Open access • Posted Content • DOI:10.1101/2021.09.17.460735

Exploring the structural basis to develop efficient multi-epitope vaccines displaying interaction with HLA and TAP and TLR3 molecules to prevent NIPAH infection, a global threat to human health — Source link

Sukrit Srivastava, Ajay K. Saxena, Michael Kolbe

Institutions: Jawaharlal Nehru University, University of Hamburg

Published on: 20 Sep 2021 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Epitope, Helper T lymphocyte, CTL* and Linear epitope

Exploring the structural basis to develop efficient multi-epitope vaccines displaying interaction with HLA and TAP and TLR3 molecules to prevent NIPAH infection, a global threat to human health

Sukrit Srivastava*[a,b], Ajay Kumar Saxena[b], Michael Kolbe*[c,d],

a Infection Biology Group, Indian Foundation for Fundamental Research, RaeBareli 229001, India
b Molecular Medicine Lab., School of Life Science, Jawaharlal Nehru University, 110067 New Delhi, India.
c Department for Structural Infection Biology, Centre for Structural Systems Biology (CSSB) & Helmholtz-Centre for Infection Research, Notkestraße 85, 22607 Hamburg, Germany.
d Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Rothenbaumchaussee 19, 20148 Hamburg, Germany.

*Correspondence should be addressed to S.S. (srivastava.sukrit@iffr.in; srivastav.sukrit@gmail.com); and M.K. (Michael.Kolbe@helmholtz-hzi.de)

ABSTRACT

Nipah virus (NiV) is an emerging zoonotic virus responsible to cause several serious outbreaks in South Asian region with high mortality rate of 40 to 90% since 2001. NiV infection causes lethal encephalitis and respiratory disease with the symptom of endothelial cell-cell fusion. No specific vaccine has yet been reported against NiV infection. Recently, some Multi-Epitope Vaccines (MEV) has been proposed but they involve limited number of epitopes which further limits the potential of vaccine. To address the urgent need for a specific and effective vaccine against NiV infection, in the present study, we have design two multi-epitope vaccines (MEVs) composed of 33 Cytotoxic T lymphocyte (CTL) epitopes and 38 Helper T lymphocyte (HTL) epitopes. Both the MEVs carry

potential B cell linear epitope overlapping regions, B cell discontinuous epitopes as well as IFN-γ inducing epitopes. Hence the designed MEVs carry potential to elicit cell-mediated as well as humoral immune response. Selected CTL and HTL epitopes were validated for their stable molecular interactions with HLA class I and II alleles as well as in case of CTL epitopes, with human transporter associated with antigen processing (TAP). Human β-defensin 2 and β-defensin 3 were used as adjuvants to enhance the immune response of both the MEVs. The molecular dynamics simulation study of MEVs-TLR3(ECD) (Toll-Like Receptor 3 Ectodomain) complex indicated stable molecular interaction. Further, the codon optimized cDNA of both the MEVs has shown high expression potential in the mammalian host cell line (Human). Hence for further studies, both the design of CTL and HTL MEVs could be cloned, expressed and tried for *in-vivo* validations (animal trails) as potential vaccine candidates against NiV infection.

Key words: Nipah virus (NiV), Human Transporter associated with antigen processing (TAP), Toll-Like Receptor 3 (TLR-3), Epitope, Immunoinformatics, Molecular Docking, Molecular Dynamics (MD) simulation, Multi-epitope Vaccine (MEV)

INTRODUCTION

Nipah virus (NiV) is an emerging zoonotic virus of the genus Henipavirus of the Paramyxoviridae family [1]. NiV infection causes fatal encephalitis and respiratory disease with a particular symptom of endothelial cell-cell fusion [2].

The first NiV infection to human was first reported in Malaysia in 1998. Later NiV outbreak was reported from Meherpur, Bangladesh in 2001. In the Malaysia NiV infection, the transmission was primarily due being in contact of pigs, whereas in later outbreaks of Bangladesh and India the transmission was associated with contaminated date palm sap and human-to-human contact [3]. Bats are identified as the main reservoir for the NiV and they are responsible for the transmission of the infection to both humans and animals [4]. After 2001 NiV outbreak has been reported from different district of Bangladesh almost every year (2003-05, 2007-12). Till March 31, 2012, a total of 209 confirmed cases of NiV infections were reported out of which 161 people died resulting in the mortality rate as high as 77%. After several outbreaks in Bangladesh, in total three NiV outbreaks have also been reported from India. Two of them occurred in the state of West Bengal in 2001 and 2007 [5]. The most recent NiV outbreak was reported from the Kerala state of India during the period of May to June-2018. The Kerala outbreak claimed 17 lives leaving only two survivors out of 19 confirmed cases [6]. Till present, there has been no specific vaccine reported against NiV infection, and the pathogenesis mechanism of NiV to human cells is largely unknown. Hence, an immune-informatics approach investigating the potential of different NiV proteins for vaccine design would be an important and essential step forward for vaccine development.

NiV infection of human cells involves several protein-protein interactions and protein cluster formation on the host cell surface. Essential proteins involved in NiV pathogenesis include C protein, Fusion glycoproteins (F), Glycoproteins

77 (G), Matrix proteins (M), Nucleocapsid protein, Phosphoprotein, Polymerase, V
78 protein and the W protein [7-21]. The C protein regulates the early host pro-
79 inflammatory response as well as the pathogen virulence thus providing a
80 conducive environment for a successful NiV infection [7]. The attachment
81 glycoprotein (G), the fusion protein (F) and the matrix protein (M) together form a
82 cluster on the human cell membrane facilitating virus particle assembly and
83 pathogenesis [8-12]. The G and F proteins of NiV have been shown to be
84 immunogenic by inducing protective immune responses in hamsters [22, 23]. The
85 NiV matrix protein is observed to play a central role in virus particle formation and
86 is essentially required for viral budding from the infected human cells [13-15].
87 The NiV Polymerase is responsible for the initiation of RNA synthesis, primer
88 extension, and transition to elongation mode and hence the enzyme facilitates
89 viral pathogenesis and survival in host cells [16]. The phosphoprotein and the
90 glycoprotein of NiV are crucially involved in the regulation of viral replication [17,
91 18] while the V protein of NiV is responsible for the host interferon (IFN)
92 signalling evasion during pathogenesis [19, 20]. Interestingly, the identical N-
93 terminal region of the pathogen's V and W protein is sufficient to exert the IFN-
94 antagonist activity [21]. Hence, all the nine above mentioned NiV proteins are
95 crucial in different ways for viral pathogenesis and are important drug and
96 vaccine candidates.

97 The Nipah virus is a zoonotic RNA virus and it infects human respiratory
98 epithelium cells as well as differentiated neurons (in the brain and spinal cord).
99 Thus, as understood by previous animal model studies, recovery from viral

infection and the clearance of viral RNA requires the presence of virus-specific antibodies and interferon gamma (IFN-γ) secretion from T cells [24-26]. Along with the B cell, the T cell also play a critical role in immune response against NiV infection. In recent studies a number of B cell and T cell epitopes from the NIPAH proteome have been reported [27-38]. Most of these epitopes show strong interaction tendency with their respective HLA molecule binders. Further different approaches were proposed for the design of multi epitope vaccines [39, 40]. The proposed vaccines utilized a limited number (6 to 8) of T and B cell epitopes. The use of limited number of epitopes could be challenging for the successful presentation of the exogenous vaccine candidates in view of the proteolytic cleavage by Antigen Presenting Cells (APC).

In present study we have screened out the most potential Cytotoxic T lymphocyte (CTL) epitopes, Helper T lymphocyte (HTL) and B cell epitopes from the NiV proteome. We have shortlisted and priotized the most potential and highest scoring 33 CTL, 38 HTL and 16 B cell epitopes. We further studied several critical properties (like IC(50), Immunogenicity, Conservancy, Non-toxic etc) to identify the most potential epitopes against Nipah virus. The shortlisted epitopes were utilized for the design of CTL and HTL multi-epitope vaccines against Nipah virus. The designed vaccines were further studied for their stable interaction with immunological receptor the Toll-Like Receptor 3 (TLR3). The analysis of cDNA of the designed multi-epitope vaccine has predicted to be highly favorable for expression in mammalian cell line. Overall in the present study we have designed and proposed potential multi-epitope vaccines against

Nipah virus infection.

METHODOLOGY

In the present study, we have designed two multi-epitope vaccines (MEVs) composed of thoroughly screened most potential Cytotoxic T lymphocyte (CTL) epitopes and Helper T lymphocyte (HTL) epitopes derived from the nine NiV proteins (glycoprotein (gi-253559848); C protein (gi-1859635642); fusion protein (gi-13559813); matrix protein (gi-13559811); nucleoprotein (gi-1679387250); phosphoprotein (gi-1802790259); V protein (gi-1802790260); RNA polymerase (gi-15487370); W protein (gi-374256971)). CTL and HTL epitopes would be the most potential vaccine candidates since they are responsible for cell-mediated immune response by their presentation on the surface of antigen presenting cells (APCs) by their respective Class I and II human leukocyte antigen (HLA) allele binders. Both the CTL and HTL epitopes chosen for MEV design also carry overlapping regions of linear B cell epitopes. Moreover, both the MEVs also carry potential discontinuous B cell epitopes as well as IFN-γ inducing epitopes in their tertiary structure model. Hence both the designed MEVs carry potential to elicit cell-mediated as well as humoral immune response. Furthermore, both the MEVs were designed with human β Defensin 2 and human β Defensin 3 as adjuvant at their N and C termini to enhance the immunogenic response [41-42]. The β-defensins have considerable immunological adjuvant activity. The β-defensins 2 & 3 have been shown in previous studies to generate potent humoral immune responses when fused with B-cell lymphoma epitopes in

146 mouse model [43-46]. Since, the pro-inflammatory mediators enhance the
147 expression of β-defensins 2 & 3 in airway epithelial cells we chose β-defensins in
148 our study. The selected CTL and HTL epitopes were validated for their stable
149 molecular interactions with their respective HLA alleles binders; for CTL epitopes,
150 their molecular interaction with human Transporter associated with antigen
151 processing (TAP) was also analyzed [47, 48]. This analysis validated the CTL
152 epitopes that get transported through the TAP cavity for their presentation on cell
153 surface or not. The human TAP selectively pumps cytosolic peptides into the
154 lumen of the endoplasmic reticulum in an ATP-dependent manner [49]. The
155 tertiary structure models of both the MEVs were generated, refined, and further
156 docked with the human Toll-Like Receptor 3 (TLR3), which is an essential
157 immunoreceptor in this pathway [50, 51]. The nuclear localization of Nipah virus
158 W protein inhibits the signaling pathway of TLR3 upon pathogenesis to suppress
159 the TLR3 induced activation of the IFN response to eventually prevent relay of
160 the warning signals to uninfected cells. TLR3 is preferentially expressed by
161 human astrocytes of the central nervous system (CNS) upon infection. The NiV
162 infection involves invasion of the neurons of CNS and hence causes infection in
163 CNS. These studies indicate the importance of the TLR3 responses in immune
164 response and hence TLR3 has been chosen to be studied for its stable binding
165 with the designed multi-epitope vaccines [52-55]. The complexes of CTL and
166 HTL MEVs formed with the human TLR3 were further analysed for their stable
167 molecular interaction by a molecular dynamics simulation study. The cDNA of the
168 designed MEVs were generated and analysed for their high expression tendency

169 in the mammalian host cell line (Human). Overall, from the present *in-silico* study,
170 we may put forward the design of two MEVs, which qualify all the significant
171 criterions for being a potential vaccine candidate against NiV infection. The
172 corresponding workflow is shown in Supplementary figure S1.
173
174 *NiV proteins selected for potential epitope screening.* In the present study, nine
175 NiV proteins were used for epitope screening. They include C protein, Fusion
176 glycoproteins (F), Glycoproteins (G), Matrix proteins (M), Nucleocapsid protein,
177 Phosphoprotein, Polymerase, V protein and the W protein. The full-length protein
178 sequences of NiV proteins were retrieved from the NCBI database (National
179 Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/protein). A
180 total of 161 protein sequences available at NCBI, belonging to different strains
181 and origins of NiV, were retrieved. For structural based epitope screenings
182 available tertiary structures of NiV proteins were retrieved from Protein Data
183 Bank (PDB) (http://www.rcsb.org/pdb/home/home.do). For the NiV proteins with
184 no structure available, homology models were generated by Swiss-model,
185 (http://swissmodel.expasy.org/) [56] (Supplementary table S1).
186
187 **Screening of Potential Epitopes**
188 ***T cell Epitope Prediction***
189 *Screening of Cytotoxic T lymphocyte (CTL) Epitope.* The screening of Cytotoxic
190 T lymphocyte epitopes was performed by the IEDB (Immune Epitope Database)
191 tool "Proteasomal cleavage/TAP transport/MHC class I combined predictor"

192 (http://tools.iedb.org/processing/) [57-59]. Proteasome cleavage score depend on
193 the total amount of cleavage site in the protein. TAP score estimates an effective
194 log -(IC50) values (half maximal inhibitory concentration (IC50)) for binding to
195 TAP of a peptide or its N-terminal prolonged precursors. The MHC binding
196 prediction score is -log(IC50) values for binding to MHC of a peptide [60]. The
197 tool gives a "Total Score" which is a combined score of the proteasome, MHC,
198 TAP (N-terminal interaction), and processing analysis scores. The total score is
199 generated by using the combination of six different methods viz. Consensus, NN-
200 align, SMM-align, Combinatorial library, Sturniolo and NetMHCIIpan. The IC(50)
201 (nM) value for each epitope and MHC allele binding pairs were also obtained by
202 this IEDB tool. Epitopes having high, intermediate, and least affinity of binding to
203 their HLA allele binders have IC50 values < 50 nM, < 500 nM and < 5000 nM
204 respectively. Immunogenicity of all the screened CTL epitopes was also obtained
205 by using "MHC I Immunogenicity" tool of IEDB
206 (http://tools.iedb.org/immunogenicity/) with all the parameters set to default
207 analyzing 1st, 2nd, and C-terminus amino acids of the given screened epitope
208 [60]. The tool predicts the immunogenicity of a given peptide-MHC (pMHC)
209 complex on the basis of the physiochemical properties of constituting amino acid
210 and their position within the peptide sequence.
211
212 *Screening of Helper T lymphocyte (HTL) Epitopes.* To identify the Helper T
213 lymphocyte epitopes from NiV proteins, the IEDB tool "MHC-II Binding
214 Predictions" (http://tools.iedb.org/mhcii/) was used. Peptides with IC50 values

215 <50 nM are considered high affinity, <500 nM intermediate affinity and <5000 nM
216 low affinity [61-64]. The tool generates percentile rank for each potential peptide.
217 This percentile rank is generated by the combination of three different methods
218 viz. combinatorial library, SMM_align and Sturniolo and by comparing the score
219 of the peptide against the scores of other random five million 15-mer peptides of
220 SWISSPROT database [61-64]. The rank from the consensus of all three
221 methods was generated by the median percentile rank of the three methods.
222 Lower the value of percentile, higher would be the rank.
223
224 *Population Coverage by CTL and HTL epitopes.* The "Population Coverage" tool
225 of IEDB (http://tools.iedb.org/population/) was used to elucidate the world human
226 population coverage by the shortlisted 33 CTL and 38 HTL epitopes derived from
227 nine NiV proteins [65]. The T cells recognize complex between a specific major
228 MHC molecule and a particular pathogen-derived epitope. The given epitope will
229 elicit a response only in an individual that express an MHC molecule, which is
230 capable of binding that particular epitope. This denominated MHC restriction of T
231 cell responses and the MHC polymorphism provide basis for population coverage
232 study. The MHC types are expressed at dramatically different frequencies in
233 different ethnicities. Hence a vaccine with larger population coverage could be of
234 greater importance [65]. Clinical administration of multiple-epitopes involving both
235 the CTL and the HTL epitopes are predicted here to have a greater probability of
236 larger human population coverage worldwide.
237

238 ***B Cell Epitope Prediction***

239 *Sequence-based B Cell epitope prediction.* Protein sequence-based six different
240 methods were utilized to screen linear B cell epitopes from nine different NiV
241 proteins. These methods are available at "B Cell Epitope Prediction Tools" of
242 IEDB server (http://tools.iedb.org/bcell/). In this screening the parameters such
243 as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and
244 antigenic propensity of the polypeptides are correlated with their localization in
245 the protein. This allows the search for continuous epitopes prediction from
246 protein sequence. The prediction is based on the propensity scales for each of
247 the 20 amino acids. For a window size n, the i - (n-1)/2 neighboring residues on
248 each side of residue i are used to compute the score for the residue i [66-71].
249 These methods utilize the propensity scale method as well as the physiochemical
250 properties of the given antigenic sequence to screen potential epitopes using
251 "Bepipred Linear Epitope Prediction", "Chou & Fasman Beta-Turn Prediction",
252 "Emini Surface Accessibility Prediction", "Karplus & Schulz Flexibility Prediction",
253 "Kolaskar & Tongaonkar Antigenicity" and "Parker Hydrophilicity Prediction" tools
254 [66-71].

255

256 *Structure-based B cell epitope prediction.* The Ellipro (ElliPro: Antibody Epitope
257 Prediction tool; http://tools.iedb.org/ellipro/) and the DiscoTope2.0 (DiscoTope:
258 Structure-based Antibody Prediction tool; http://tools.iedb.org/discotope/)
259 methods available at IEDB, were used to screen the linear and the discontinuous
260 B cell epitopes [72, 73]. The ElliPro method analyses on the basis of the location

261 of residue in the protein's 3D structure. The residues lying outside of an ellipsoid
262 covering 90% of the inner core protein residues score highest Protrusion Index
263 (PI) of 0.9; and so on. The discontinuous epitopes predicted by the ElliPro tool
264 are clustered on the basis of the distance "R" in Å between two residue's centers
265 of mass lying outside of the largest possible ellipsoid. The larger value of R
266 indicates larger distant residues (residue discontinuity) are screened in the
267 epitopes. The Discotope 2.0 method is based on the "contact number" of the
268 residue's Cα carbon atom as well as on the propensity of a given residue to be a
269 part of an epitope [72, 73]. The residue "contact number" is the number of Cα
270 atoms in the antigen within a distance of 10 Å of the particular residue's Cα atom.
271 A low contact number would indicate the residue being close to the surface or in
272 protruding regions of the antigen's structures.
273
274 **Characterisation of potential epitopes**
275 *Epitope conservation analysis.* The shortlisted CTL, HTL and B cell epitopes
276 screened from nine NiV proteins were analysed for the conservancy of their
277 amino acid sequence by "Epitope Conservancy Analysis" tool
278 (http://tools.iedb.org/conservancy/) of IEDB. The epitope conservancy is the
279 fraction of protein sequences that contain that particular epitope. The analysis
280 was done against their entire respective source protein sequences of NiV
281 proteins retrieved from the NCBI protein database [74].
282
283 *Epitope Toxicity prediction.* The tool ToxinPred

284 (http://crdd.osdd.net/raghava/toxinpred/multi_submit.php) was used to analyse
285 the toxicity of shortlisted CTL, HTL and B cell epitopes. The tool allows to
286 identifying highly toxic or non-toxic short peptides. The toxicity check analysis
287 was done by the "SVM (Swiss-Prot) based" (support vector machine) method
288 utilizing dataset of 1805 sequences as positive, 3593 negative sequences from
289 Swissprot as well as an alternative dataset comprises the same 1805 positive
290 sequences and 12541 negative sequences from TrEMBLE [75].
291
292 *Overlapping residue analysis.* The overlapping residue analysis for the shortlisted
293 CTL, HTL and the B cell linear epitopes was performed by the Multiple Sequence
294 Alignment (MSA) analysis by Clustal Omega tool
295 (https://www.ebi.ac.uk/Tools/msa/clustalo/) of EBI (European Bioinformatics
296 Institute) [76]. The Clustal Omega multiple sequence alignment tool virtually
297 aligns any number of protein sequences and delivers an accurate alignment.
298
299 *Epitope selected for molecular interaction study with HLA allele and TAP*
300 *transporter.* On the basis of the overlapping residue analysis of shortlisted CTL,
301 HTL and linear B cell epitopes few numbers of CTL and HTL epitopes were
302 chosen for further analysis involving stable interaction with their respective HLA
303 allele binders and TAP cavity interaction (Supplementary table S3 & S4, Figure
304 3). These epitopes were chosen on the basis of them having partial or full
305 overlapping sequence region amongst all three types of epitopes (CTL, HTL and
306 B Cell), or having full sequence overlap amongst any of the two types of

307 epitopes, or having the highest number of the HLA allele binders.

308

309 **Molecular interaction analysis of selected epitopes with HLA allele and TAP**
310 **transporter.**

311 *Tertiary structure modelling of HLA alleles and selected T cell epitopes.* The
312 Swiss-model was used for homology modelling of the HLA class I and II allele
313 binders of shortlisted epitopes [56]. The amino acid sequence of the HLA allele
314 binders were retrieved from Immuno Polymorphism Database (IPD-IMGT/HLA)
315 (https://www.ebi.ac.uk/ipd/imgt/hla/allele.html). Templates for homology
316 modelling were chosen on the basis of highest amino acid sequence similarity.
317 All the HLA allele models were further validated by their QMEAN value. The
318 QMEAN value gives a composite quality estimate involving both global as well as
319 local analysis of the model [77]. Generated Model having acceptable QMEAN
320 value (cutoff -4.0) were chosen for further studies (Supplementary table S2).

321 The "Natural Peptides Module for Beginners" module of PEPstrMOD
322 (http://osddlinux.osdd.net/raghava/pepstrmod/nat_ss.php) was utilized to
323 generate tertiary structures for the selected CTL and HTL epitopes [78]. The time
324 window for simulation was set to 100 picoseconds (ps) in a vacuum environment.

