

4. Vita R, Mahajan S, Overton JA, et al. (2019). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 8; 47(D1):D339-D343. https://doi.org/10.1093/nar/gky1006

5. Larsen MV, Lundegaard C, Lamberth K, et al. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics.* 8:424. https://doi.org/10.1186/1471-2105-8-424

6. Shahid F, Ashfaq UA, Javaid A, et al. (2020). Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. *InfectGenet Evol*. 80:104199.https://doi.org/10.1016/j.meegid.2020.104199

7. Vita R, Zarebski L, Greenbaum JA, et al. (2010) The immune epitope database 2.0. *Nucleic Acids Res.*38(Databaseissue):D854-62.https://doi.org/10.1093/nar/gkp1004

8. Bui HH, Sidney J, Dinh K, et al. (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics.* 7:153. https://doi.org/10.1186/1471-2105-7-153

9. Morris GM, Huey R, Lindstrom W, et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 30(16):2785-91.https://doi.org/10.1002/jcc.21256

10. Rapin N, Lund O, Bernaschi M, et al. (2010). Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One.* 5(4):e9862. https://doi.org/10.1371/journal.pone.0009862