325

326 *Molecular interaction analysis of selected CTL and HTL epitopes with HLA*
327 *alleles.* The AutoDock 4.2 (ADT) and the AutoDock Vina were used for *in-silico*
328 molecular docking study of the selected CTL and HTL epitopes with their
329 respective HLA class I and II allele binders [79, 80]. The generated docked

330 complexes were studied for their stable nature by molecular dynamics simulation.
331 MD simulation was performed by the Gromacs 5.1.4 using the Optimized
332 Potential for Liquid Simulations - all-atom force field (OPLS-AA) [81, 82].
333
334 *Molecular interaction analysis of selected CTL epitopes with TAP transporter.*
335 TAP transporter plays an important role in the presentation of CTL epitope. From
336 the cytosol after proteasome processing, the fragmented peptide of foreign
337 protein gets transported to endoplasmic reticulum (ER) through the TAP
338 transporter, from the ER these short peptides reach to the Golgi bodies and then
339 get presented on the cell surface [83]. Molecular interaction study of the
340 shortlisted CTL epitopes with the TAP transporter cavity was performed by
341 molecular docking using AutoDock Vina [79, 80]. As structural model the cryo-
342 EM structure of TAP transporter (PDB ID: 5u1d) after removing the antigen from
343 TAP cavity of the original structure [48] was used for epitope-TAP interaction
344 study.
345
346 **Design, characterisation and molecular interaction analysis of Multi-**
347 **Epitope Vaccines with immunological receptor.**
348 *Design of Multi-Epitope Vaccines.* All the screened CTL and HTL epitopes were
349 utilized to design CTL and HTL Multi-Epitope vaccines. Short peptides EAAAK
350 and GGGGS were used as rigid and flexible linkers respectively (Figure 1). The
351 GGGGS linker provides proper conformational flexibility to the vaccine tertiary
352 structure and hence facilitates stable conformation to the vaccine. The EAAAK

353 linker facilitates in domain formation and hence facilitates the vaccine to obtain
354 its final stable structure [84-86]. The human beta-defensin 2 (hBD-2) (PDB ID:
355 1FD3, Sequence: GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKP)
356 and the human beta-defensin 3 (hBD-3) (PDB ID: 1KJ6, Sequence:
357 GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK) were used
358 as adjuvants in the design of both the MEVs at N and C terminals respectively
359 [41-46, 87,]. Human Beta-defensins (hBD) have an important role in the
360 chemotactic activity memory T cells, immature dendritic cells and monocytes.
361 Beta-defensins are also involved in degranulation of the mast cells. Due to the
362 important role of hBDs in immune response enhancement, hBDs have been
363 chosen and utilized as adjuvants for the MEV designs.
364
365 **Figure 1. Design of Multi-Epitope Vaccine (MEVs). (A)** CTL and **(B)** HTL
366 epitopes were linked by the short peptide linker 'GGGGS'. Human β Defensin 2
367 and β Defensin 3 were used as an adjuvant at the N and C terminals
368 respectively. The short peptide EAAAK was used to link the β Defensin 2 and β
369 Defensin 3. Epitopes from different proteins were coloured in different colours. C
370 terminal 6xHis is designed as His tag. * Epitopes common to Phosphoprotein, V
371 Protein and W protein.
372
373 ***Characterisation of designed Multi-Epitope Vaccines***
374 *Interferon gamma inducing epitope prediction.* From the designed amino acid
375 sequence of both the MEVs potential interferon gamma (IFN-γ) epitopes were

screened by the "IFNepitope" server (http://crdd.osdd.net/raghava/ifnepitope/scan.php) using "Motif and SVM hybrid", (MERCI: Motif-EmeRging and with Classes-Identification, and SVM: support vector machine) method. The tool predicts peptides from protein sequences having the capacity to induce IFN-gamma release from CD4+T cells. This module generates overlapping IFN-gamma inducing peptides from query sequence. . For the screening, IEDB database with 3705 IFN-gamma inducing and 6728 non-inducing MHC class II binders is utilized [88, 89].

MEVs allergenicity and antigenicity prediction. The designed MEVs were further analysed for allergenicity and antigenicity prediction by utilizing the AlgPred (http://crdd.osdd.net/raghava/algpred/submission.html) and the Vaxigen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) tools respectively [90, 91]. The AlgPred prediction is based on the similarity of already known epitope with any region of the submitted protein. For the screening of allergenicity, the Swiss-prot dataset consisting of 101725 non-allergens and 323 allergens is utilized [90]. The VaxiJen utilizes an alignment-free approach, solely based on the physicochemical properties of the query amino acid sequence. For prediction of antigenicity, the Bacterial, viral and the tumour protein datasets were used to derive models for the prediction of whole protein antigenicity. Every set consisted of known 100 antigens and 100 non-antigens [91].

Physicochemical property analysis of designed MEVs. The ProtParam

399 (https://web.expasy.org/protparam/) tool was utilized to analyse the
400 physiochemical properties of the designed CTL and HTL MEVs [92]. The
401 ProtParam analysis performs an empirical investigation for the given query amino
402 acid sequence. ProtParam computes various physicochemical properties derived
403 from a given protein sequence.
404
405 *Tertiary structure modelling and refinement of MEVs.* The tertiary structure of the
406 designed CTL and HTL MEVs were calculated by homology modelling utilizing
407 the RaptorX structure prediction tool
408 (http://raptorx.uchicago.edu/StructurePrediction/predict/) [93]. RaptorX predicts
409 template-based secondary and tertiary structures, contacts, solvent accessibility,
410 disordered regions and binding sites for given protein sequence, with or with out
411 having close homologs in the Protein Data Bank (PDB). RaptorX also assigns
412 confidence scores to indicate the quality of a predicted 3D model [94-96]. Quality
413 assessment for both the generated homology models of CTL and HTL MEVs was
414 performed by their respective P-values. The P-value for a predicted homology
415 model is a probability score of the generated model being worse than the best.
416 Hence the P-value indicates a relative quality of the generated model in terms of
417 modelling error, combining the global distance test (GDT) and the un-normalized
418 global distance test (uGDT) indicating the error involved at each residue. The
419 smaller the P-value the greater the quality of a predicted model.
420 The refinement of both the generated MEV models was performed by
421 ModRefiner (https://zhanglab.ccmb.med.umich.edu/ModRefiner/) and

422 GalaxyRefine tool (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE)
423 [97, 98]. Modrefiner improves the physical realism and structural accuracy of the
424 model using a Two-step Atomic-level Energy Minimization. ModRefiner is an
425 algorithm for the atomic-level, high-resolution protein tertiary structure
426 refinement. Both the side-chain and the backbone atoms of protein structure are
427 completely flexible during the structure refinement simulations. The
428 conformational search is guided by a composite of physics and knowledge based
429 force field. The tool generates significant improvement in the physical quality of
430 the local structures [97]. TM-score generated by ModRefiner indicates the
431 structural similarity of the refined model with the original input model. Closer the
432 TM-Score to 1, higher would be the similarity of original and the refined model.
433 The GalaxyRefine tool refines the query tertiary structure by repeated structure
434 perturbation as well as by utilizing the subsequent structural relaxation by the
435 molecular dynamics simulation. The tool GalaxyRefine generates reliable core
436 structures from multiple templates and then re-builds unreliable loops or termini
437 by using an optimization-based refinement method [96, 99]. To avoid any breaks
438 in the 3D model GalaxyRefine uses triaxial loop closure method. The MolProbity
439 score generated for a given refined model indicates the log-weighted
440 combination of the clash score (the number of atomic clashes per 1000 atoms),
441 the Ramachandran favored backbone torsion angles and the percentage of bad
442 side-chain rotamers (the percentages of rotamer outliers). The 'GDT-HA' (Global
443 Distance Test-High Accuracy) generated by the tool indicates the backbone
444 structure accuracy; 'RMSD' (Root mean Square Deviation) indicates the overall

445 structural deviation in refined model from the initial model and the 'Rama favored'
446 indicates percentage of Ramachandran favored residues.
447
448 *Validation of CTL and HTL MEVs refined models.* Both the refined CTL and HTL
449 MEV 3D models were further validated by RAMPAGE analysis tool
450 (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) [100, 101]. The generated
451 Ramachandran plots for the MEV models show the sterically allowed and
452 disallowed residues along with their dihedral psi (ψ) and phi (φ) angles.
453
454 *Discontinuous B-cell epitope prediction from MEVs.* Both the generated tertiary
455 models of designed CTL and HTL MEVs were subjected to discontinuous B cell
456 epitopes prediction. The structure-based discontinuous B cell epitopes were
457 screened from both the MEV models by utilizing the ElliPro method as described
458 earlier [72].
459
460 ***Molecular interaction analysis of MEVs with TLR-3***
461 Molecular interaction analysis of both the designed MEVs with Toll-Like receptor-
462 3 (TLR-3), was performed by molecular docking and molecular dynamics
463 simulation. Molecular docking was performed by PatchDock server
464 (http://bioinfo3d.cs.tau.ac.il/PatchDock/) [102-104]. PatchDock utilizes algorithm
465 for unbound (real life) docking of molecules for protein-protein. The algorithm
466 carries out the rigid docking, with the surface variability/flexibility implicitly
467 addressed through liberal intermolecular penetration. The algorithm focuses on

468 the (i) initial molecular surface fitting on localized, curvature based surface
469 patches (ii) use of Geometric Hashing and Pose Clustering for initial
470 transformation detection (iii) computation of shape complementarity utilizing the
471 Distance Transform (iv) efficient steric clash detection and geometric fit scoring
472 based on a multi-resolution shape representation and (v) utilization of biological
473 information by focusing on hot spot rich surface patches [102-104]. For molecular
474 docking, the 3D structure of human TLR-3 ectodomain (ECD) was retrieved from
475 the PDB databank (PDB ID: 2A0Z). Further, the molecular dynamics simulation
476 study of the MEVs-TLR-3 complexes were performed by Gromacs 5.1.4, utilizing
477 the Optimized Potential for Liquid Simulations - all-atom force field (OPLS-AA)
478 [81, 82]. MD simulation was performed to understand the properties of the MEVs-
479 TLR3 complexes in terms of their structure and the microscopic interactions in
480 the complex. The study provides dynamical properties of the designed system
481 with MEVs-TLR3 complexes with a guess at the interactions between the
482 molecules, and also it gives 'exact' predictions of bulk properties. MD Simulations
483 act as a bridge between theory and experiment [81, 82].
484

485 ***In-silico* analysis of MEVs for cloning and expression**

486 *Analysis of cDNA of the MEVs for cloning and expression in mammalian cell line.*
487 Complementary DNA of both the MEVs, codon optimized for expression in
488 Mammalian cell line (Human) were generated by Java Codon Adaptation Tool
489 (http://www.jcat.de/). The generated cDNA were further analysed by GenScript
490 Rare Codon Analysis Tool (https://www.genscript.com/tools/rare-codon-analysis).

The tool analyses the GC content, Codon Adaptation Index (CAI) and the Tandem rare codon frequency for a given cDNA [105-107]. The CAI indicates the possibility of cDNA expression in a chosen expression system. The tandem rare codon frequency indicates the presence of low-frequency codons in the given cDNA.

RESULTS & DISCUSSION

Screening of potential epitopes

T cell Epitope Prediction

Screening of Cytotoxic T lymphocyte (CTL) Epitope. Cytotoxic T lymphocyte (CTL) epitopes screened were shortlisted according to the highest "Total Score", low IC(50) (nM) value for epitope-HLA class I allele complexes, and epitopes with the larger number of the HLA class I allele binders. The immunogenicity of the shortlisted CTL epitopes was also determined; the higher immunogenicity score indicates the greater immunogenic potential of the given epitope (Supplementary table S3, S7). A total of 33 CD8+ T cell epitopes were finally chosen. Out of the 33 CTL epitopes reported here 10 epitopes (Fusion Protein: FALSNGVLF; Glycoprotein: TVYHCSAVY; Nucleocapsid: YPALALNEF; Phosphoprotein: VSDAKMLSY; Polymerase: YPECNNILF, FPVMGNRIY, AEFFSFFRTF, IPFLFLSAY, ETDDYNGIY, SQNLLVTSY) show a match with previous studies [35-38], indicating consensus of epitope screening by different approaches and methods (Supplementary table S3).

514 *Screening of Helper T lymphocyte (HTL) epitopes.* The screening of helper T
515 lymphocyte (HTL) epitopes from nine different proteins of NiV was performed on
516 the basis of "Percentile rank". The smaller the value of percentile rank the higher
517 would be the affinity of the peptide with its respective HLA allele binders. In our
518 initial screening, we got several potential CD4+ T cell epitopes with high scoring.
519 38 epitopes out of initial screening were shortlisted on the basis that they had
520 highest percentile rank and highest number of HLA class II allele binders
521 (Supplementary table S4, S8).
522
523 *Population Coverage by CTL and HTL epitopes.* The population coverage by the
524 shortlisted epitopes was also studied, in particular involving countries of South
525 Asia, East Asia, Northeast Asia and the Middle East. From this study, we may
526 conclude that the combined use of all the shortlisted CTL and HTL epitopes
527 would have an average worldwide population coverage as high as 97.88%, with
528 a standard deviation of 21.97 (Supplementary table S5).
529
530 ***B Cell epitope prediction***
531 *Sequence-based B Cell epitope prediction.* In our initial study, we screened a
532 total of 116 B Cell epitope from nine different NiV proteins, with the epitope
533 length of at least four amino acids utilizing the Bepipred Linear Epitope Prediction
534 method. B cell epitopes predicted by another five different methods based on
535 different physiochemical properties were found to have significant consensus
536 with the epitope amino acid sequences predicted by Bepipred Linear Epitope

537 Prediction. Here, 16 out of the 116 epitopes were shortlisted having a length of 4
538 to 19 amino acids (Supplementary table S6, Figure 2). One of these 16 B cell
539 epitopes (Matrix Protein: SIPREFMIY) matches with a previous study [39],
540 indicating epitope screening consensus using different approaches and methods
541 (Supplementary table S6).
542
543 **Figure 2. Overlapping regions amongst the linear B cell epitopes predicted**
544 **by the BepiPred method and other seven different B cell epitopes**
545 **prediction methods.** B cell epitopes predicted by the BepiPred method and
546 other different protein sequence based (Chou.., Emini.., Karplus.., Kolaskar..,
547 and Parker..) and protein structure based (DiscoTope and ElliPro) prediction
548 methods were found to have significant consensus. Consensus overlapping
549 regions of BepiPred epitopes are underlined by the different colour,
550 corresponding to respective prediction method.
551
552 *Structure-based B cell epitope prediction.* Structure-based discontinuous and
553 linear epitopes predicted by the DiscoTope 2.0 and the Ellipro methods have
554 shown a significant consensus of overlapping amino acid sequence with the
555 linear epitopes predicted by Bepipred linear epitopes method (Supplementary
556 table S6, Figure 2). This result further confirms that the shortlisted B Cell
557 Bepipred Linear Epitopes have a high chance of being highly immunogenic
558 epitopes.
559

560 **Characterisation of potential epitopes**

561 *Epitope conservation analysis.* Sequence conservation analysis of the shortlisted
562 33 CTL, 38 HTL and 16 B cell epitopes showed that the '100% amino acid
563 sequence' conservancy of CTL, HTL and B cell epitopes amongst all retrieved
564 NiV protein sequences is mostly 100%, as shown in Supplementary table S3, S4,
565 S6. This result indicates the high conservancy nature of the amino acid sequence
566 of the shortlisted CTL, HTL and B cell epitopes.

567

568 *Epitope toxicity prediction.* Toxicity analysis of all the shortlisted CTL, HTL and B
569 Cell epitopes was also performed. The ToxinPred study indicated the non-toxic
570 nature of all the shortlisted epitopes (Supplementary table S3, S4, S6).

571

572 *Overlapping residue analysis.* Amino acid sequence overlap analysis amongst
573 the shortlisted CTL, HTL and B cell epitopes from nine NiV proteins was
574 performed by using the Multiple Sequence Alignment (MSA) analysis tool Clustal
575 Omega. Our analysis showed that several epitopes of CTL, HTL and B cell have
576 overlapping amino acid sequence. The CTL, HTL and B cell epitopes having two
577 or more amino acids overlap are shown in Figure 3.

578

579 **Figure 3. Overlapping CTL, HTL and B cell epitopes**. Multiple sequence
580 alignment performed by Clustal Omega at EBI to identify the consensus
581 overlapping regions of CTL (red), HTL (blue) and B cell epitopes (green)
582 amongst shortlisted epitopes. Epitopes with overlapping region amongst all the

583 three types of epitopes (CTL, HTL and B Cell epitopes), epitopes with full
584 sequence overlap and epitopes with the highest number of HLA allele binders
585 were chosen for further studies (encircled).
586
587 *Epitope selected for molecular interaction study with HLA allele and TAP*
588 *transporter.* Amongst all the shortlisted epitope peptides seven CTL and
589 seventeen HTL epitope peptides have partial or full overlapping sequences or
590 have the highest number of HLA allele binders were shortlisted for further studies
591 (Figure 3, Supplementary table S3 & S4).
592
593 **Molecular interaction analysis of selected epitopes with HLA allele and TAP**
594 **transporter.**
595 *Molecular interaction analysis of selected CTL and HTL epitopes with HLA*
596 *alleles.* The molecular docking study of the shortlisted CTL and HTL epitopes
597 with their respective HLA class I and II allele binders was performed using
598 AutoDock Vina. Docking studies revealed for all epitopes significant molecular
599 interaction with their HLA allele binders having low binding energies and multiple
600 hydrogen bonds formated (Figure 4A & 4B). The stability of the obtained docking
601 complexes was further tested by molecular dynamics (MD) simulation studies.
602 MD simulations were performed over a time interval of 0.5-1 ns at the invariable
603 temperature of ~ 300 K and at invariable pressure of ~ 1 bar. All the complexes
604 showed reasonably invariant root mean square deviation (RMSD) value
605 (between ~ 0.2 to 0.4 nm) indicating the stable nature of the tested epitope-HLA

606 allele complexes (Figure 5A & 5B). Moreover, the reasonably invariant Rg (radius
607 of gyration) of the complexes, throughout the MD simulation (Supplementary
608 figure S2), and the root mean square fluctuation (RMSF) for all the atoms of the
609 complexes (Supplementary figure S3) again indicate the stable nature of the
610 epitopes and HLA allele complexes. Furthermore, the B-factor analysis of all the
611 epitope-HLA allele complexes indicated most of the complex regions to be stable
612 (blue) with a very small region being acceptably fluctuating (yellow and orange)
613 (VIBGYOR colour presentation) Supplementary figure S4.
614
615 **Figure 4. (A) Molecular Docking analysis of CTL epitopes and HLA alleles.**
616 Molecular docking of selected CTL epitopes (cyan sticks) with their respective
617 HLA class I allele binders (gray sticks). The study shows the docked complexes
618 to have significantly negative binding energy along with hydrogen bonds (green
619 dots) formation in the complex interface. **(B) Molecular Docking analysis of**
620 **HTL epitopes and HLA alleles.** Molecular docking of selected HTL epitopes
621 (cyan sticks) with their respective HLA class II allele binders (gray sticks). The
622 study shows the docked complexes to have significantly negative binding energy
623 along with hydrogen bonds (green dots) formation in the complex interface.
624
625 **Figure 5. (A) Molecular Dynamics simulation analysis of CTL epitopes and**
626 **HLA allele complexes.** Molecular Dynamics simulation study reveals a stable
627 nature of the CTL-HLA allele complexes throughout 0.5-1 ns time window with
628 reasonably invariable RMSD. **(B) Molecular Dynamics simulation analysis of**

CTL epitopes and HLA allele complexes. Molecular Dynamics simulation study reveals a stable nature of the HTL-HLA allele complexes throughout 0.5-1 ns time window with reasonably invariable RMSD.

Molecular interaction analysis of selected CTL epitopes with TAP transporter. The molecular docking interaction analysis of the chosen CTL epitopes with the TAP transporter cavity showed a significantly strong molecular interaction with low binding energy and several hydrogen bonds formed at different sites of the TAP transporter cavity. Two sites of interaction were of particular interest, one located near the cytoplasmic end and the other in the vicinity of the ER lumen (Figure 6). Our study confirms the transportation feasibility of the chosen CTL epitopes from the cytoplasm into the ER lumen which is essential for the representation of peptides by the HLA allele molecules on the surface of antigen presenting cells.

Figure 6. Molecular docking analysis of CTL epitopes within the TAP transporter cavity. Molecular interaction of CTL epitopes (cyan sticks) within the TAP cavity (gray ribbon/sticks) is shown in detail. For every panel of epitope-TAP complex, **(A)** shows the binding of epitope at two different sites within TAP cavity, **(B)** and **(C)** show detailed molecular interaction between epitopes and TAP cavity; (a, b) show chain A and B of TAP transporter. H-bonds are shown in yellow dots. (*) Binding energy, shown in kcal/mol.

652 **Characterisation and molecular interaction analysis of designed Multi-**
653 **Epitope Vaccines with immunological receptor**

654 ***Characterisation of designed Multi-Epitope Vaccines***

655 *Interferon-gamma inducing epitope prediction.* Interferon-gamma (IFN-γ)
656 inducing epitopes are involved in both the adaptive and the innate immune
657 response. The IFN-γ inducing 15-mer peptide epitopes were screened from the
658 CTL and HTL MEVs by utilizing the IFNepitope server. A total of 33 CTL MEV
659 and 43 HTL MEV INF-γ inducing POSITIVE epitopes with a score of 1 or more
660 than 1 were shortlisted (Supplementary table S9, Figure 7D & 7I).

661

662 **Figure 7. Tertiary structure modelling of CTL and HTL Multi-Epitope**
663 **Vaccines. (A) & (F)**: Tertiary structural models of CTL and HTL MEVs showing
664 epitopes in different colours corresponds to as in Figure 1. **(B) & (G)**: Show the
665 different domains of CTL and HTL MEVs. **(C) & (H)**: The overlapping linear B cell
666 epitope region present in CTL and HTL MEVs, shown by spheres. **(D) & (I):** From
667 the CTL and HTL MEVs, the INF-γ inducing epitopes are shown in cyan,
668 discontinuous B Cell epitopes are shown in magenta and the region common
669 amongst INF-γ and discontinuous B Cell epitopes are shown in wheat colour. **(E)**
670 **& (J)**: RAMPAGE analysis of the refined CTL and HTL MEV models.

671

672 *MEVs allergenicity and antigenicity prediction.* Both the CTL and HTL MEVs
673 were analyzed to be NON-ALLERGEN by the AlgPred analysis scoring -
674 0.61243421 and -0.93493027 respectively while default threshold value being -

675 0.4. CTL and HTL MEVs were also analyzed by VaxiJen to be probable
676 ANTIGENS with the prediction score of 0.4447 and 0.4836 respectively, while the
677 default threshold value for viral proteins being 0.4. Hence, with the mentioned
678 analysis tools both the CTL and HTL MEVs are predicted to be non-allergic as
679 well as antigenic in nature.
680
681 *Physicochemical property analysis of designed MEVs.* ProtParam analysis were
682 performed for both the CTL and HTL MEVs to analyse their physiochemical
683 properties. The CTL MEV is composed of 576 amino acids, has a molecular
684 weight of 58.57 kDa and a theoretical pI of 8.19. The expected half-life of the
685 CTL MEV in *E.coli*, yeast and mammalian reticulocytes were predicted with 10 h,
686 20 min, and 30 h respectively; the aliphatic index of CTL MEV was found to be
687 58.42, and grand average of hydropathicity (GRAVY) of CTL MEV was found to
688 be -0.010, both indicating globular and hydrophilic nature of the CTL MEV. The
689 instability index score of the CTL MEV was 48.03 indicating its stable folding
690 under native conditions.
691 Further, the ProtParam analysis of the HTL MEV showed for the 857
692 amino acids, a molecular weight of 87.62 kDa and a theoretical pI of 5.99. The
693 expected half-life of HTL MEV in *E.coli*, yeast and mammalian reticulocytes was
694 predicted to be 10 h, 20 min and 30 h, respectively. The aliphatic index of HTL
695 MEV was calculated as 82.99, and the grand average of hydropathicity (GRAVY)
696 of the HTL MEV was found to be 0.188, both indicating that HTL MEV has a
697 globular and hydrophilic nature. The instability index of the HTL MEV was 45.66

698 indicating its stable nature.

699

700 *Tertiary structure modelling and refinement of MEVs.* 3D homology models were
701 generated for both the CTL and HTL MEVs by utilizing the RaptorX modelling
702 tool (Figure 7A, 7F). The model obtained for CTL MEV has 6% helix, 27% β-
703 sheet, 66% coil content and structural elements are 23% exposed, 39% medium
704 and 36% buried. The structural model has three domains ranging from amino
705 acid 1 to 46 (1st domain, template-1fd3:A), 47 to 520 (2nd domain, templates-
706 1yrzA, 1y7bA, 5jozA, 3zxjA, 5z5dA) and 521 to 576 (3rd domain , template-
707 1kj6:A) (Figure 7B). Similarly, the 3D model calculated with RaptorX for the HTL
708 MEV has 21% helix, 22% β-sheet, 56% coil content with 24% of the amino acids
709 exposed, 39% medium and 36% buried. The structural model has three domains
710 ranging from amino acid 1 to 46 (1st domain, template-1fd3:A), 47 to 801 (2nd
711 domain, templates-5m5zA, 3eqnA), 802 to 857 (3rd domain, template-1kj6:A)
712 (Figure 7G). The P-Value for the best template based CTL and HTL MEV
713 homology models were 2.79e-04 and 5.99e-03 respectively. Good quality, mostly
714 alpha proteins have a P-value of less than 10^{-3} and that of mostly beta proteins
715 has a P-value of less than 10^{-4}. Hence both the homology models of CTL and
716 HTL MEVs are predicted to be of good quality. Since for the CTL and HTL MEV
717 design, the CTL and HTL epitopes used also show overlapping common regions
718 with the linear B cell epitopes (Figure 2), both the generated CTL and HTL MEV
719 models also carry the overlapping regions of linear B Cell epitopes (Figure 7C,
720 7H).

721 The generated CTL and HTL 3D models were further refined using
722 ModRefiner to fix structural gaps followed by GalaxyRefine refinement.
723 Refinement with ModRefiner showed a TM-score of 0.9703 and 0.8934 for the
724 CTL and HTL models respectively, both being close to 1 indicating the initial and
725 the refined models were structurally similar. For the CTL MEV model refinement,
726 the sore of models 1 for different parameters were, Rama favored was 90.8%,
727 GDT-HA was 0.9596, RMSD was 0.385, MolProbity was 2.673, Clash score was
728 29.9, and Poor rotamers was 1.8. Likewise for HTL MEV model refinement, the
729 sore of models 1 for different parameters were, Rama favoured was 88.5%,
730 GDT-HA was 0.9463, RMSD was 0.419, MolProbity was 2.811, Clash score was
731 38.4, and Poor rotamers was 1.6. Here, MolProbity shows the log-weighted
732 combination of the clash score, percentage Ramachandran not favoured and the
733 percentage bad side-chain rotamers. After refinement, all the mentioned
734 parameters were significantly improved in comparison to the initial CTL and HTL
735 MEV models (Supplementary table S10).
736
737 *Validation of CTL and HTL MEVs refined models.* Both the CTL and HTL model
738 were analysed with the RAMPAGE analysis tool after refinement. The refined
739 CTL MEV model has 91.5% residues in favored regions, 6.8% residues in
740 allowed regions, and only 1.7% residues in the outlier region; while the refined
741 HTL MEV model was found to have 89.5% of residues in favored region, 8.7%
742 residues in allowed region, and 1.9% residues in the outlier region (Figure 7E,
743 7J).

744

745 *Discontinuous B-cell epitope prediction from MEVs.* Discontinuous B-cell
746 epitopes were further predicted from the final refined 3D models of CTL and HTL
747 MEVs utilizing the ElliPro tool on IEDB server. The screening revealed that the
748 CTL MEV carries 3 and the HTL MEV has 2 potential discontinuous epitopes.
749 The PI (Protrusion Index) score of the CTL MEV discontinuous B cell epitopes
750 ranges from 0.682 to 0.747 and that of HTL MEV it ranges from 0.687 to 0.745
751 (Supplementary table S11, Figure 7D & 7I). The higher PI score indicates a
752 greater potential of the discontinuous B cell epitope.

753

754 ***Molecular interaction analysis of MEVs with TLR-3 .***

755 The refined models of CTL and HTL MEVs were further studied for their
756 molecular interaction with the ectodomain (ECD) of human TLR-3. Therefore,
757 molecular docking of CTL and HTL MEVs model with the TLR-3 crystal structure
758 model (PDB ID: 2A0Z) was performed utilizing the PatchDock tool. Generated
759 docking conformation with highest scores of 22382 and 18264 for CTL and HTL
760 MEVs, respectively were selected for further studies. The highest docking score
761 predicted with the PatchDock tool indicates the best geometric shape
762 complementarity fitting conformation of MEVs and the TLR-3 receptor. Both, the
763 CTL and HTL MEVs were fitting into the ectodomain region of TLR-3 after
764 docking (Figue 8A & 8E). The CTL and HTL MEVs have shown to form multiple
765 hydrogen bonds within the ectodomain cavity region of TLR-3. Further, the
766 molecular dynamics simulation study was also performed for the docked

767 complexes of both the MEVs and TLR-3. In MD simulations both the complexes
768 have shown reasonably stable RMSD value between ~ 0.2 to 0.5 nm for a given
769 time window of 10 ns at invariable pressure (~ 1 bar) and temperature (~ 300 K)
770 (Figure 8B & 8F). The reasonably invariant radius of gyration (Rg) of both MEVs-
771 TLR-3 complexes (Figure 8C & 8G), and RMS fluctuation (RMSF) for all the
772 atoms in both the complexes (Figure 8D & 8H) indicate that the MEVs-TLR3
773 complexes are stable. The B-factor analysis of MEVs-TLR3 complexes was also
774 performed. The B-factor indicates the displacement of the atomic positions from
775 an average (mean) value i.e. the more flexible an atom is the larger the
776 displacement from the mean position will be (mean-squares displacement)
777 (Figure 8A and 8E). The areas with high B-factors are colored red (hot), while low
778 B-factors are colored blue (cold) (VIBGYOR presentation). The B-factor of most
779 of the regions of MEVs-TLR3 complexes indicates the stable nature of the
780 complexes while a very small region is found to be fluctuating. The results
781 suggest a stable complex formation tendency for both the CTL and HTL MEVs
782 with the ectodomain of the human TLR-3 receptor.
783
784 **Figure 8. Molecular Docking and dynamics simulation study of CTL and**
785 **HTL MEVs with TLR-3. (A)** CTL and **(E)** HTL MEVs (VIBGYOR) docked
786 complex with TLR-3 (gray). Both the complexes are forming several hydrogen
787 bonds in the MEV and TLR-3 interface, as shown by green dots. B-Factor of the
788 docked MEVs is shown by a rainbow (VIBGYOR) presentation. The regions in
789 blue being indicated stable and the region in red indicate unstable. In the above

790 complexes, most of the region of docked MEVs is in blue and with the very small
791 region is green, yellow or orange, hence the complexes are predicted to be very
792 stable. **(B)** and **(F),** RMSD as generated by the molecular Dynamics simulation
793 study of CTL, HTL MEVs and TLR-3 complexes. **(C)** & **(G)** Rg (radius of gyration)
794 across the time window of 10 nanosecond. **(D)** & **(H)**, RMS fluctuation for all the
795 atoms of the CTL, HTL MEVs and TLR-3 complexes.
796
797 ***In-silico* analysis of MEVs for cloning and expression**
798 *Analysis of cDNA of both the MEVs for cloning and expression in mammalian cell*
799 *line.* Complementary DNA codon optimized for CTL and HTL expression in
800 mammalian host cell lines (Human) was generated with the Java Codon
801 Adaptation Tool. The generated optimized cDNA's for both the MEVs were also
802 analysed by utilizing the GenScript Rare Codon Analysis Tool revealing aGC
803 content of optimized CTL-MEV cDNA of 69.79% and a CAI (Codon Adaptation
804 Index) score of 1.00 with 0% tandem rare codons. Likewise, the GC content of
805 the optimized HTL-MEV cDNA was 70.69%, CAI score was 1.00 with 0% tandem
806 rare codons. Since for higher possibility for cDNA expression in human
807 expression system, the GC content of a cDNA should be within the range of 30%
808 to 70%, the CAI score should be between 0.8-1.0, and the tandem rare codon
809 frequency that indicates the presence of low-frequency codons, should be <30%,
810 the cDNA constructs of CTL and HTL MPVs are expected to have high potential
811 for expression in human expression system. The tandem rare codons may hinder
812 the proper expression of the cDNA or even interrupt the translational machinery

813 of the chosen expression system. According to the GenScript Rare Codon
814 Analysis the cDNA of both the MEVs satisfy all the mentioned parameters and
815 are predicted to have high expression in the mammalian host cell line (Human).
816

817 **CONCLUSION**

818 In the present study, we have designed and validated two multi-epitope
819 vaccines derived from CTL and HTL epitopes. The selected peptides show
820 significant sequence overlap with screened linear B cell epitopes. Both the CTL
821 and HTL MEVs tertiary models carry potential discontinuous B cell epitopes and
822 INF-γ epitopes. Consequently, the designed MEVs might have the potential to
823 elicit profound humoral and cellular immune responses. Human β-Defensin 2 and
824 3 fused to the N and C terminal ends of both the MEVs serving as adjuvants to
825 enhance the immune response. The identified epitopes for the CTL and HTL
826 MEVs were also validated by molecular docking and MD simulation studies to
827 test the interaction with their respective HLA allele binders. Molecular interaction
828 of the selected CTL epitopes within the TAP transporter cavity was also
829 evaluated indicating a favorable transport of epitopes from cytoplasm to lumen of
830 Endoplasmic Reticulum for further presentation on cell surface by Golgi bodies.
831 Our analysis of the shortlisted CTL and HTL epitopes combined revealed
832 coverage of 97.88% world human population. The molecular interaction analysis
833 of both the CTL and HTL MEVs with the immunoreceptor TLR3 showed
834 structural fit of the MEVs into the ectodomain of TLR3 receptor cavity and the MD
835 simulations indicate a very stable complex formation. Since both the CTL and

836 HTL MEVs carry CTL, HTL as well as discontinuous B cell epitopes, the
837 combined administration of both the MEVs, is predicted to elicit both the humoral
838 as well as cell-mediated immune response. The cDNA for both the MEVs were
839 designed considering codon-bias for the expression in mammalian host cell lines
840 (Human). The cDNAs were optimized with respect to their GC content and zero
841 tandem rare codons to gain high expression in mammalian host cell lines
842 (Human). In future experimental studies the designed CTL and HTL MEVs
843 should be cloned, expressed and validated *in-vivo* and in animal trails as
844 potential vaccine candidates against NiV infection.
845

846 **Supplementary figure S1.** Workflow chart.
847
848 **Supplementary figure S2. (A)** Rg (radius of gyration) for the CTL epitope – HLA
849 class I allele complexes, across the time window of 1 nano second. **(B)** Rg for
850 the HTL epitope – HLA class II allele complexes, across the time window of 1
851 nano second.
852
853 **Supplementary figure S3. (A)** RMS fluctuation in nanometers for all the atoms
854 of the CTL epitope – HLA class I allele complexes. **(B)** RMS fluctuation in
855 nanometers for all the atoms of the HTL epitope – HLA class II allele complexes.
856
857 **Supplementary figure S4. (A)** B-Factor of CTL epitope – HLA class I allele
858 complexes **(B)** B-Factor of HTL epitope – HLA class II allele complexes. Epitopes
859 are shown in sticks and HLA alleles are shown in gray cartoons. B-factor is
860 indicated by rainbow (VIBGYOR) colour, blue for stable region and red for most
861 unstable region.
862

863 **Supplementary table S1. Protein sequence retrieval, tertiary structures**
864 **retrieval and homology modeling of nine Nipah proteins.** Nipah protein
865 sequences were retrieved from NCBI. Available structure files (pdb) for Nipah
866 proteins were retrieved from RCSB PDB. Nipah proteins with no tertiary structure
867 available were subjected to homology modeling by Swissmodel.
868

869 **Supplementary table S2. Homology modeling for HLA alleles.** Tertiary
870 structure of HLA alleles were modeled by homology modeling using SwissModel

server. Templates were chosen with highest sequence identity. Generated models with acceptable QMEAN value were chosen for further studies.

Supplementary table S3. Shortlisted high scoring CTL epitopes. Selected high scoring CTL epitopes and their respective HLA alleles binders are listed. *In-silico* analysis have shown all the selected epitopes to be non-toxic (Non-Toxin) as well as they show significant conservancy. ToxinPred analysis is based on the ToxinPred main dataset used by "ToxinPred" algorithm to predict toxicity of any unknown peptide. # Epitope match with previous studies indicating consensus in epitope screening by different approaches and methods.

Supplementary table S4. Shortlisted high scoring HTL epitopes. Selected high scoring HTL epitopes and their respective HLA alleles binders are listed above. *In-silico* analysis have shown all the selected epitopes to be non-toxic (Non-Toxin) as well as they show significant conservancy.

Supplementary table S5. World population coverage by the shortlisted CTL and HTL epitopes combined. With a standard deviation of 21.97 on an average 97.88% of world population could be covered by the joint administration of selected CTL and HTL epitopes as vaccine candidate.

Supplementary table S6. Shortlisted B Cell epitopes. BepiPred Linear B Cell epitopes showing sequence overlap with CTL and HTL epitopes are shortlisted above. *In-silico* analysis have shown all the selected epitopes to be non-toxic (Non-Toxin) as well as they show significant amino acid sequence conservancy. # Epitope match with previous studies indicating consensus in epitope screening by different approaches and methods.

Supplementary table S7. CTL epitope prediction. Detailed scoring of all screened CTL epitopes and their respective HLA class I allele binders. CTL epitopes were chosen on the basis of high "Total score" and higher number of HLA allele binders. Total score is a combined score of TAP score, MHC score, Proteasome score and Processing score.

Supplementary table S8. HTL epitope prediction. Percentile rank of HTL epitopes and their respective HLA class II allele binders. HTL epitopes were screened on the basis of percentile rank (lower the percentile number, higher the rank) and larger number of HLA allele binders. Last column show the method used for epitope screening.

Supplementary table S9. INF-γ epitopes from CTL and HTL MEVs. INF-γ inducing (POSITIVE) epitopes from CTL and HTL MEVs were screened by using "Motif and SVM hybrid" (MERCI & SVM) approaches.

Supplementary table S10. Refinement models of CTL and HTL MEVs. CTL and HTL MEVs models were refined by GalaxyWEB server and used for further studies. After refinement in particular Rama favored residues increased significantly.

Supplementary table S11. B Cell discontineous epitopes of CTL & HTL MEVs. Discontinuous B Cell epitopes predicted by ElliPro (IEDB) from CTL & HTL MEVs.

ACKNOWLEDGEMENTS:

We acknowledge the Advance Instrumentation Research Facility (AIRF) at Jawaharlal Nehru University, New Delhi for providing the advance computational facility to conduct experiments. We also acknowledge Indian Foundation for Fundamental Research (IFFR) for providing resources and funding

FUNDING:

Indian Foundation for Fundamental Research (IFFR)

AUTHOR CONTRIBUTION:

Idea conceived, methodology designed and performed by S.S., critical data analysis and scientific writing was done by S.S. and M.K., facility of MD simulation was provided by A.K.S, draft was finalized by S.S., M.K., A.K.S.

ADDITIONAL INFORMATION:

Authors declare to have no competing interests.

FIGURES:

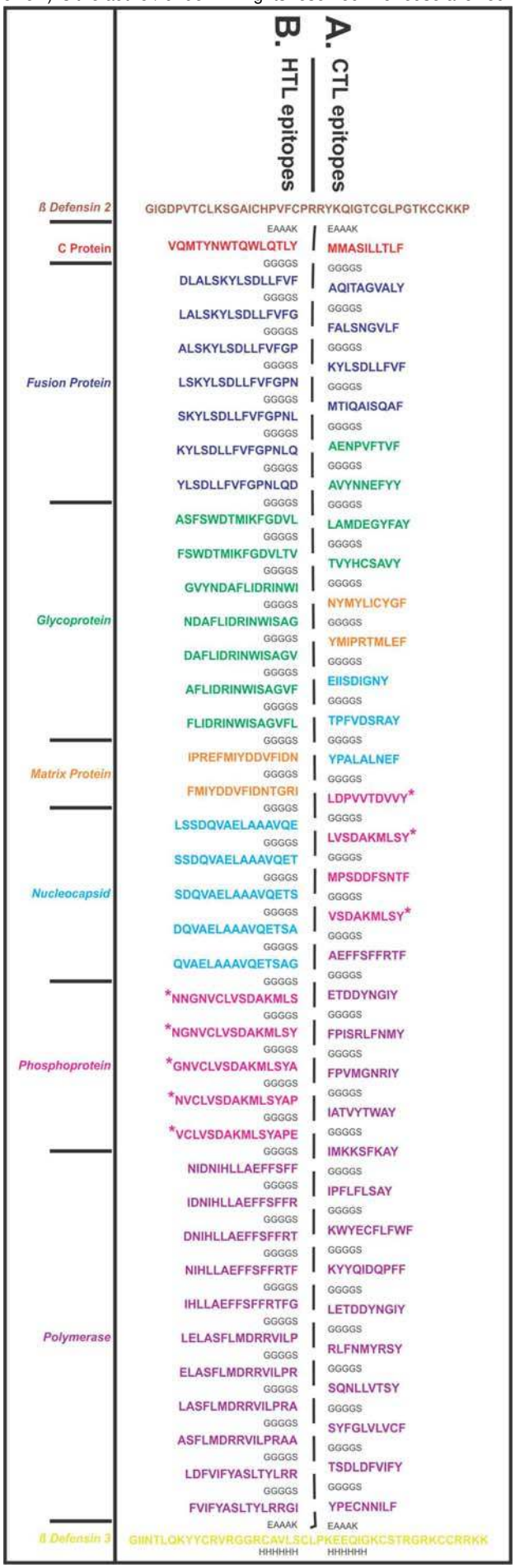


Figure 1. Design of Multi-Epitope Vaccine (MEVs). (A) CTL and **(B)** HTL epitopes were linked by the short peptide linker 'GGGGS'. Human β Defensin 2 and β Defensin 3 were used as an adjuvant at the N and C terminals respectively. The short peptide EAAAK was used to link the β Defensin 2 and β Defensin 3. Epitopes from different proteins were coloured in different colours. C terminal 6xHis is designed as His tag. * Epitopes common to Phosphoprotein, V Protein and W protein.

Overlapping regions of predicted B cell Bepipred linear Epitope and epitopes predicted by other protein sequence and structure based methods		Overlapping regions of sequence based B cell linear epitope prediction					Overlapping regions of structure based B cell epitope prediction		
		Chou & Fasman Beta-Turn	Emini Surface Accessibility	Karplus & Schulz Flexibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity	DiscoTope: discontinuous epitopes	EliiPro	
								Linear	Discontinuous
Fusion Protein	215 GPNLQDPVSNSM 226	215-226		215-226	215-216	217-226			
Glycoprotein	271 WTPPNPNT 278	271-278	271-278	272-278		272-278			271-278
Glycoprotein	459 SWDTMI 464	459-460							463-464
Glycoprotein	529 NQTAE 533	529-532	529-533	529-533		529-533			
Matrix Protein	184 SGIYM 188	184-188		184-188		184-188			
Matrix Protein	330 SIPREFMIY 338	330-335		330-336		331-334	332-333	332-333	333-338
Matrix Protein	340 DVFIDNTGRI 349	341-349		340-349		340-344			340-349
Phosphoprotein	125 GYGFTSSPERGWSDYTSGA 143	125-143	130-136	128-143		127-143			
Phosphoprotein	163 IAVSKEDR 170	166-170	166-170	163-170	163-166	165-170			
Polymerase	346 NIDN 349	348-349							
Polymerase	364 GHPILE 369	364-367			365-369				366-369
Polymerase	1095 DKSFDEDLEL 1104	1096-1103	1096-1101	1095-1104		1096-1103			1095-1104
Polymerase	1108 LMDR 1111								1108-1111
Polymerase	1409 LRLETDDYNG 1418	1412-1418	1410-1416	1412-1418		1411-1418	1409-1417	1409-1417	
Polymerase	1982 GFPIS 1986	1982-1986		1982-1986	1982-1986				1182-1986
Polymerase	2004 PVYSNPD 2010	2004-2010	2006-2010	2007-2010	2004-2008	2007-2010			2004-2010

Figure 2. Overlapping regions amongst the linear B cell epitopes predicted by the BepiPred method and other seven different B cell epitopes prediction methods. B cell epitopes predicted by the BepiPred method and other different protein sequence based (Chou.., Emini.., Karplus.., Kolaskar.., and Parker..) and protein structure based (DiscoTope and ElliPro) prediction methods were found to have significant consensus. Consensus overlapping regions of BepiPred epitopes are underlined by the different colour, corresponding to respective prediction method.

Clusters of overlapping CTL, HTL & B cell epitopes

Fusion Protein	----------MTIQAISQAF- ------KYLSDLLFVF---------- -------YLSDLLFVFGPNLQD---------- ------KYLSDLLFVFGPNLQ---------- -----SKYLSDLLFVFGPNL---------- ----LSKYLSDLLFVFGPN---------- ---ALSKYLSDLLFVFGP---------- --LALSKYLSDLLFVFG---------- -DLALSKYLSDLLFVF---------- ----------GPNLQDPVSNSM---------- 199-236 LDLALSKYLSDLLFVFGPNLQDPVSNSMTIQAISQAFG
Glycoprotein	----------AENPVFTVF- -------FLIDRINWISAGVFL---------- ------AFLIDRINWISAGVF---------- -----DAFLIDRINWISAGV---------- ----NDAFLIDRINWISAG---------- -GVYNDAFLIDRINWI---------- --------TVYHCSAVY------- ----------AVYNNEFYY- -WTPPNPNT---------- 270-293 VWTPPNPNTVYHCSAVYNNEFYYV ----FSWDTMIKFGDVLTV-- --ASFSWDTMIKFGDVL---- -----SWDTMI---------- 454-474 YQASFSWDTMIKFGDVLTVNP ----------NQTAE-------- 505-541 EGVYNDAFLIDRINWISAGVFLDSNQTAENPVFTVFK
Matrix Protein	-----FMIYDDVFIDNTGRI- -IPREFMIYDDVFIDN----- ---YMIPRTMLEF- ----------DVFIDNTGRI- -SGIYM---------- SIPREFMIY---------- 183-197 DSGIYMIPRTMLEFR 330-350 SIPREFMIYDDVFIDNTGRIL
Nucleocapsid	------QVAELAAAVQETSAG-- -----DQVAELAAAVQETSA--- ----SDQVAELAAAVQETS---- ---SSDQVAELAAAVQET----- --LSSDQVAELAAAVQE------ 382-404 LGLSSDQVAELAAAVQETSAGRQ
Phosphoprotein	----------LVSDAKMLSY---------- ----------VSDAKMLSY---------- ----------VCLVSDAKMLSYAPE---------- ----------NVCLVSDAKMLSYAP---------- ----------GNVCLVSDAKMLSYA---------- ----------NGNVCLVSDAKMLSY---------- ----------NNGNVCLVSDAKMLS---------- ----------IAVSKEDR-- --GYGFTSSPERGWSDYTSGA---------- 123-172 CTGYGFTSSPERGWSDYTSGANNGNVCLVSDAKMLSYAPEIAVSKEDRET
Polymerase	--------AEFFSFFRTF------- -----IHLLAEFFSFFRTFG------ ----NIHLLAEFFSFFRTF------- ---DNIHLLAEFFSFFRT-------- --IDNIHLLAEFFSFFR--------- -NIDNIHLLAEFFSFF---------- -NIDN---------- ----------GHPILE- 345-370 LNIDNIHLLAEFFSFFRTFGHPILEA ----------ASFLMDRRVILPRAA----- ----------LASFLMDRRVILPRA------ ---------ELASFLMDRRVILPR------- ---------LELASFLMDRRVILP-------- ----------LMDR---------- --DKSFDEDLEL---------- 1093-1124 FHDKSFDEDLELASFLMDRRVILPRAAHEILD ----ETDDYNGIY-- ---LETDDYNGIY-- -LRLETDDYNG----- 1408-1422 RLRLETDDYNGIYHL
Polymerase	----------SYFGLVLVCF--------- ------RLFNMYRSY---------- -TSDLDFVIFY---------- --FPISRLFNMY---------- ------FVIFYASLTYLRRGI- ----------PVYSNPD-- ----LDFVIFYASLTYLRR--- -GFPIS---------- 1715-1736 NTSDLDFVIFYASLTYLRRGII 1981-2012 PGFPISRLFNMYRSYFGLVLVCFPVYSNPDST

Figure 3. Overlapping CTL, HTL and B cell epitopes. Multiple sequence alignment performed by Clustal Omega at EBI to identify the consensus overlapping regions of CTL (red), HTL (blue) and B cell epitopes (green) amongst shortlisted epitopes. Epitopes with overlapping region amongst all the three types of epitopes (CTL, HTL and B Cell epitopes), epitopes with full sequence overlap and epitopes with the highest number of HLA allele binders were chosen for further studies (encircled).

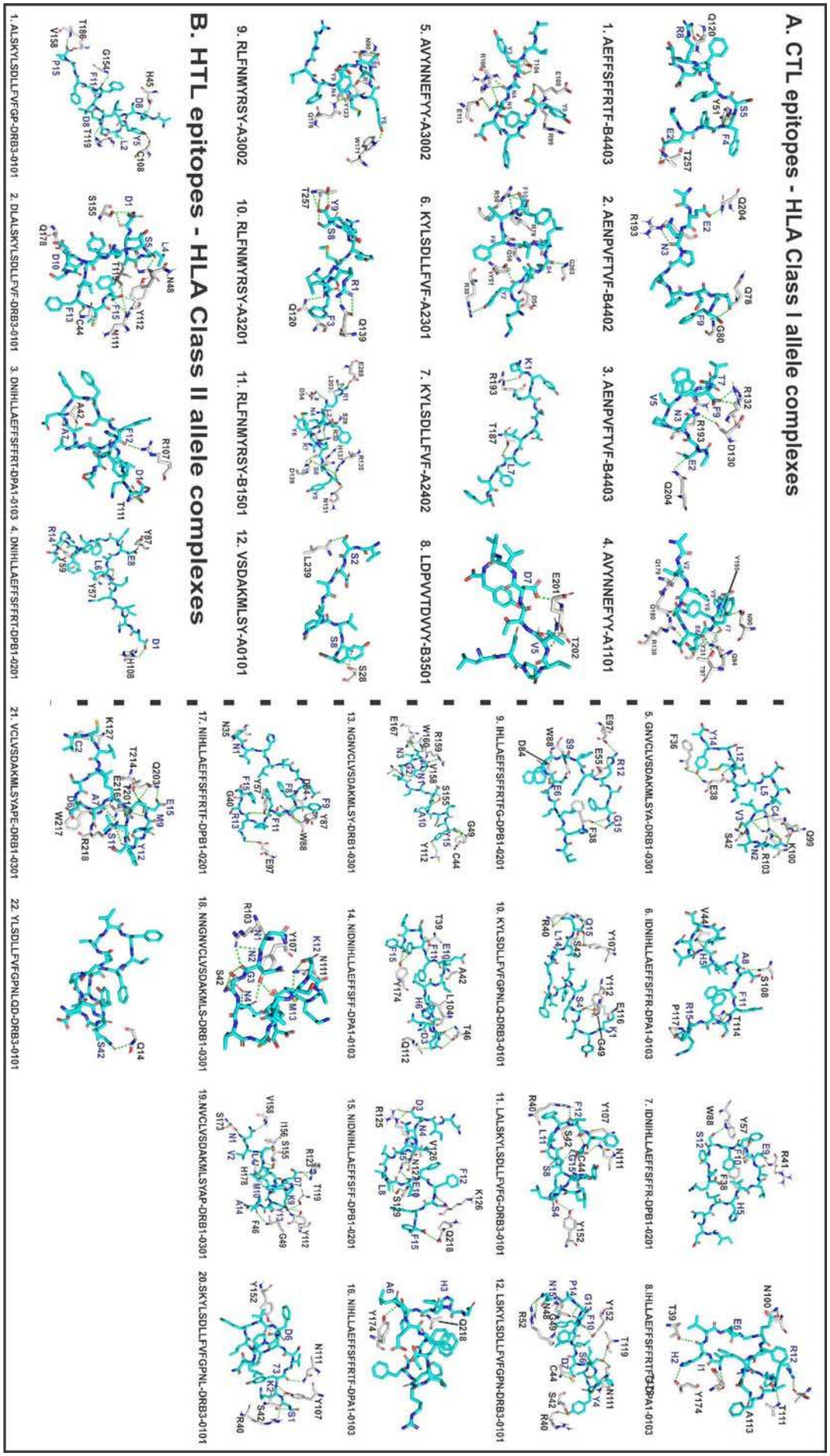


Figure 4 (A). Molecular Docking analysis of CTL epitopes and HLA alleles. Molecular docking of selected CTL epitopes (cyan sticks) with their respective HLA class I allele binders (gray sticks). The study shows the docked complexes to have significantly negative binding energy along with hydrogen bonds (green dots) formation in the complex interface. **(B) Molecular Docking analysis of HTL epitopes and HLA alleles.** Molecular docking of selected HTL epitopes (cyan sticks) with their respective HLA class II allele binders (gray sticks). The study shows the docked complexes to have significantly negative binding energy along with hydrogen bonds (green dots) formation in the complex interface.

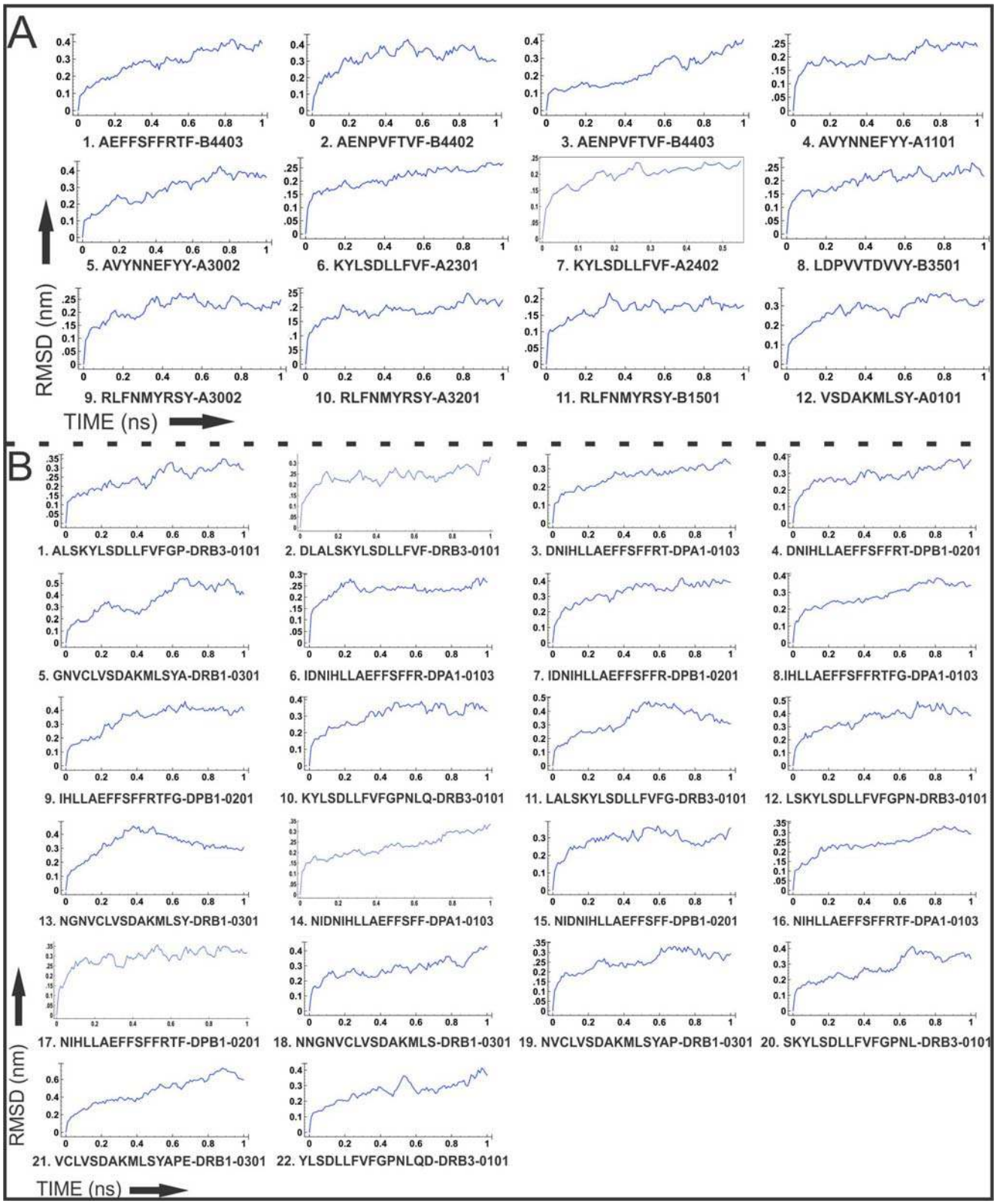


Figure 5 (A). Molecular Dynamics simulation analysis of CTL epitopes and HLA allele complexes. Molecular Dynamics simulation study reveals a stable nature of the CTL-HLA allele complexes throughout 0.5-1 ns time window with reasonably invariable RMSD. **(B) Molecular Dynamics simulation analysis of CTL epitopes and HLA allele complexes.** Molecular Dynamics simulation study reveals a stable nature of the HTL-HLA allele complexes throughout 0.5-1 ns time window with reasonably invariable RMSD.

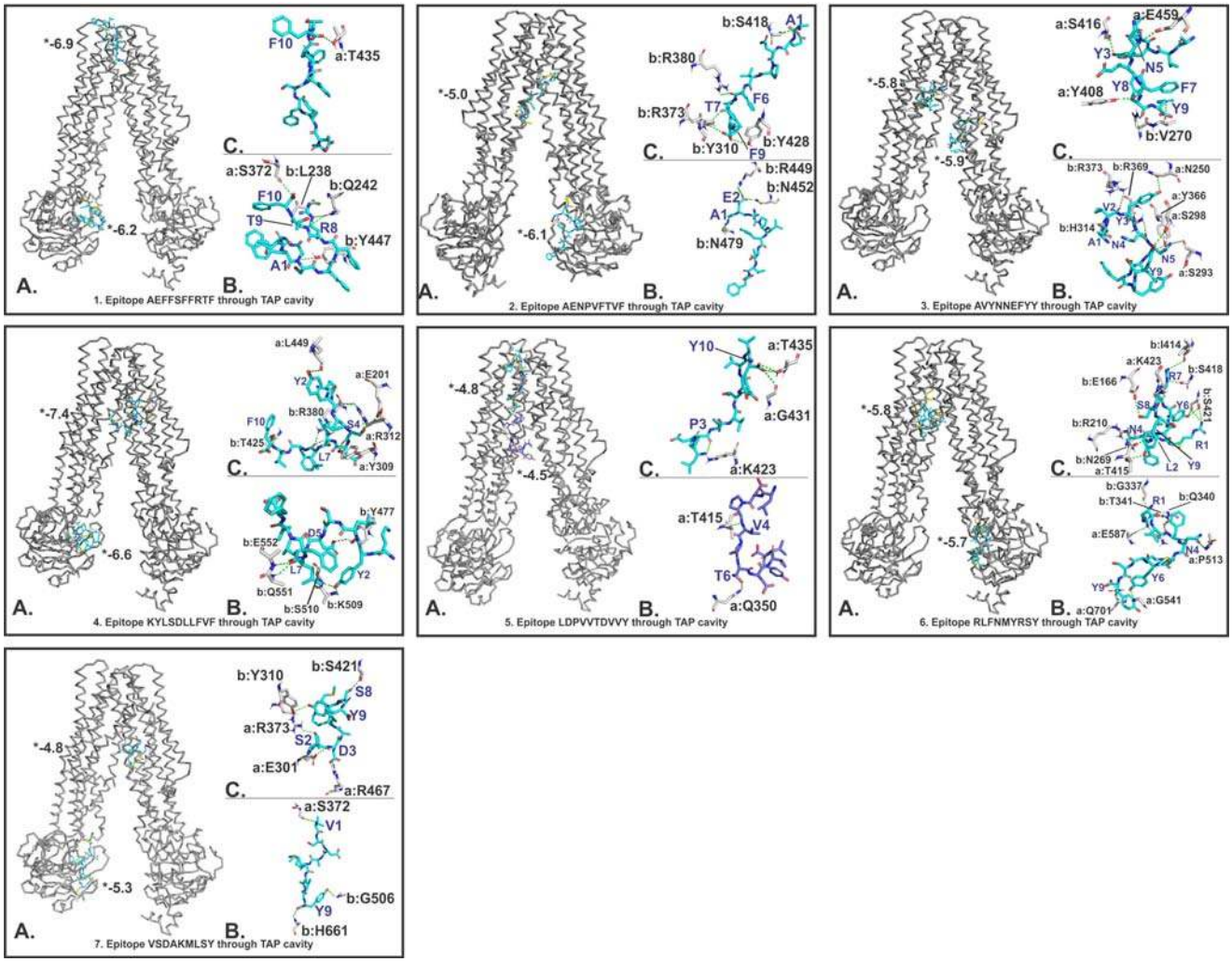
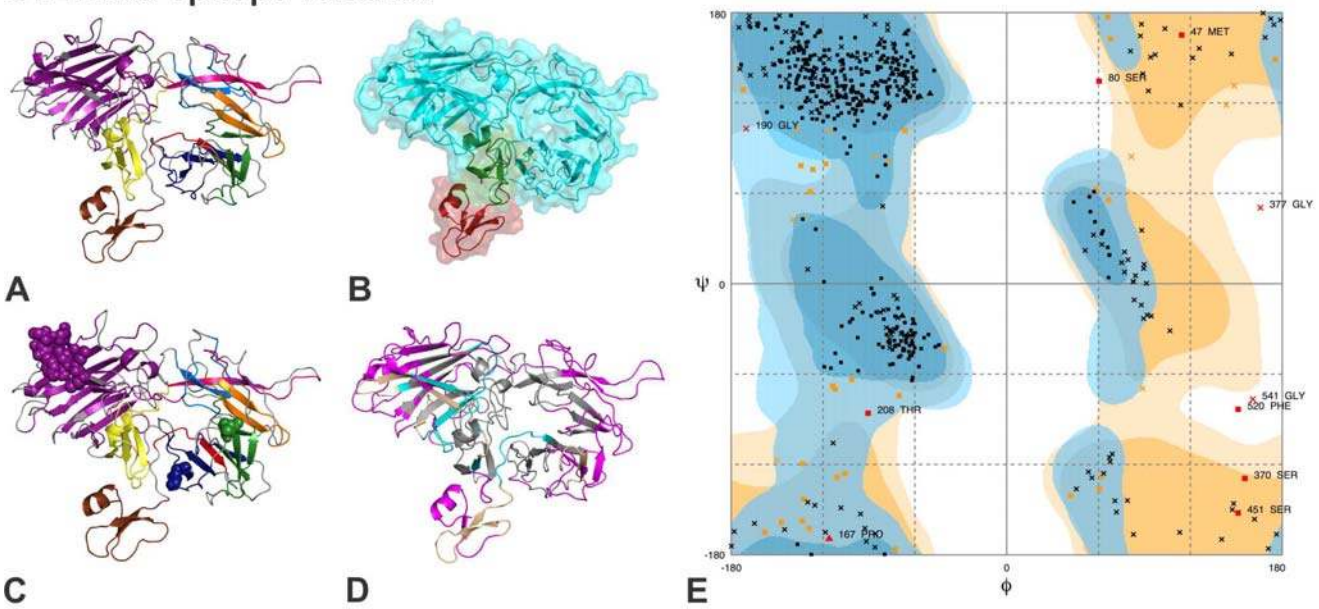


Figure 6. Molecular docking analysis of CTL epitopes within the TAP transporter cavity. Molecular interaction of CTL epitopes (cyan sticks) within the TAP cavity (gray ribbon/sticks) is shown in detail. For every panel of epitope-TAP complex, **(A)** shows the binding of epitope at two different sites within TAP cavity, **(B)** and **(C)** show detailed molecular interaction between epitopes and TAP cavity; (a, b) show chain A and B of TAP transporter. H-bonds are shown in yellow dots. (*) Binding energy, shown in kcal/mol.

CTL multi-epitope vaccine:



HTL multi-epitope vaccine:

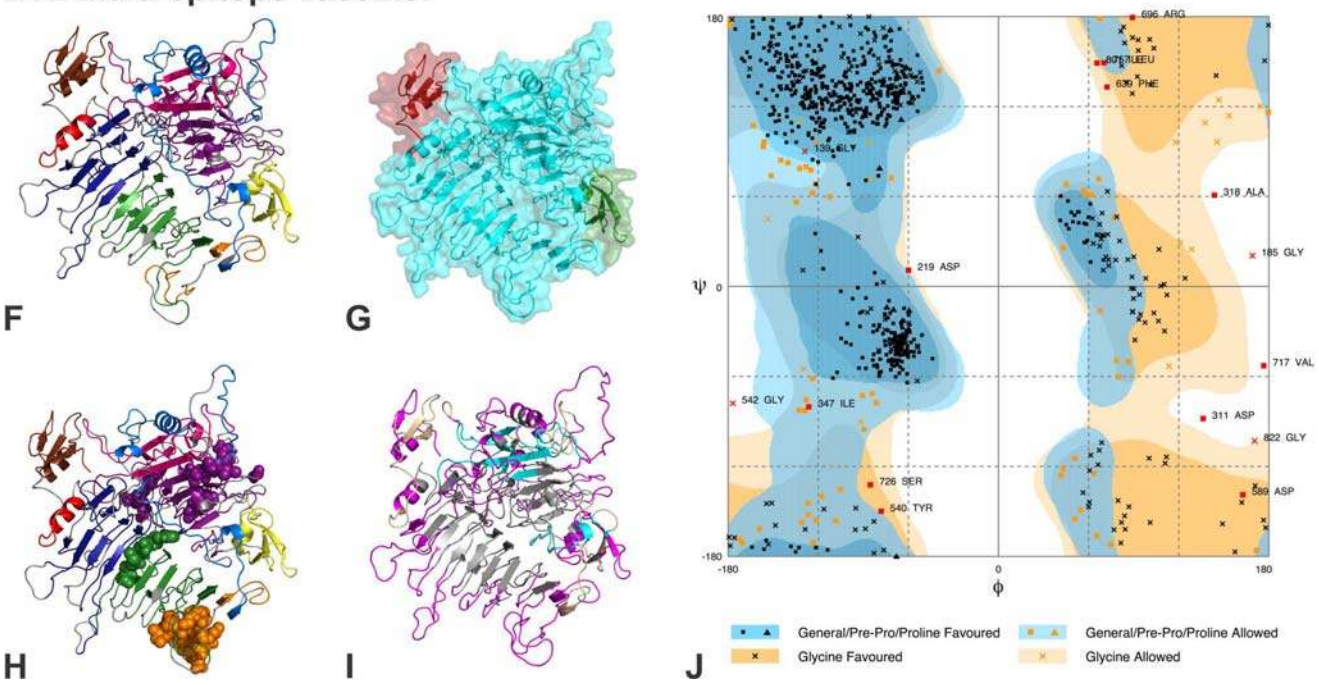


Figure 7. Tertiary structure modelling of CTL and HTL Multi-Epitope Vaccines. (A) & (F): Tertiary structural models of CTL and HTL MEVs showing epitopes in different colours corresponds to as in Fig. 1. **(B) & (G)**: Show the different domains of CTL and HTL MEVs. **(C) & (H)**: The overlapping linear B cell epitope region present in CTL and HTL MEVs, shown by spheres. **(D) & (I):** From the CTL and HTL MEVs, the INF-γ inducing epitopes are shown in cyan, discontinuous B Cell epitopes are shown in magenta and the region common amongst INF-γ and discontinuous B Cell epitopes are shown in wheat colour. **(E) & (J)**: RAMPAGE analysis of the refined CTL and HTL MEV models.

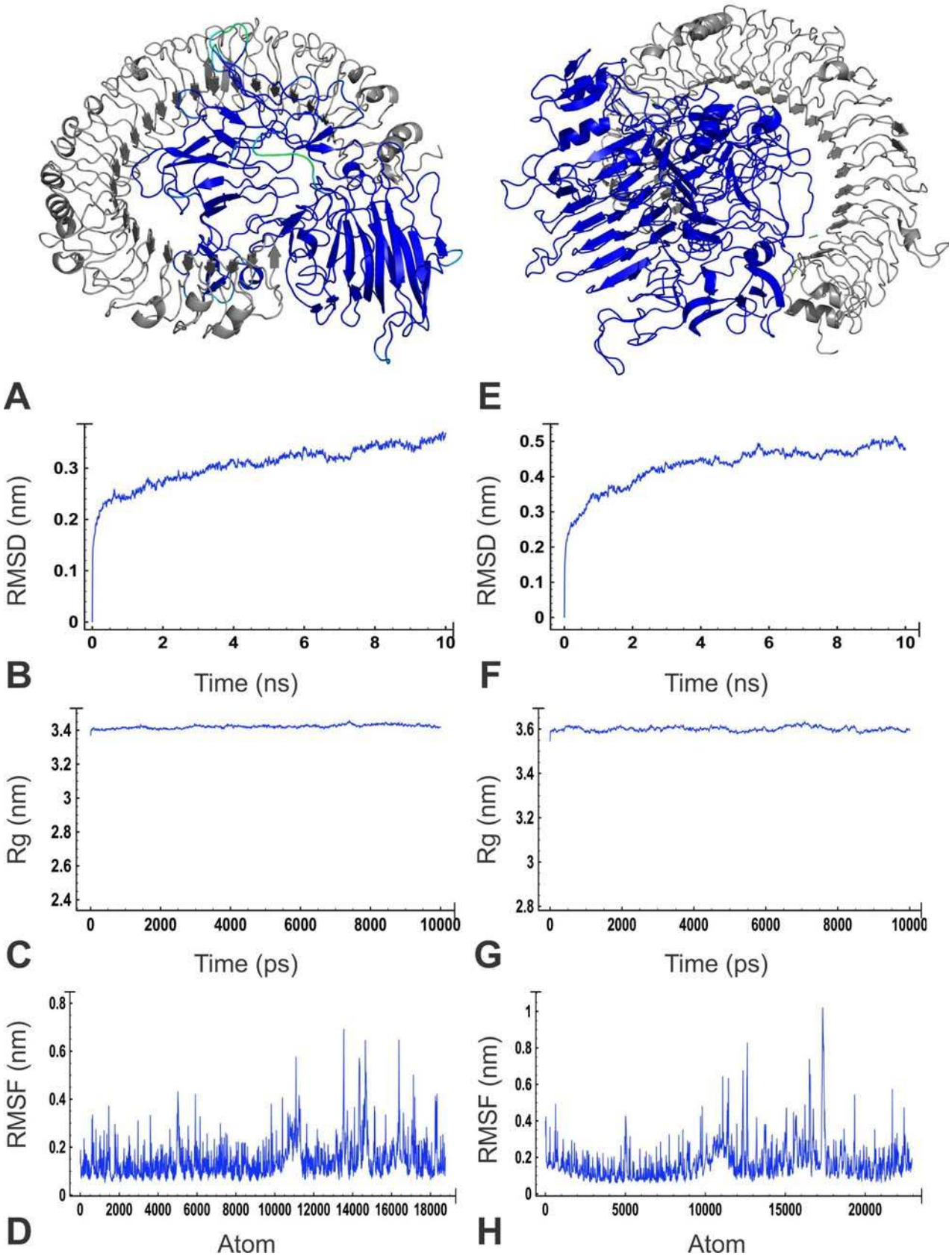


Figure 8. Molecular Docking and dynamics simulation study of CTL and HTL MEVs with TLR-3. (A) CTL and **(E)** HTL MEVs (VIBGYOR) docked complex with TLR-3 (gray). Both the complexes are forming several hydrogen bonds in the MEV and TLR-3 interface, as shown by green dots. B-Factor of the docked MEVs is shown by a rainbow (VIBGYOR) presentation. The regions in blue being indicated stable and the region in red indicate unstable. In the above complexes, most of the region of docked MEVs is in blue and with the very small region is green, yellow or orange, hence the complexes are predicted to be very stable. **(B)** and **(F),** RMSD as generated by the molecular Dynamics simulation study of CTL, HTL MEVs and TLR-3 complexes. **(C)** & **(G)** Rg (radius of gyration) across the time window of 10 nanosecond. **(D)** & **(H)**, RMS fluctuation for all the atoms of the CTL, HTL MEVs and TLR-3 complexes.

1548 **REFERENCES:**

1549 1. Angeletti, S., Presti, A.L., Cella, E. and Ciccozzi, M., 2016. Molecular
1550 epidemiology and phylogeny of nipah virus infection: a mini review. Asian
1551 Pacific journal of tropical medicine, 9(7), pp.630-634.
1552 2. Aguilar, H.C., Henderson, B.A., Zamora, J.L. and Johnston, G.P., 2016.
1553 Paramyxovirus glycoproteins and the membrane fusion process. Current
1554 clinical microbiology reports, 3(3), pp.142-154.
1555 3. Ang, B.S., Lim, T.C. and Wang, L., 2018. Nipah Virus Infection. Journal of
1556 clinical microbiology, pp.JCM-01875.
1557 4. WHO Report, Surveillance and outbreak alert, Nipah virus;
1558 https://www.who.int/health-topics/nipah-virus-infection#tab=tab_1
1559 5. Plowright RK, Becker DJ, Crowley DE, Washburne AD, Huang T, Nameer PO,
1560 Gurley ES, Han BA. Prioritizing surveillance of Nipah virus in India. PLoS
1561 Negl Trop Dis. 2019 Jun 27;13(6):e0007393. doi:
1562 10.1371/journal.pntd.0007393. PMID: 31246966; PMCID: PMC6597033.
1563 6. Thomas, B., Chandran, P., Lilabi, M. P., George, B., Sivakumar, C. P.,
1564 Jayadev, V. K., Bindu, V., Rajasi, R. S., Vijayan, B., Mohandas, A., &
1565 Hafeez, N. (2019). Nipah Virus Infection in Kozhikode, Kerala, South
1566 India, in 2018: Epidemiology of an Outbreak of an Emerging Disease.
1567 Indian journal of community medicine : official publication of Indian
1568 Association of Preventive & Social Medicine, 44(4), 383–387.
1569 https://doi.org/10.4103/ijcm.IJCM_198_19
1570 7. Mathieu, C., Guillaume, V., Volchkova, V.A., Pohl, C., Jacquot, F., Looi, R.Y.,
1571 Wong, K.T., Legras-Lachuer, C., Volchkov, V.E., Lachuer, J. and Horvat,
1572 B., 2012. Nonstructural Nipah virus C protein regulates both the early host
1573 proinflammatory response and viral virulence. Journal of virology, pp.JVI-
1574 01203.
1575 8. Liu, Q., Chen, L., Aguilar, H.C. and Chou, K.C., 2018. A stochastic assembly
1576 model for Nipah virus revealed by super-resolution microscopy. Nature
1577 communications, 9(1), p.3050.
1578 9. Johnston, G.P., Contreras, E.M., Dabundo, J., Henderson, B.A., Matz, K.M.,
1579 Ortega, V., Ramirez, A., Park, A. and Aguilar, H.C., 2017. Cytoplasmic
1580 motifs in the nipah virus fusion protein modulate virus particle assembly
1581 and egress. Journal of virology, pp.JVI-02150.
1582 10. Satterfield, B.A., Cross, R.W., Fenton, K.A., Borisevich, V., Agans, K.N.,
1583 Deer, D.J., Graber, J., Basler, C.F., Geisbert, T.W. and Mire, C.E., 2016.
1584 The Nipah virus C and W proteins contribute to respiratory disease in
1585 ferrets. Journal of virology, pp.JVI-00215.
1586 11. Lamp, B., Dietzel, E., Kolesnikova, L., Sauerhering, L., Erbar, S., Weingartl,
1587 H. and Maisner, A., 2013. Nipah virus entry and egress from polarized
1588 epithelial cells. Journal of virology, pp.JVI-02696.
1589 12. Weise, C., Erbar, S., Lamp, B., Vogt, C., Diederich, S. and Maisner, A., 2010.
1590 Tyrosine residues in the cytoplasmic domains affect sorting and fusion
1591 activity of the Nipah virus glycoproteins in polarized epithelial cells.
1592 Journal of virology, 84(15), pp.7634-7641.
1593 13. Ciancanelli, M.J. and Basler, C.F., 2006. Mutation of YMYL in the Nipah virus

1594 matrix protein abrogates budding and alters subcellular localization.
1595 Journal of virology, 80(24), pp.12070-12078.
1596 14. Patch, J.R., Crameri, G., Wang, L.F., Eaton, B.T. and Broder, C.C., 2007.
1597 Quantitative analysis of Nipah virus proteins released as virus-like
1598 particles reveals central role for the matrix protein. Virology journal, 4(1),
1599 p.1.
1600 15. Patch, J.R., Han, Z., McCarthy, S.E., Yan, L., Wang, L.F., Harty, R.N. and
1601 Broder, C.C., 2008. The YPLGVG sequence of the Nipah virus matrix
1602 protein is required for budding. Virology journal, 5(1), p.137.
1603 16. Jordan, P.C., Liu, C., Raynaud, P., Lo, M.K., Spiropoulou, C.F., Symons,
1604 J.A., Beigelman, L. and Deval, J., 2018. Initiation, extension, and
1605 termination of RNA synthesis by a paramyxovirus polymerase. PLoS
1606 pathogens, 14(2), p.e1006889.
1607 17. Ranadheera, C., Proulx, R., Chaiyakul, M., Jones, S., Grolla, A., Leung, A.,
1608 Rutherford, J., Kobasa, D., Carpenter, M. and Czub, M., 2018. The
1609 interaction between the Nipah virus nucleocapsid protein and
1610 phosphoprotein regulates virus replication. Scientific reports, 8(1),
1611 p.15994.
1612 18. Baronti, L., Erales, J., Habchi, J., Felli, I.C., Pierattelli, R. and Longhi, S.,
1613 2015. Dynamics of the intrinsically disordered C-terminal domain of the
1614 Nipah virus nucleoprotein and interaction with the X domain of the
1615 phosphoprotein as unveiled by NMR spectroscopy. ChemBioChem, 16(2),
1616 pp.268-276.
1617 19. Uchida, S., Horie, R., Sato, H., Kai, C. and Yoneda, M., 2018. Possible role
1618 of the Nipah virus V protein in the regulation of the interferon beta
1619 induction by interacting with UBX domain-containing protein1. Scientific
1620 reports, 8(1), p.7682.
1621 20. Ludlow, L.E., Lo, M.K., Rodriguez, J.J., Rota, P.A. and Horvath, C.M., 2008.
1622 Henipavirus V protein association with Polo-like kinase reveals functional
1623 overlap with STAT1 binding and interferon evasion. Journal of virology,
1624 82(13), pp.6259-6271.
1625 21. Park, M.S., Shaw, M.L., Munoz-Jordan, J., Cros, J.F., Nakaya, T., Bouvier,
1626 N., Palese, P., García-Sastre, A. and Basler, C.F., 2003. Newcastle
1627 disease virus (NDV)-based assay demonstrates interferon-antagonist
1628 activity for the NDV V protein and the Nipah virus V, W, and C proteins.
1629 Journal of virology, 77(2), pp.1501-1511.
1630 22. Sakib, M.S., Islam, M., Hasan, A.K.M. and Nabi, A.H.M., 2014. Prediction of
1631 epitope-based peptides for the utility of vaccine development from fusion
1632 and glycoprotein of nipah virus using in silico approach. Advances in
1633 bioinformatics, 2014.
1634 23. Guillaume, V., Contamin, H., Loth, P., Georges-Courbot, M.C., Lefeuvre, A.,
1635 Marianneau, P., Chua, K.B., Lam, S.K., Buckland, R., Deubel, V. and
1636 Wild, T.F., 2004. Nipah virus: vaccination and passive protection studies in
1637 a hamster model. Journal of virology, 78(2), pp.834-840.
1638 24. Koyuncu, O.O., Hogue, I.B. and Enquist, L.W., 2013. Virus infections in the
1639 nervous system. Cell host & microbe, 13(4), pp.379-393.

1640 25. Griffin, D.E. and Metcalf, T., 2011. Clearance of virus infection from the CNS.
1641 Current opinion in virology, 1(3), pp.216-221.
1642 26. Kong, D., Wen, Z., Su, H., Ge, J., Chen, W., Wang, X., Wu, C., Yang, C.,
1643 Chen, H. and Bu, Z., 2012. Newcastle disease virus-vectored Nipah
1644 encephalitis vaccines induce B and T cell responses in mice and long-
1645 lasting neutralizing antibodies in pigs. Virology, 432(2), pp.327-335.
1646 27. Kamthania, M. and Sharma, D.K., 2015. Screening and structure-based
1647 modeling of T-cell epitopes of Nipah virus proteome: an immunoinformatic
1648 approach for designing peptide-based vaccine. *3 Biotech*, *5*(6), pp.877-
1649 882.
1650 28. Kamthania, M. and Sharma, D.K., 2016. Epitope-based peptides prediction
1651 from proteome of nipah virus. *International Journal of Peptide Research*
1652 *and Therapeutics*, *22*(4), pp.465-470.
1653 29. Ali, M.T., Morshed, M.M. and Hassan, F., 2015. A computational approach
1654 for designing a universal epitope-based peptide vaccine against Nipah
1655 virus. *Interdisciplinary Sciences: Computational Life Sciences*, *7*(2),
1656 pp.177-185.
1657 30. kumar Sharma, S., Srivastava, S., Kumar, A. and Srivastava, V., 2021.
1658 Anticipation of Antigenic Sites for the Goal of Vaccine Designing Against
1659 Nipah Virus: An Immunoinformatics Inquisitive Quest. International Journal
1660 of Peptide Research and Therapeutics, pp.1-13.
1661 31. Dey, S., Roy, P., Dutta, T., Nandy, A. and Basak, S.C., 2018. Rational
1662 Design of Peptide Vaccines for the Highly Lethal Nipah and Hendra
1663 Viruses. *bioRxiv*, p.425819.
1664 32. Krishnamoorthy, P.K., Subasree, S., Arthi, U., Mobashir, M., Gowda, C. and
1665 Revanasiddappa, P.D., 2020. T-cell Epitope-based Vaccine Design for
1666 Nipah Virus by Reverse Vaccinology Approach. *Combinatorial chemistry &*
1667 *high throughput screening*, *23*(8), pp.788-796.
1668 33. Sakib, M.S., Islam, M., Hasan, A.K.M. and Nabi, A.H.M., 2014. Prediction of
1669 epitope-based peptides for the utility of vaccine development from fusion
1670 and glycoprotein of nipah virus using in silico approach. *Advances in*
1671 *bioinformatics*, *2014*.
1672 34. Eshaghi, M., Tan, W.S. and Yusoff, K., 2005. Identification of epitopes in the
1673 nucleocapsid protein of Nipah virus using a linear phage-displayed
1674 random peptide library. *Journal of medical virology*, *75*(1), pp.147-152.
1675 35. Mohammed, A.A., Shantier, S.W., Mustafa, M.I., Osman, H.K., Elmansi, H.E.,
1676 Osman, I.A.A., Mohammed, R.A., Abdelrhman, F.A., Elnnewery, M.E.,
1677 Yousif, E.M. and Mustafa, M.M., 2020. Epitope-based peptide vaccine
1678 against glycoprotein G of Nipah henipavirus using immunoinformatics
1679 approaches. *Journal of immunology research*, *2020*.
1680 36. Gupta, A.K., Kumar, A., Rajput, A., Kaur, K., Dar, S.A., Thakur, A., Megha, K.
1681 and Kumar, M., 2020. NipahVR: a resource of multi-targeted putative
1682 therapeutics and epitopes for the Nipah virus. *Database*, *2020*.
1683 37. Habib, P.T., 2021. Learning from COVID-19 Pandemic: In Silico Vaccine and
1684 Cloning Design Against Nipah Virus by Studying and Analyzing the Whole
1685 Nipah Virus Proteome.

1686 38. Singh, R.K., Dhama, K., Chakraborty, S., Tiwari, R., Natesan, S., Khandia,
1687 R., Munjal, A., Vora, K.S., Latheef, S.K., Karthik, K. and Singh Malik, Y.,
1688 2019. Nipah virus: epidemiology, pathology, immunobiology and advances
1689 in diagnosis, vaccine designing and control strategies–a comprehensive
1690 review. *Veterinary Quarterly*, *39*(1), pp.26-55.
1691 39. Majee, P., Jain, N. and Kumar, A., 2021. Designing of a multi-epitope vaccine
1692 candidate against Nipah virus by in silico approach: a putative prophylactic
1693 solution for the deadly virus. *Journal of Biomolecular Structure and*
1694 *Dynamics*, *39*(4), pp.1461-1480.
1695 40. Ojha, R., Pareek, A., Pandey, R.K., Prusty, D. and Prajapati, V.K., 2019.
1696 Strategic development of a next-generation multi-epitope vaccine to
1697 prevent Nipah virus zoonotic infection. *ACS omega*, *4*(8), pp.13069-
1698 13079.
1699 41. Wilson, S.S., Wiens, M.E. and Smith, J.G., 2013. Antiviral mechanisms of
1700 human defensins. Journal of molecular biology, 425(24), pp.4965-4980.
1701 42. Duits, L.A., Nibbering, P.H., Strijen, E., Vos, J.B., Mannesse-Lazeroms, S.P.,
1702 Sterkenburg, M.A. and Hiemstra, P.S., 2003. Rhinovirus increases human
1703 β-defensin-2 and-3 mRNA expression in cultured bronchial epithelial cells.
1704 Pathogens and Disease, 38(1), pp.59-64.
1705 43. Yang, D., Biragyn, A., Kwak, L.W. and Oppenheim, J.J., 2002. Mammalian
1706 defensins in immunity: more than just microbicidal. Trends in immunology,
1707 23(6), pp.291-296.
1708 44. Biragyn, A., Surenhu, M., Yang, D., Ruffini, P.A., Haines, B.A.,
1709 Klyushnenkova, E., Oppenheim, J.J. and Kwak, L.W., 2001. Mediators of
1710 innate immunity that target immature, but not mature, dendritic cells
1711 induce antitumor immunity when genetically fused with nonimmunogenic
1712 tumor antigens. The Journal of Immunology, 167(11), pp.6644-6653.
1713 45. Duits, L.A., Nibbering, P.H., van Strijen, E., Vos, J.B., Mannesse-Lazeroms,
1714 S.P., van Sterkenburg, M.A. and Hiemstra, P.S., 2003. Rhinovirus
1715 increases human β-defensin-2 and-3 mRNA expression in cultured
1716 bronchial epithelial cells. FEMS Immunology & Medical Microbiology,
1717 38(1), pp.59-64.
1718 46. Kohlgraf, K.G., Pingel, L.C., Dietrich, D.E. and Brogden, K.A., 2010.
1719 Defensins as anti-inflammatory compounds and mucosal adjuvants.
1720 Future microbiology, 5(1), pp.99-113.
1721 47. Antoniou, A.N., Powis, S.J. and Elliott, T., 2003. Assembly and export of
1722 MHC class I peptide ligands. Current opinion in immunology, 15(1), pp.75-
1723 81.
1724 48. Oldham, M.L., Grigorieff, N. and Chen, J., 2016. Structure of the Transporter
1725 associated with antigen processing trapped by herpes simplex virus. eLife,
1726 5, p.e21829.
1727 49. Meena, S.R., Gangwar, S.P. and Saxena, A.K., 2012. Purification,
1728 crystallization and preliminary X-ray crystallographic analysis of the
1729 ATPase domain of human TAP in nucleotide-free and ADP-, vanadate-
1730 and azide-complexed forms. Acta Crystallographica Section F: Structural
1731 Biology and Crystallization Communications, 68(6), pp.655-658.

1732 50. Delneste, Y., Beauvillain, C. and Jeannin, P., 2007. Innate immunity:
1733 structure and function of TLRs. Medecine sciences: M/S, 23(1), pp.67-73.
1734 51. Totura, A.L., Whitmore, A., Agnihothram, S., Schäfer, A., Katze, M.G., Heise,
1735 M.T. and Baric, R.S., 2015. Toll-like receptor 3 signaling via TRIF
1736 contributes to a protective innate immune response to severe acute
1737 respiratory syndrome coronavirus infection. MBio, 6(3), pp.e00638-15.
1738 52. Shaw, M.L., Cardenas, W.B., Zamarin, D., Palese, P. and Basler, C.F.,
1739 2005. Nuclear localization of the Nipah virus W protein allows for inhibition
1740 of both virus-and toll-like receptor 3-triggered signaling pathways. Journal
1741 of virology, 79(10), pp.6078-6088.
1742 53. Seto, J., Qiao, L., Guenzel, C.A., Xiao, S., Shaw, M.L., Hayot, F. and
1743 Sealfon, S.C., 2010. Novel Nipah virus immune-antagonism strategy
1744 revealed by experimental and computational study. Journal of virology,
1745 84(21), pp.10965-10973.
1746 54. Farina, C., Krumbholz, M., Giese, T., Hartmann, G., Aloisi, F. and Meinl, E.,
1747 2005. Preferential expression and function of Toll-like receptor 3 in human
1748 astrocytes. Journal of neuroimmunology, 159(1-2), pp.12-19.
1749 55. Weingartl, H., Czub, S., Copps, J., Berhane, Y., Middleton, D., Marszal, P.,
1750 Gren, J., Smith, G., Ganske, S., Manning, L. and Czub, M., 2005. Invasion
1751 of the central nervous system in a porcine host by Nipah virus. Journal of
1752 virology, 79(12), pp.7528-7534.
1753 56. Arnold K, Bordoli L, Kopp J, and Schwede T (2006). The SWISS-MODEL
1754 Workspace: A web-based environment for protein structure homology
1755 modelling. Bioinformatics.,22,195-201.
1756 57. Tenzer S, Peters B, Bulik S, Schoor O, Lemmel C, Schatz MM, Kloetzel PM,
1757 Rammensee HG, Schild H, Holzhutter HG. 2005. Modeling the MHC class
1758 I pathway by combining predictions of proteasomal cleavage, TAP
1759 transport and MHC class I binding. Cell Mol Life Sci 62:1025-1037.
1760 58. Peters B, Bulik S, Tampe R, Van Endert PM, Holzhutter HG. 2003. Identifying
1761 MHC class I epitopes by predicting the TAP transport efficiency of epitope
1762 precursors. J Immunol171:1741-1749.
1763 59. Hoof, I., Peters, B., Sidney, J., Pedersen, L.E., Sette, A., Lund, O., Buus, S.
1764 and Nielsen, M., 2009. NetMHCpan, a method for MHC class I binding
1765 prediction beyond humans. Immunogenetics, 61(1), p.1.
1766 60. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A,
1767 Kesmir C, Peters B. 2013. Properties of MHC class I presented peptides
1768 that enhance immunogenicity. PloS Comp. Biol. 8(1):361.
1769 61. Wang, P., Sidney, J., Kim, Y., Sette, A., Lund, O., Nielsen, M. and Peters,
1770 B., 2010. Peptide binding predictions for HLA DR, DP and DQ molecules.
1771 BMC bioinformatics, 11(1), p.568.
1772 62. Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A. and
1773 Peters, B., 2008. Quantitative peptide binding motifs for 19 human and
1774 mouse MHC class I molecules derived using positional scanning
1775 combinatorial peptide libraries. Immunome research, 4(1), p.2.
1776 63. Nielsen, M., Lundegaard, C. and Lund, O., 2007. Prediction of MHC class II
1777 binding affinity using SMM-align, a novel stabilization matrix alignment

1778 method. BMC bioinformatics, 8(1), p.238.
1779 64. Sturniolo, T., Bono, E., Ding, J., Raddrizzani, L., Tuereci, O., Sahin, U.,
1780 Braxenthaler, M., Gallazzi, F., Protti, M.P., Sinigaglia, F. and Hammer, J.,
1781 1999. Generation of tissue-specific and promiscuous HLA ligand
1782 databases using DNA microarrays and virtual HLA class II matrices.
1783 Nature biotechnology, 17(6), p.555.
1784 65. Bui H. H, Sidney J, Dinh K, Southwood S, Newman M. J, Sette A. 2006.
1785 Predicting population coverage of T-cell epitope-based diagnostics and
1786 vaccines. BMC Bioinformatics 17:153.
1787 66. Larsen JE, Lund O, Nielsen M. 2006. Improved method for predicting linear
1788 B-cell epitopes. Immunome Res 2:2.
1789 67. Chou PY, Fasman GD. 1978. Prediction of the secondary structure of
1790 proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol
1791 Biol 47:45-148.
1792 68. Emini EA, Hughes JV, Perlow DS, Boger J. 1985. Induction of hepatitis A
1793 virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol
1794 55:836-839.
1795 69. Karplus PA, Schulz GE. 1985. Prediction of chain flexibility in proteins.
1796 Naturwissenschaften 72:212-213.
1797 70. Kolaskar AS, Tongaonkar PC. 1990. A semi-empirical method for prediction
1798 of antigenic determinants on protein antigens. FEBS Lett276:172-174.
1799 71. Parker JM, Guo D, Hodges RS. 1986. New hydrophilicity scale derived from
1800 high-performance liquid chromatography peptide retention data:
1801 correlation of predicted surface residues with antigenicity and X-ray-
1802 derived accessible sites. Biochemistry 25:5425-5432.
1803 72. J. V. Kringelum, C. Lundegaard, O. Lund, M. Nielsen. 2012. Reliable B cell
1804 epitope predictions: impacts of method development and improved
1805 benchmarking. PLoS Comput Biol. 8:(12):e1002829.
1806 73. Ponomarenko JV, Bui H, Li W, Fusseder N, Bourne PE, Sette A, Peters B.
1807 2008. ElliPro: a new structure-based tool for the prediction of antibody
1808 epitopes. BMC Bioinformatics 9:514.
1809 74. Bui HH, Sidney J, Li W, Fusseder N, Sette A. 2007. Development of an
1810 epitope conservancy analysis tool to facilitate the design of epitope-based
1811 diagnostics and vaccines. BMC Bioinformatics 8:361.
1812 75. Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava,
1813 G.P. and Open Source Drug Discovery Consortium, 2013. In silico
1814 approach for predicting toxicity of peptides and proteins. PLoS One, 8(9),
1815 p.e73957.
1816 76. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R.,
1817 McWilliam, H., Remmert, M., Söding, J. and Thompson, J.D., and Higgins
1818 D.G., 2011. Fast, scalable generation of high-quality protein multiple
1819 sequence alignments using Clustal Omega. Molecular systems biology,
1820 7(1), p.539.
1821 77. Benkert, P., Tosatto, S.C. and Schomburg, D., 2008. QMEAN: A
1822 comprehensive scoring function for model quality assessment. Proteins:
1823 Structure, Function, and Bioinformatics, 71(1), pp.261-277.

78. Singh, S., Singh, H., Tuknait, A., Chaudhary, K., Singh, B., Kumaran, S. and Raghava, G.P.S. (2015) PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. Biology Direct 10:73.
79. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. and Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 30(16), pp.2785-2791.
80. O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, Journal of Computational Chemistry 31 (2010) 455-461
81. Abraham, M. J., Murtola, T., Schulz, R., Pa´ll, S., Smith, J. C., Hess, B., Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from lap- tops to supercomputers. SoftwareX 1–2:19–25, 2015.
82. Jorgensen, W.L., Maxwell, D.S. and Tirado-Rives, J., 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. Journal of the American Chemical Society, 118(45), pp.11225-11236.
83. Abele, R. and Tampé, R., 2004. The ABCs of immunology: structure and function of TAP, the transporter associated with antigen processing. Physiology, 19(4), pp.216-224.
84. Hu W, Li F, Yang X, Li Z, Xia H, Li G, Wang Y, Zhang Z. A flexible peptide linker enhances the immunoreactivity of two copies HBsAg preS1 (21-47) fusion protein. J Biotechnol. 2004;107:83–90.
85. Hajighahramani, N., Nezafat, N., Eslami, M., Negahdaripour, M., Rahmatabadi, S.S. and Ghasemi, Y., 2017. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infection, Genetics and Evolution, 48, pp.83-94.
86. Chen, X., Zaro, J.L. and Shen, W.C., 2013. Fusion protein linkers: property, design and functionality. Advanced drug delivery reviews, 65(10), pp.1357-1369.
87. Hoover, D.M., Rajashankar, K.R., Blumenthal, R., Puri, A., Oppenheim, J.J., Chertov, O. and Lubkowski, J., 2000. The structure of human β-defensin-2 shows evidence of higher order oligomerization. Journal of Biological Chemistry, 275(42), pp.32911-32918. PDB ID: 1FD3.
88. Nagpal, G., Gupta, S., Chaudhary, K., Dhanda, S.K., Prakash, S. and Raghava, G.P., 2015. VaccineDA: Prediction, design and genome-wide screening of oligodeoxynucleotide-based vaccine adjuvants. Scientific reports, 5, p.12478.
89. Dhanda, S. K., Vir, P. & Raghava, G. P. Designing of interferon-gamma inducing MHC class-II binders. Biol. Direct. 8, 30 (2013).
90. Saha, S. & Raghava, G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic. Acids. Res. 34, W202–W209 (2006).
91. Irini A Doytchinova and Darren R Flower. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC

Bioinformatics. 2007 8:4.

92. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.E., Wilkins, M.R., Appel, R.D. and Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy server (pp. 571-607). Humana Press.

93. Morten Källberg, Haipeng Wang, Sheng Wang, Jian Peng, Zhiyong Wang, Hui Lu, and Jinbo Xu. Template-based protein structure modeling using the RaptorX web server. Nature Protocols 7, 1511-1522, 2012.

94. Ma, J., Wang, S., Zhao, F. and Xu, J., 2013. Protein threading using context-specific alignment potential. Bioinformatics, 29(13), pp.i257-i265.

95. Wang, Z., Zhao, F., Peng, J. and Xu, J., 2010, December. Protein 8-class secondary structure prediction using conditional neural fields. In 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 109-114). IEEE.

96. Wang, Z. and Xu, J., 2013. Predicting protein contact map using evolutionary and physical constraints by integer programming. Bioinformatics, 29(13), pp.i266-i273.

97. Dong Xu and Yang Zhang. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-step Atomic-level Energy Minimization. Biophysical Journal, vol 101, 2525-2534 (2011).

98. J. Ko, H. Park, L. Heo, and C. Seok, GalaxyWEB server for protein structure prediction and refinement, Nucleic Acids Res. 40 (W1), W294-W297 (2012).

99. Shin, W.H., Lee, G.R., Heo, L., Lee, H. and Seok, C., 2014. Prediction of protein structure and interaction by GALAXY protein modeling programs. Bio Design, 2(1), pp.1-11.

100. Ramakrishnan, C. and Ramachandran, G.N., 1965. Stereochemical criteria for polypeptide and protein chain conformations: II. Allowed conformations for a pair of peptide units. Biophysical journal, 5(6), pp.909-933.

101. S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson and D.C. Richardson (2002) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins: Structure, Function & Genetics. 50: 437-450.

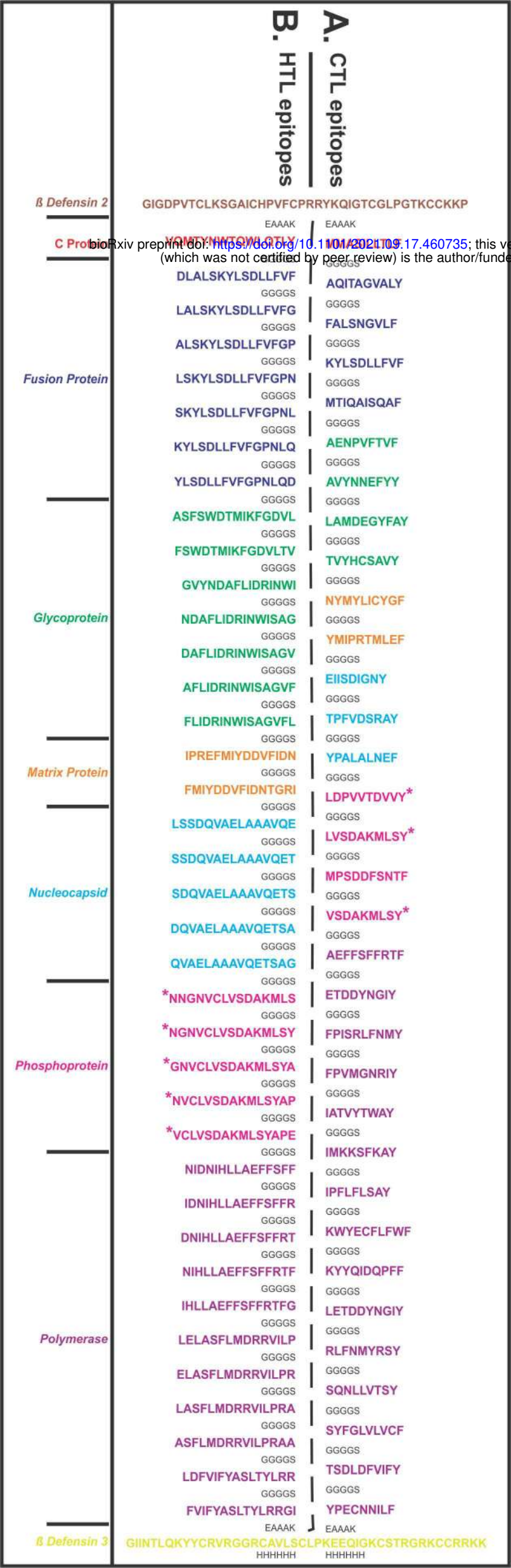
102. Bell, J.K., Botos, I., Hall, P.R., Askins, J., Shiloach, J., Segal, D.M. and Davies, D.R., 2005. The molecular structure of the Toll-like receptor 3 ligand-binding domain. Proceedings of the National Academy of Sciences of the United States of America, 102(31), pp.10976-10980.

103. Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. In Gusfield et al., Ed. Proceedings of the 2'nd Workshop on Algorithms in Bioinformatics(WABI) Rome, Italy, Lecture Notes in Computer Science 2452, pp. 185-200, Springer Verlag, 2002

104. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids. Res. 33: W363-367, 2005.

105. Nezafat, N., Eslami, M., Negahdaripour, M., Rahbar, M.R. and Ghasemi, Y., 2017. Designing an efficient multi-epitope oral vaccine against Helicobacter pylori using immunoinformatics and structural vaccinology

1916 approaches. Molecular BioSystems, 13(4), pp.699-713.
1917 106. Morla, S., Makhija, A. & Kumar, S. Synonymous codon usage pattern in
1918 glycoprotein gene of rabies virus. Gene. 584, 1–6 (2016).
1919 107. Wu, X., Wu, S., Li, D., Zhang, J., Hou, L., Ma, J., Liu, W., Ren, D., Zhu, Y.
1920 and He, F., 2010. Computational identification of rare codons of
1921 Escherichia coli based on codon pairs preference. Bmc Bioinformatics,
1922 11(1), p.61.
1923
1924



Overlapping regions of predicted B cell Bepipred linear Epitope and epitopes predicted by other protein sequence and structure based methods		Overlapping regions of sequence based B cell linear epitope prediction					Overlapping regions of structure based B cell epitope prediction		
		Chou & Fasman Beta-Turn	Emini Surface Accessibility	Karplus & Schulz Flexibility	Kolaskar & Tongaonkar Antigenicity	Parker Hydrophilicity	DiscoTope: discontinuous epitopes	ElliPro	
								Linear	Discontinuous
Fusion Protein	215 GPNLQDPVSNSM 226	215-226		215-226	215-216	217-226			
Glycoprotein	271 WTPPNPNT 278	271-278	271-278	272-278		272-278			271-278
Glycoprotein	459 SWDTMI 464	459-460							463-464
Glycoprotein	529 NQTAE 533	529-532	529-533	529-533		529-533			
Matrix Protein	184 SGIYM 188	184-188		184-188		184-188			
Matrix Protein	330 SIPREFMIY 338	330-335		330-336		331-334	332-333	332-333	333-338
Matrix Protein	340 DVFIDNTGRI 349	341-349		340-349		340-344			340-349
Phosphoprotein	125 GYGFTSSPERGWSDYTSGA 143	125-143	130-136	128-143		127-143			
Phosphoprotein	163 IAVSKEDR 170	166-170	166-170	163-170	163-166	165-170			
Polymerase	346 NIDN 349	348-349							
Polymerase	364 GHPILE 369	364-367			365-369				366-369
Polymerase	1095 DKSFDEDLEL 1104	1096-1103	1096-1101	1095-1104		1096-1103			1095-1104
Polymerase	1108 LMDR 1111								1108-1111
Polymerase	1409 LRLETDDYNG 1418	1412-1418	1410-1416	1412-1418		1411-1418	1409-1417	1409-1417	
Polymerase	1982 GFPIS 1986	1982-1986		1982-1986	1982-1986				1182-1986
Polymerase	2004 PVYSNPD 2010	2004-2010	2006-2010	2007-2010	2004-2008	2007-2010			2004-2010

Clusters of overlapping CTL, HTL & B cell epitopes

Fusion Protein

```
                                  ----------MTIQAISQAF-
                        ------KYLSDLLFVF----------
                         -------YLSDLLFVFGPNLQD----------
                         ------KYLSDLLFVFGPNLQ----------
                          -----SKYLSDLLFVFGPNL----------
                           ----LSKYLSDLLFVFGPN----------
                            ---ALSKYLSDLLFVFGP----------
                             --LALSKYLSDLLFVFG----------
                              -DLALSKYLSDLLFVF----------
                                  ----------GPNLQDPVSNSM----------
199-236 LDLALSKYLSDLLFVFGPNLQDPVSNSMTIQAISQAFG
```

Glycoprotein

```
                                                                                  ----------AENPVFTVF-
                                                                          -------FLIDRINWISAGVFL----------
                                                                           ------AFLIDRINWISAGVF----------
                                                                            -----DAFLIDRINWISAGV----------
                                                                             ----NDAFLIDRINWISAG----------
                                                                              -GVYNDAFLIDRINWI----------
                                                                                  ----------NQTAE--------
                  --------TVYHCSAVY-------                  ----FSWDTMIKFGDVLTV--
                   ----------AVYNNEFYY-                    --ASFSWDTMIKFGDVL----
                    -WTPPNPNT----------                      -----SWDTMI----------
270-293 VWTPPNPNTVYHCSAVYNNEFYYV     454-474 YQASFSWDTMIKFGDVLTVNP     505-541 EGVYNDAFLIDRINWISAGVFLDSNQTAENPVFTVFK
```

Matrix Protein

```
                                  -----FMIYDDVFIDNTGRI-
                                 -IPREFMIYDDVFIDN-----
                                  ----------DVFIDNTGRI-
                   ----YMIPRTMLEF-
                    -SGIYM---------
183-197   DSGIYMIPRTMLEFR          330-350 SIPREFMIYDDVFIDNTGRIL
```

Nucleocapsid

```
                    ------QVAELAAAVQETSAG--
                     -----DQVAELAAAVQETSA---
                      ----SDQVAELAAAVQETS----
                       ---SSDQVAELAAAVQET-----
                        --LSSDQVAELAAAVQE------
382-404 LGLSSDQVAELAAAVQETSAGRQ
```

Phosphoprotein

```
                                  ----------LVSDAKMLSY----------
                                  ----------VSDAKMLSY----------
                                  ----------VCLVSDAKMLSYAPE----------
                                  ----------NVCLVSDAKMLSYAP----------
                                  ----------GNVCLVSDAKMLSYA----------
                                  ----------NGNVCLVSDAKMLSY----------
                                  ----------NNGNVCLVSDAKMLS----------
                                  ----------IAVSKEDR--
                  --GYGFTSSPERGWSDYTSGA----------
123-172 CTGYGFTSSPERGWSDYTSGANNGNVCLVSDAKMLSYAPEIAVSKEDRET
```

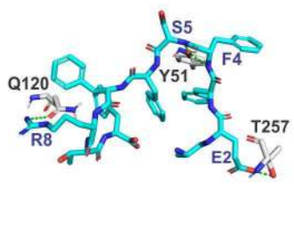
Polymerase

```
                    ---------AEFFSFFRTF-------
                     -----IHLLAEFFSFFRTFG------
                      ----NIHLLAEFFSFFRTF-------
                       ---DNIHLLAEFFSFFRT--------
                        --IDNIHLLAEFFSFFR---------
                         -NIDNIHLLAEFFSFF----------
                          -NIDN----------
                                  ----------GHPILE-
345-370 LNIDNIHLLAEFFSFFRTFGHPILEA
                                                    ----------ASFLMDRRVILPRAA-----
                                                    ----------LASFLMDRRVILPRA------
                                                    ----------ELASFLMDRRVILPR-------
                                                    ---------LELASFLMDRRVILP--------
                                                    ----------LMDR----------
                                                    --DKSFDEDLEL----------
1093-1124 FHDKSFDEDLELASFLMDRRVILPRAAHEILD
                                                                                  ----ETDDYNGIY--
                                                                                  ---LETDDYNGIY--
                                                                                  -LRLETDDYNG----
1408-1422 RLRLETDDYNGIYHL
```

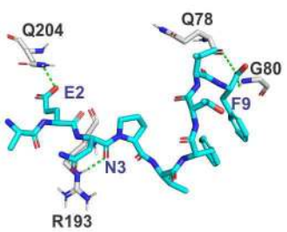
Polymerase

```
                                                    ----------SYFGLVLVCF---------
                                                    ------RLFNMYRSY----------
                                                    --FPISRLFNMY----------
                                                    ----------PVYSNPD--
                   -TSDLDFVIFY----------
                     ------FVIFYASLTYLRRGI-
                      ----LDFVIFYASLTYLRR---
                                                    -GFPIS----------
1715-1736 NTSDLDFVIFYASLTYLRRGII     1981-2012 PGFPISRLFNMYRSYFGLVLVCFPVYSNPDST
```

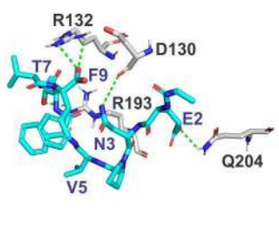
A. CTL epitopes - HLA Class I allele complexes



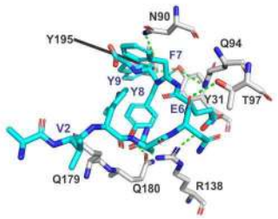
1. AEFFSFFRTF-B4403



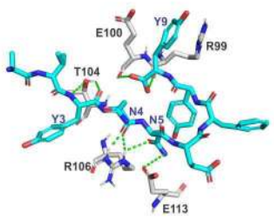
2. AENPVFTVF-B4402



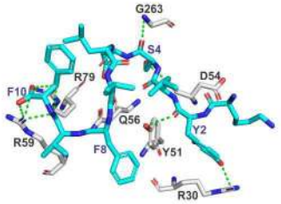
3. AENPVFTVF-B4403



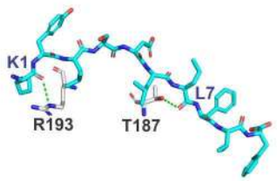
4. AVYNNEFYY-A1101



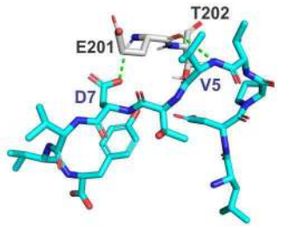
5. AVYNNEFYY-A3002



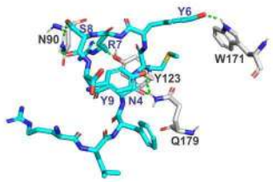
6. KYLSDLLFVF-A2301



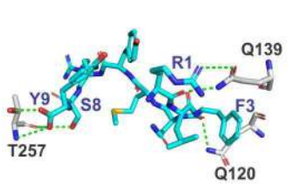
7. KYLSDLLFVF-A2402



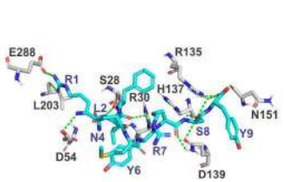
8. LDPVVTDVVY-B3501



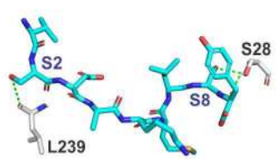
9. RLFNMYRSY-A3002



10. RLFNMYRSY-A3201

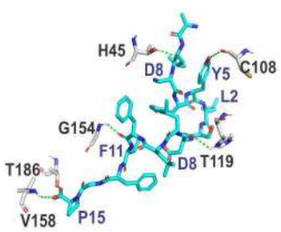


11. RLFNMYRSY-B1501

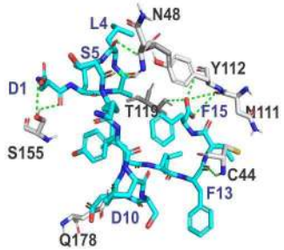


12. VSDAKMLSY-A0101

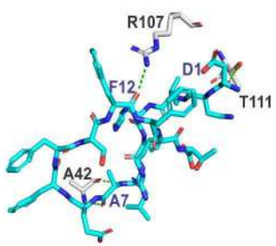
B. HTL epitopes - HLA Class II allele complexes



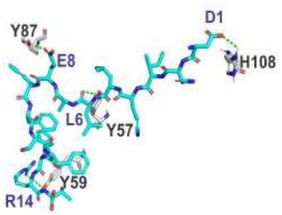
1. ALSKYLSDLLFVFGP-DRB3-0101



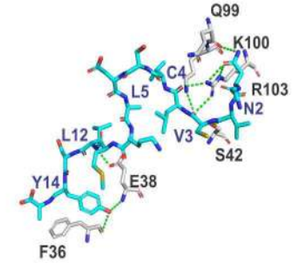
2. DLALSKYLSDLLFVF-DRB3-0101



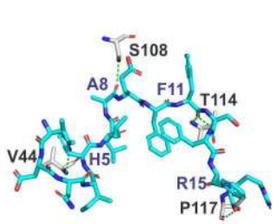
3. DNIHLLAEFFSFFRT-DPA1-0103



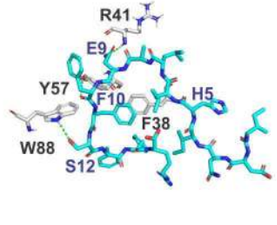
4. DNIHLLAEFFSFFRT-DPB1-0201



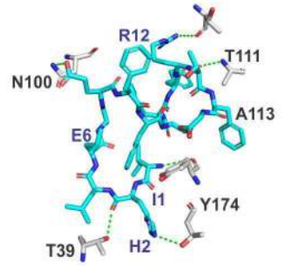
5. GNVCLVSDAKMLSYA-DRB1-0301



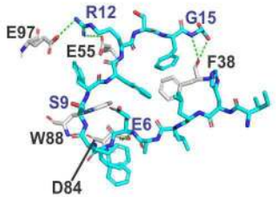
6. IDNIHLLAEFFSFFR-DPA1-0103



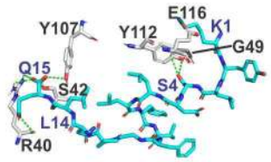
7. IDNIHLLAEFFSFFR-DPB1-0201



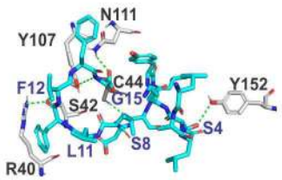
8.IHLLAEFFSFFRTFG-DPA1-0103



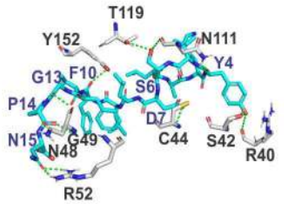
9. IHLLAEFFSFFRTFG-DPB1-0201



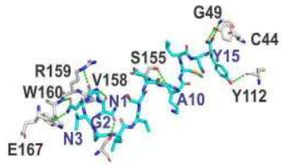
10. KYLSDLLFVFGPNLQ-DRB3-0101



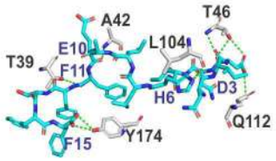
11. LALSKYLSDLLFVFG-DRB3-0101



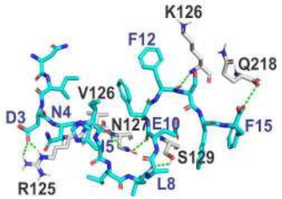
12. LSKYLSDLLFVFGPN-DRB3-0101



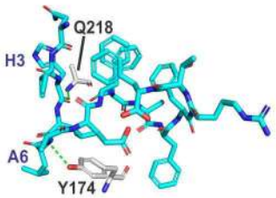
13. NGNVCLVSDAKMLSY-DRB1-0301



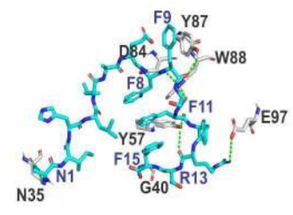
14. NIDNIHLLAEFFSFF-DPA1-0103



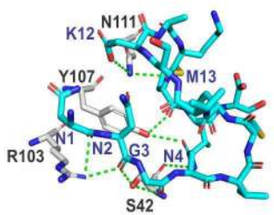
15. NIDNIHLLAEFFSFF-DPB1-0201



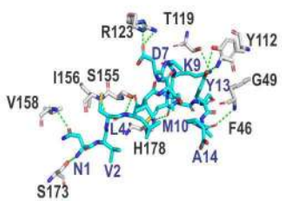
16. NIHLLAEFFSFFRTF-DPA1-0103



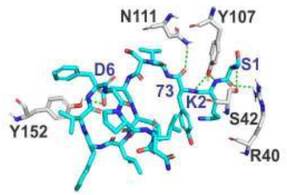
17. NIHLLAEFFSFFRTF-DPB1-0201



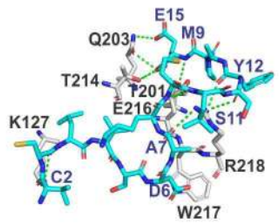
18. NNGNVCLVSDAKMLS-DRB1-0301



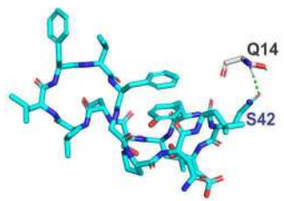
19.NVCLVSDAKMLSYAP-DRB1-0301



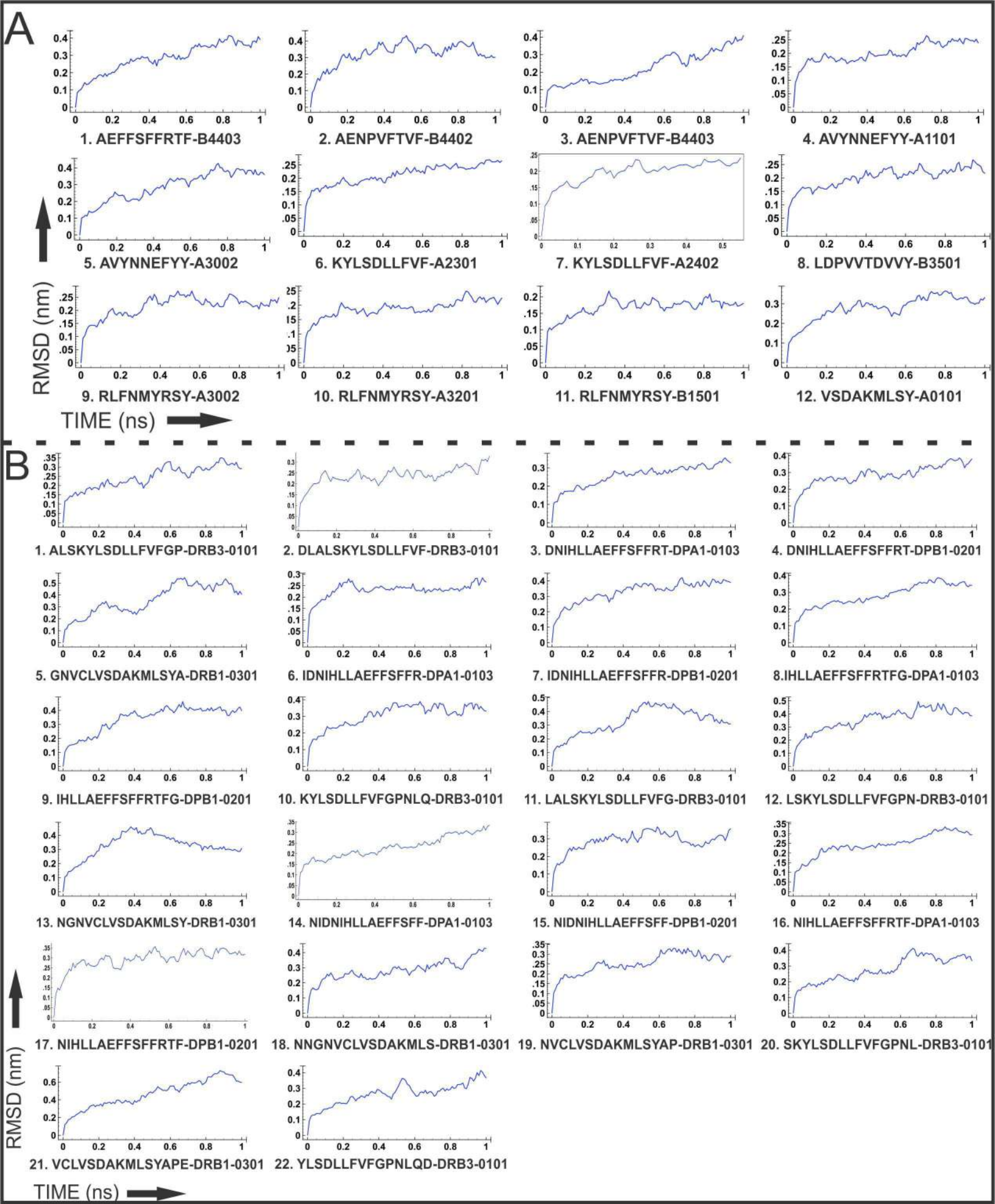
20.SKYLSDLLFVFGPNL-DRB3-0101

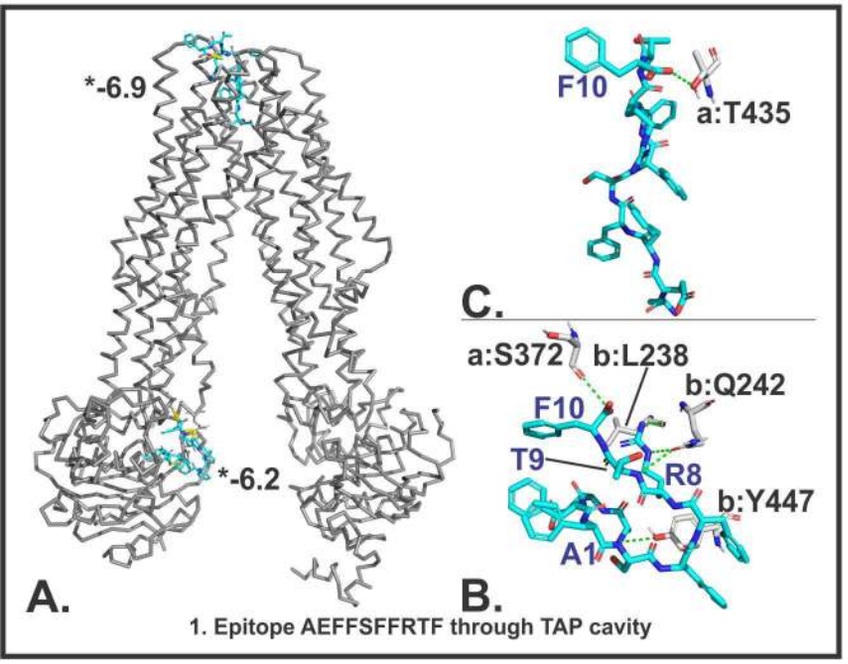


21. VCLVSDAKMLSYAPE-DRB1-0301

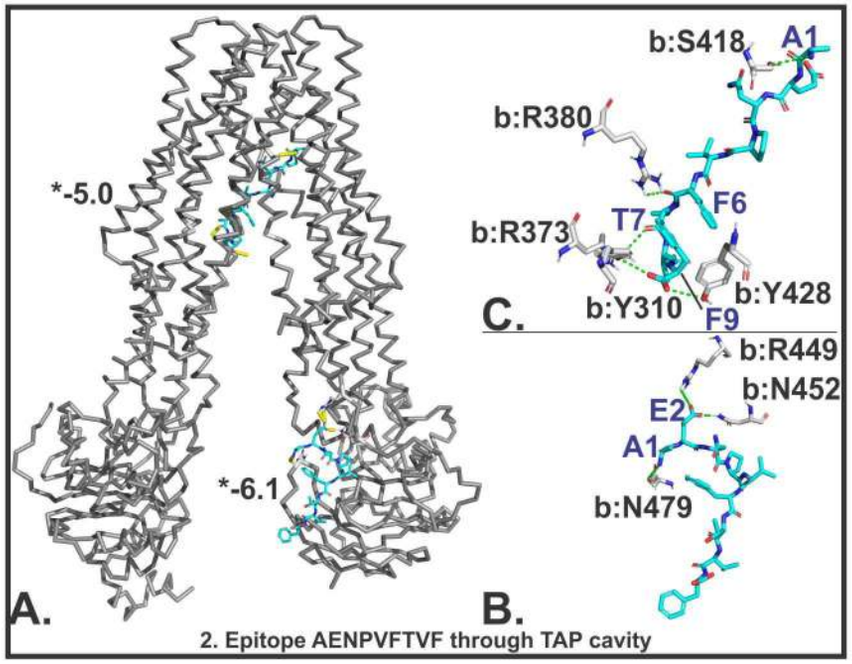


22. YLSDLLFVFGPNLQD-DRB3-0101

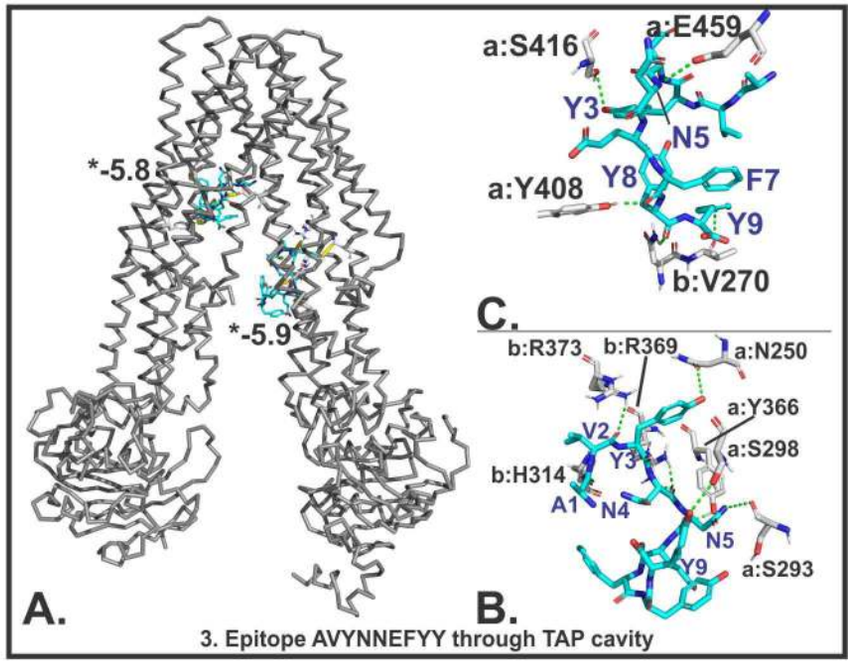

A
1. AEFFSFFRTF-B4403
2. AENPVFTVF-B4402
3. AENPVFTVF-B4403
4. AVYNNEFYY-A1101
5. AVYNNEFYY-A3002
6. KYLSDLLFVF-A2301
7. KYLSDLLFVF-A2402
8. LDPVVTDVVY-B3501
9. RLFNMYRSY-A3002
10. RLFNMYRSY-A3201
11. RLFNMYRSY-B1501
12. VSDAKMLSY-A0101
RMSD (nm)
TIME (ns)
B
1. ALSKYLSDLLFVFGP-DRB3-0101
2. DLALSKYLSDLLFVF-DRB3-0101
3. DNIHLLAEFFSFFRT-DPA1-0103
4. DNIHLLAEFFSFFRT-DPB1-0201
5. GNVCLVSDAKMLSYA-DRB1-0301
6. IDNIHLLAEFFSFFR-DPA1-0103
7. IDNIHLLAEFFSFFR-DPB1-0201
8.IHLLAEFFSFFRTFG-DPA1-0103
9. IHLLAEFFSFFRTFG-DPB1-0201
10. KYLSDLLFVFGPNLQ-DRB3-0101
11. LALSKYLSDLLFVFG-DRB3-0101
12. LSKYLSDLLFVFGPN-DRB3-0101
13. NGNVCLVSDAKMLSY-DRB1-0301
14. NIDNIHLLAEFFSFF-DPA1-0103
15. NIDNIHLLAEFFSFF-DPB1-0201
16. NIHLLAEFFSFFRTF-DPA1-0103
17. NIHLLAEFFSFFRTF-DPB1-0201
18. NNGNVCLVSDAKMLS-DRB1-0301
19. NVCLVSDAKMLSYAP-DRB1-0301
20. SKYLSDLLFVFGPNL-DRB3-0101
21. VCLVSDAKMLSYAPE-DRB1-0301
22. YLSDLLFVFGPNLQD-DRB3-0101
RMSD (nm)
TIME (ns)



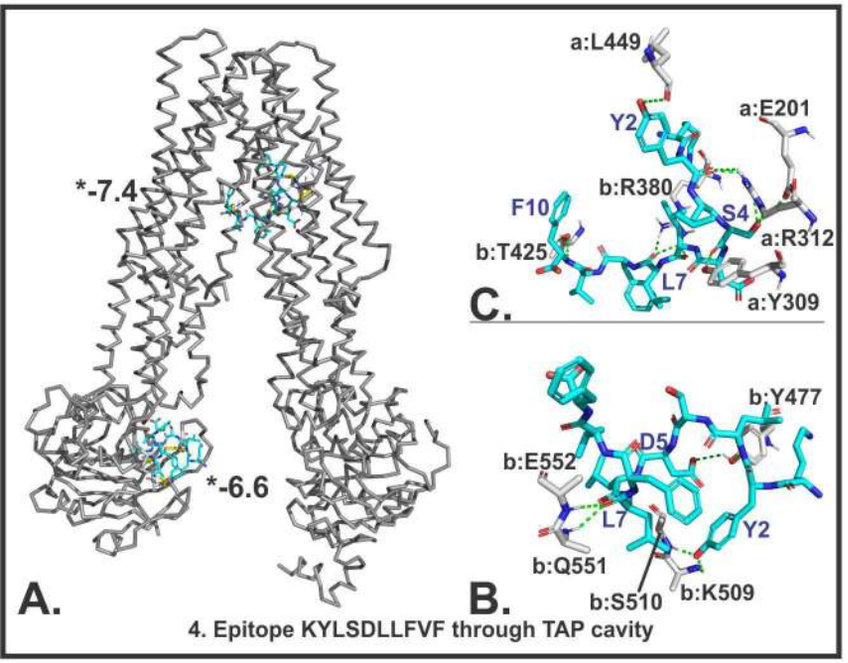
1. Epitope AEFFSFFRTF through TAP cavity



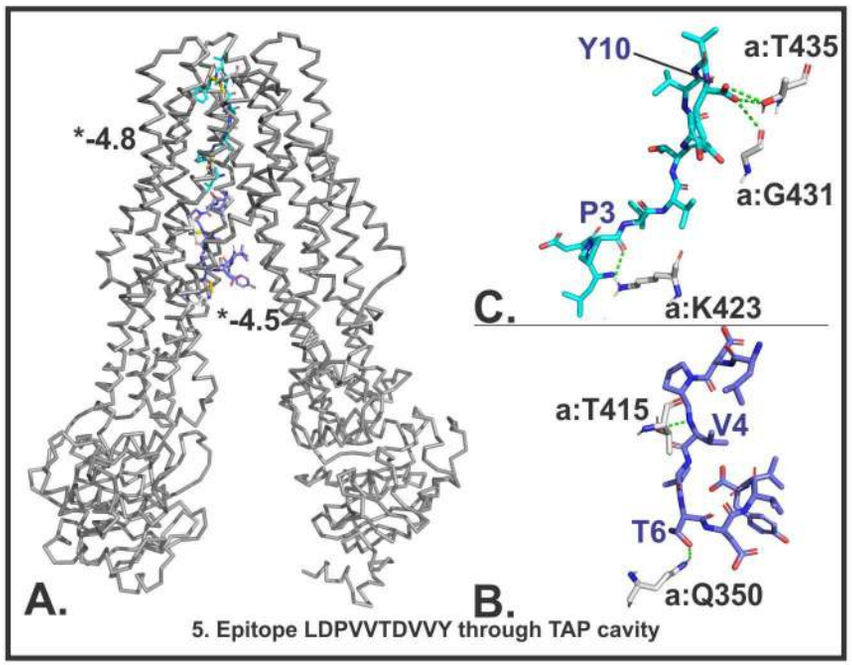
2. Epitope AENPVFTVF through TAP cavity



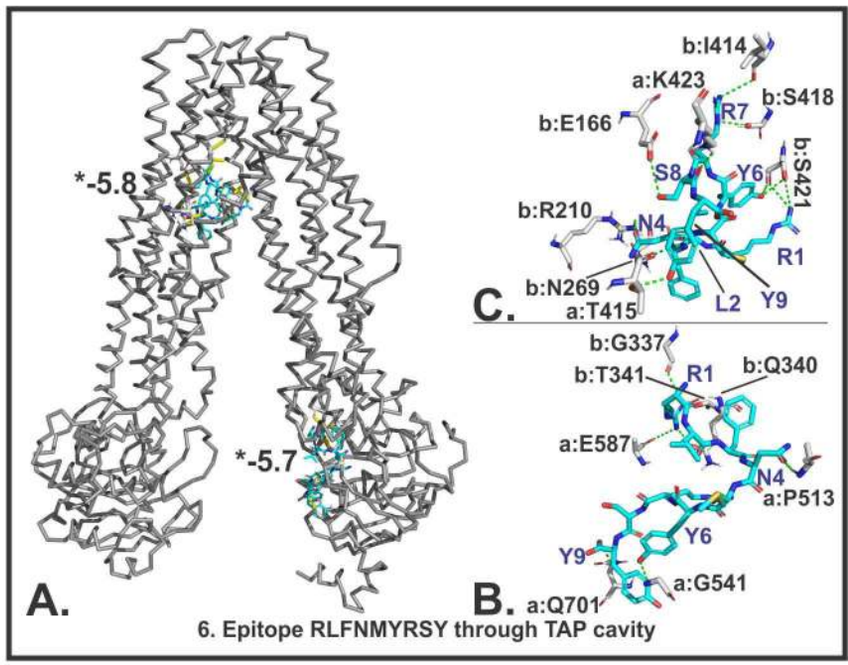
3. Epitope AVYNNEFYY through TAP cavity



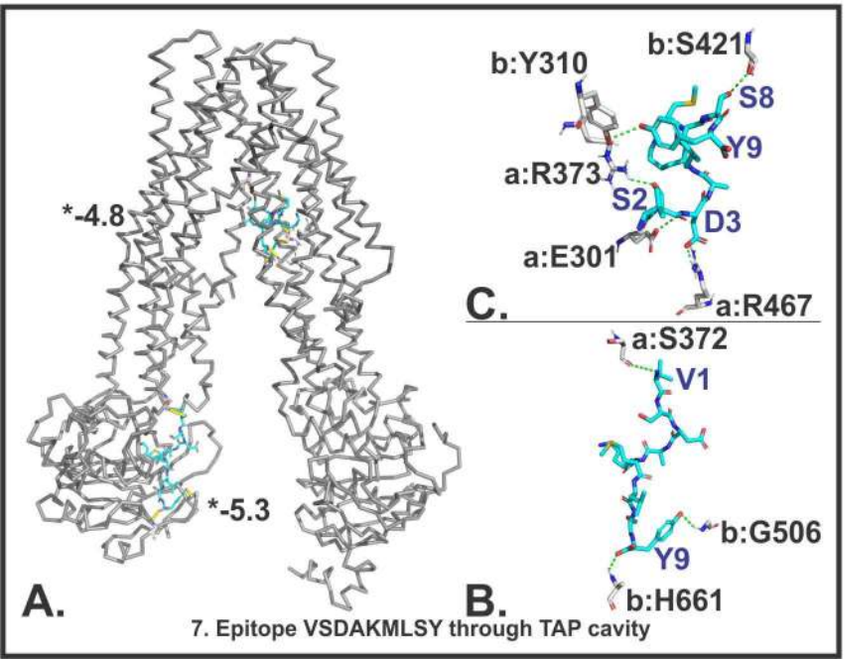
4. Epitope KYLSDLLFVF through TAP cavity



5. Epitope LDPVVTDVVY through TAP cavity

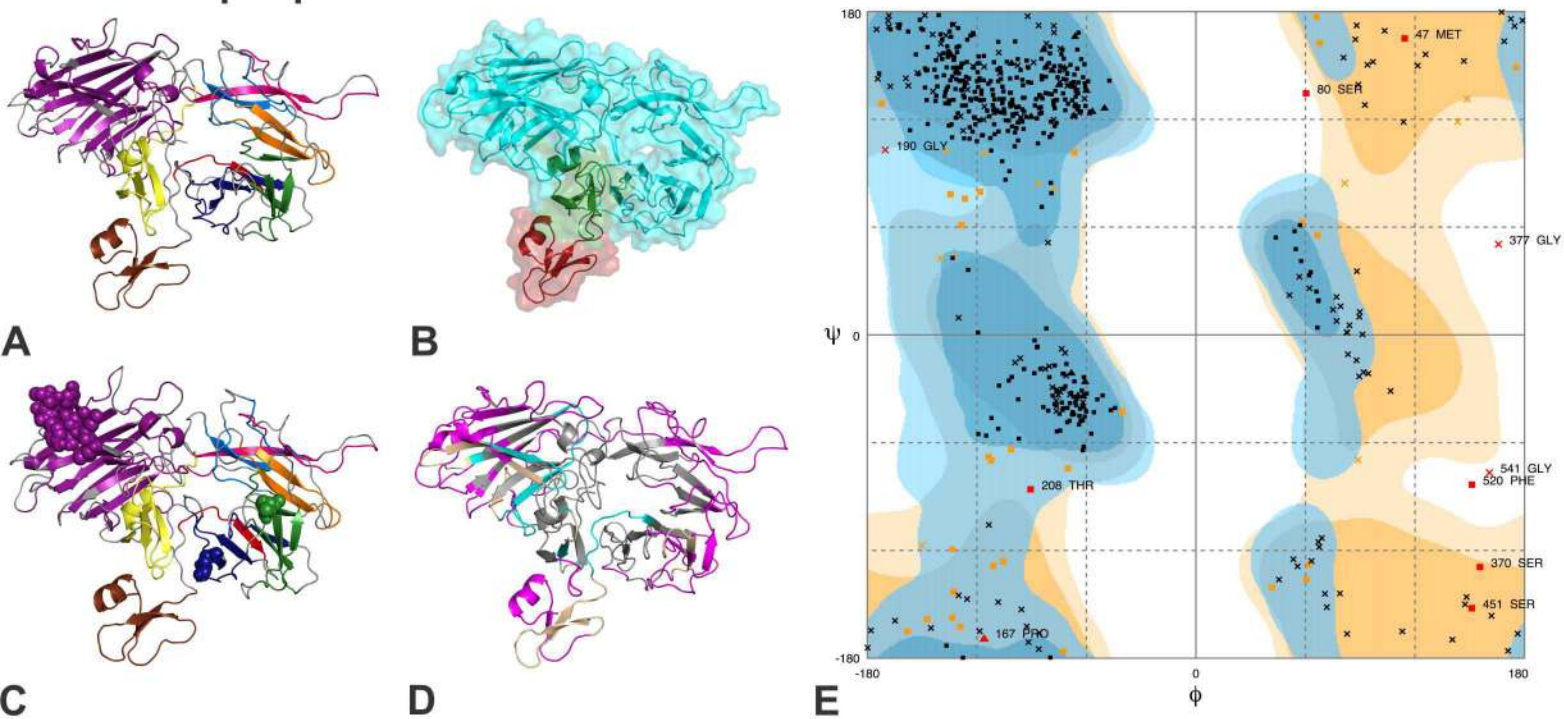


6. Epitope RLFNMYRSY through TAP cavity

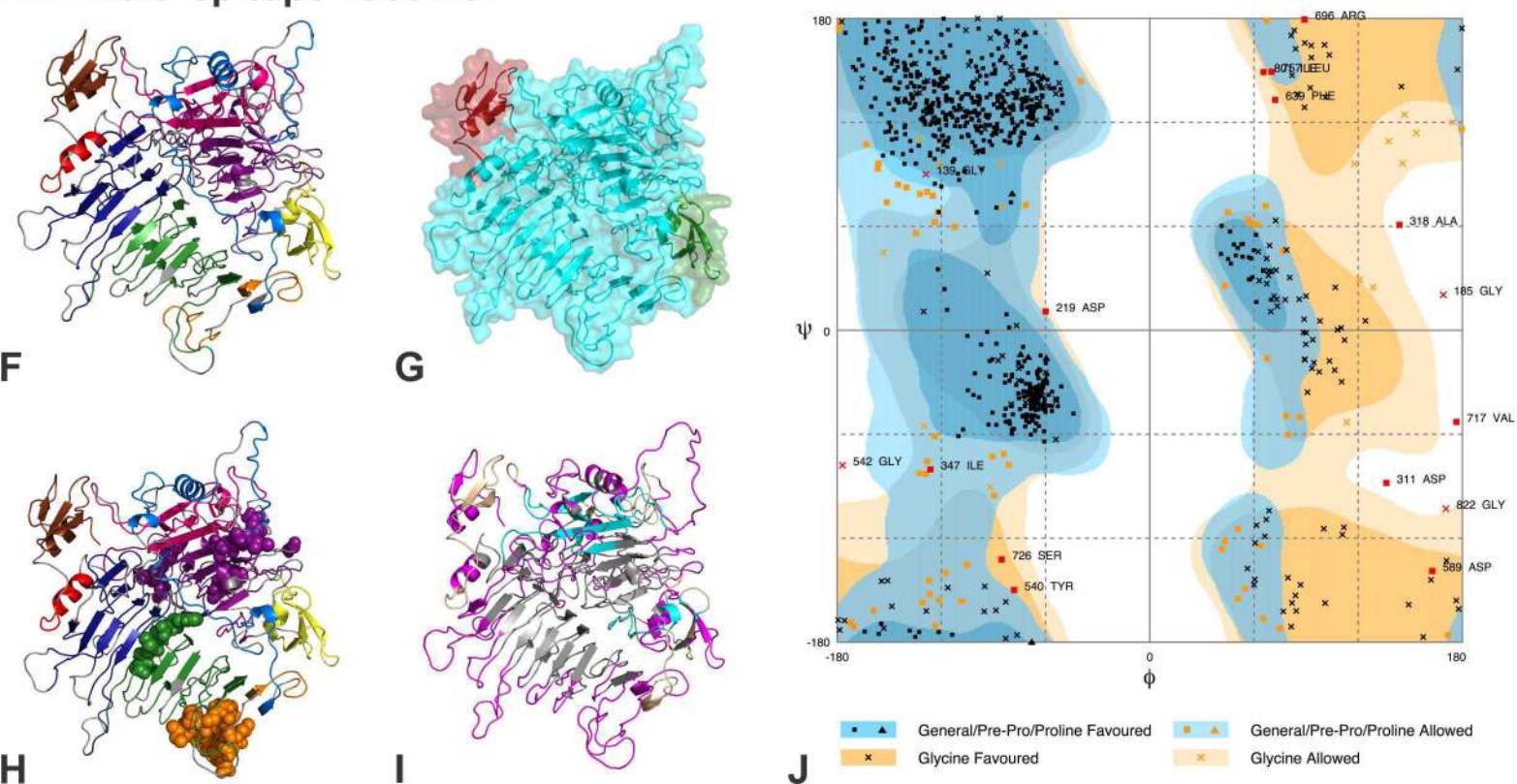


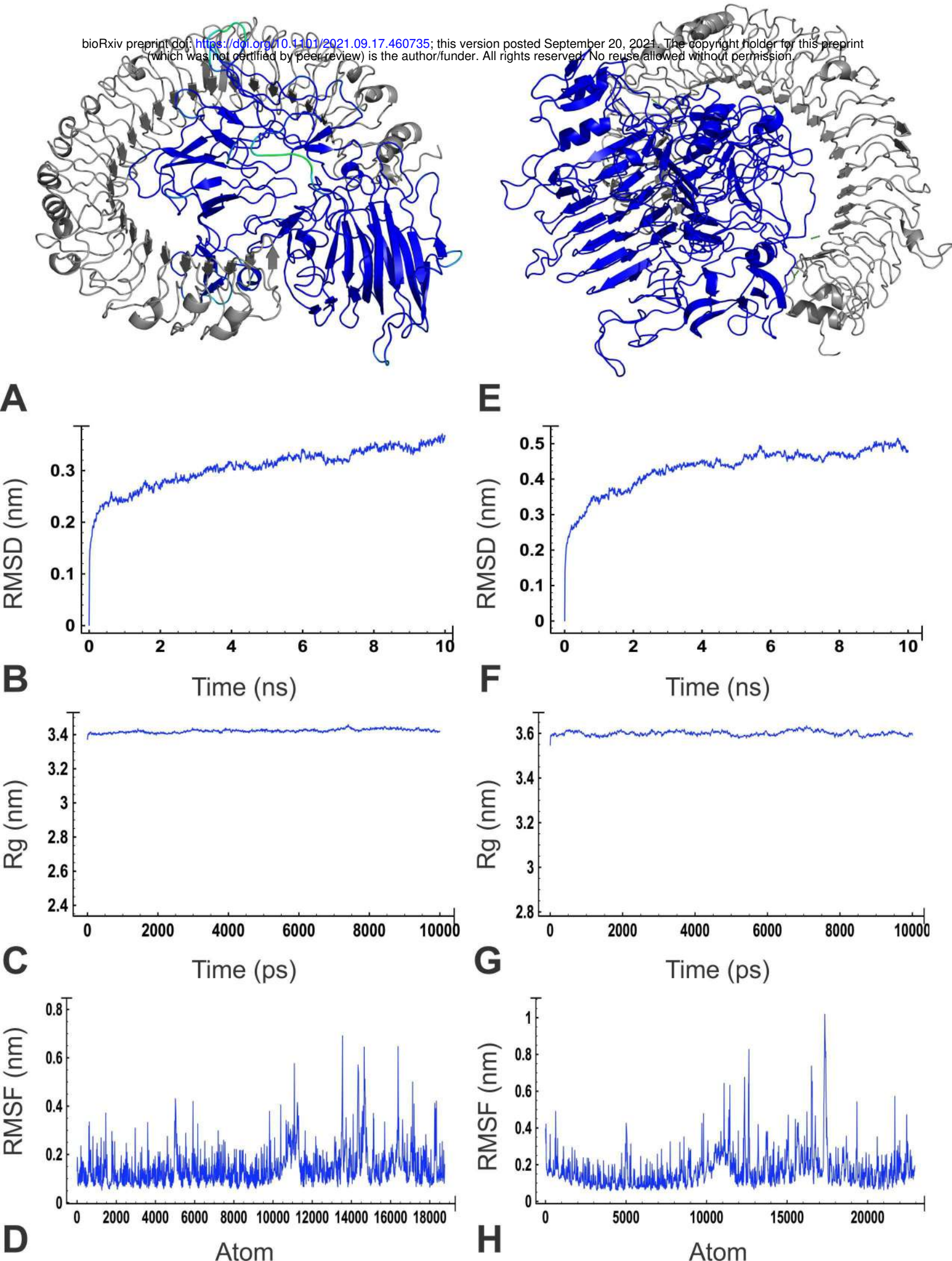
7. Epitope VSDAKMLSY through TAP cavity

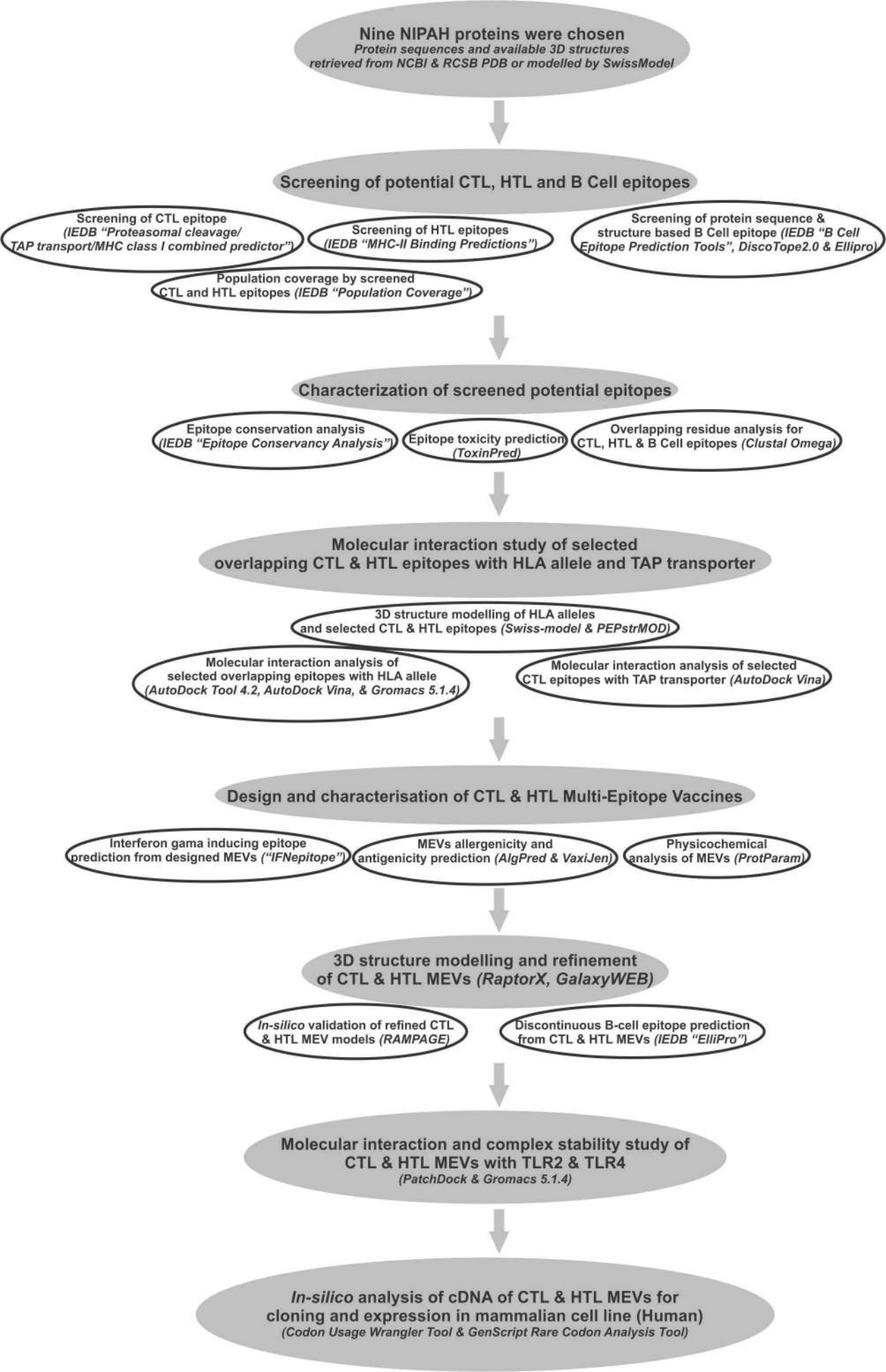
CTL multi-epitope vaccine:

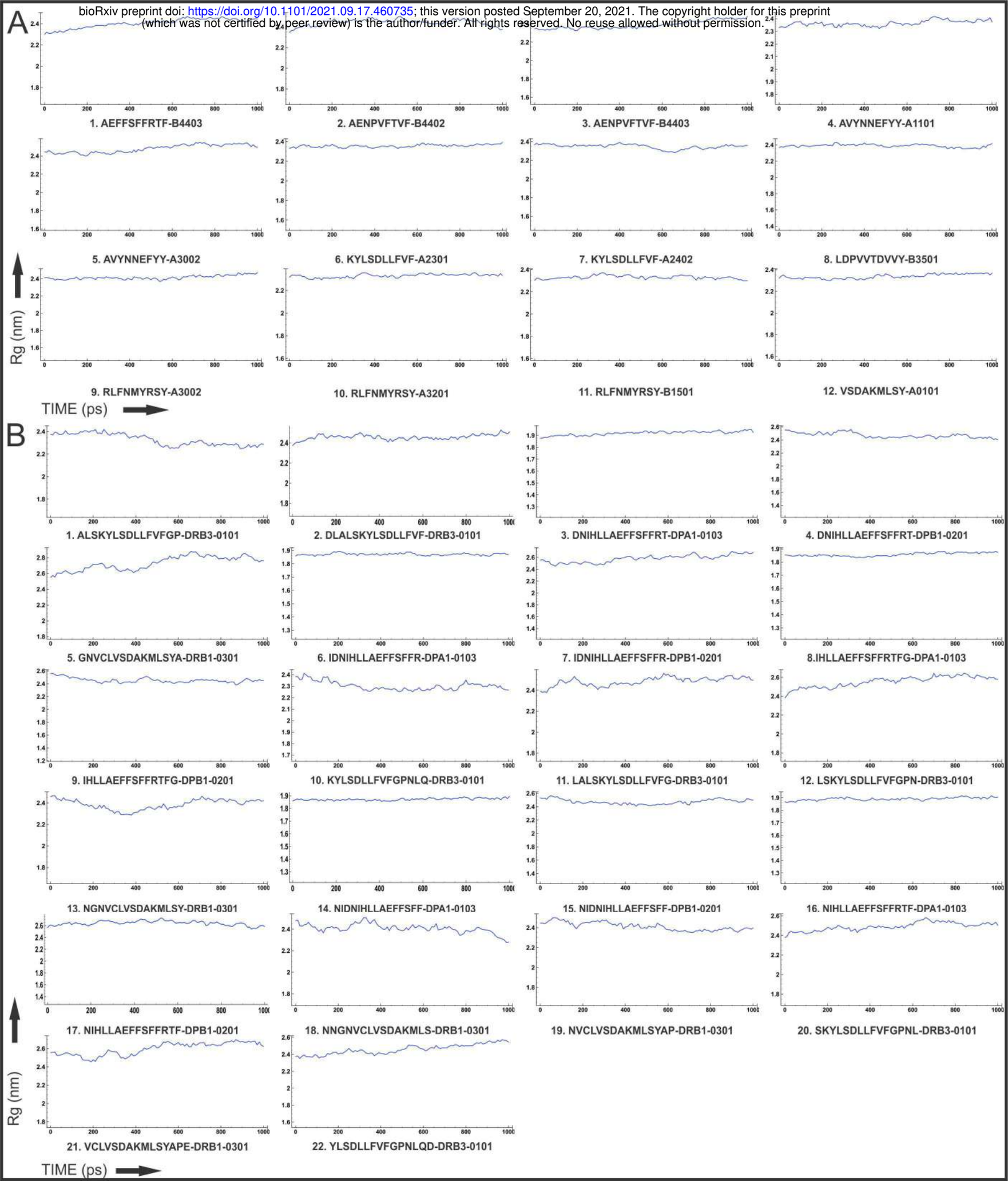


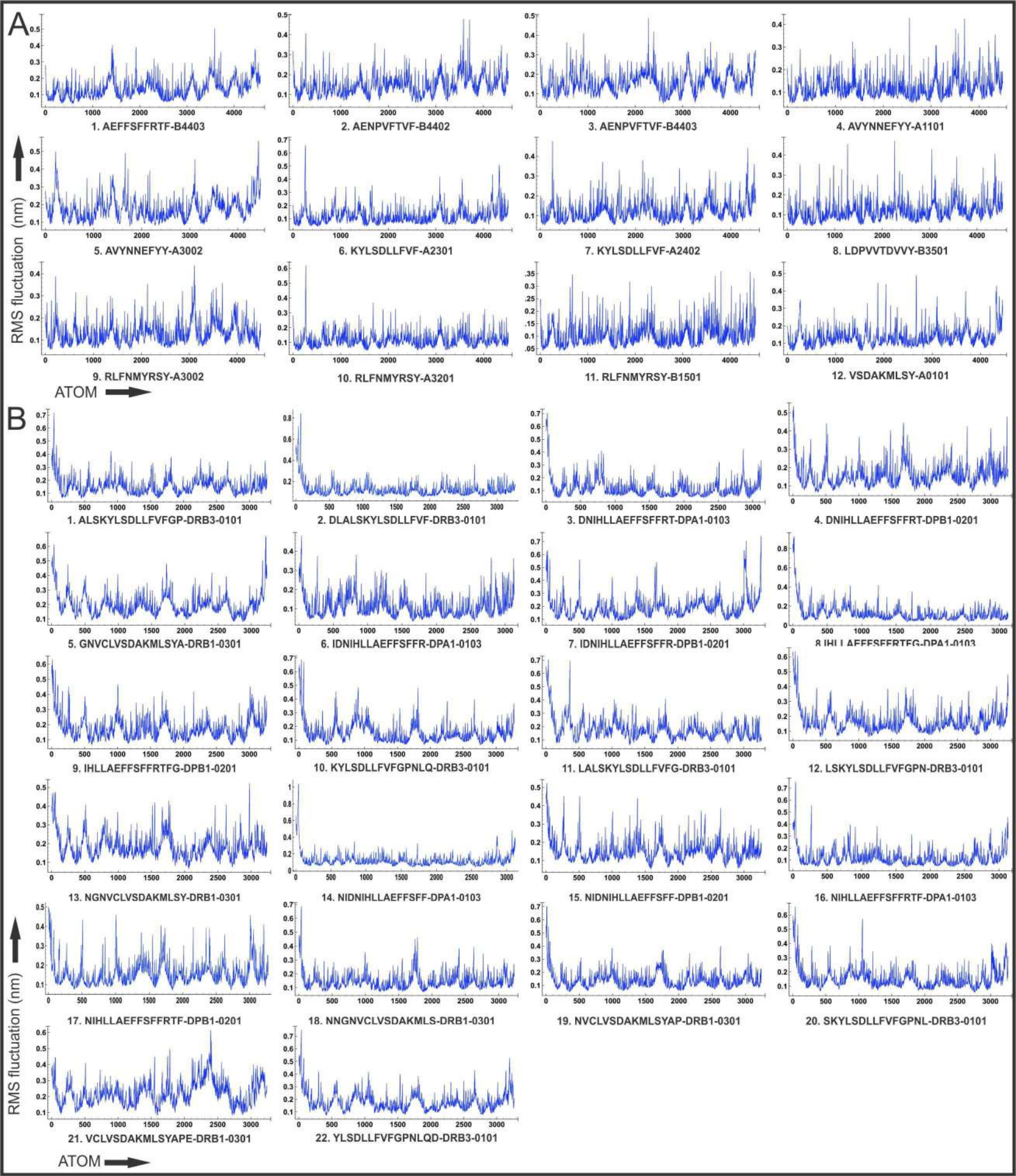
HTL multi-epitope vaccine:



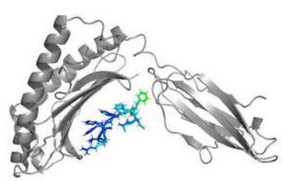


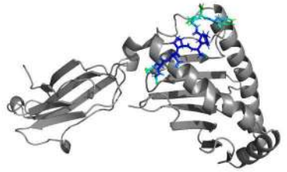

Nine NIPAH proteins were chosen
Protein sequences and available 3D structures retrieved from NCBI & RCSB PDB or modelled by SwissModel
Screening of potential CTL, HTL and B Cell epitopes
Screening of CTL epitope (IEDB "Proteasomal cleavage/ TAP transport/MHC class I combined predictor")
Screening of HTL epitopes (IEDB "MHC-II Binding Predictions")
Screening of protein sequence & structure based B Cell epitope (IEDB "B Cell Epitope Prediction Tools", DiscoTope2.0 & Ellipro)
Population coverage by screened CTL and HTL epitopes (IEDB "Population Coverage")
Characterization of screened potential epitopes
Epitope conservation analysis (IEDB "Epitope Conservancy Analysis")
Epitope toxicity prediction (ToxinPred)
Overlapping residue analysis for CTL, HTL & B Cell epitopes (Clustal Omega)
Molecular interaction study of selected overlapping CTL & HTL epitopes with HLA allele and TAP transporter
3D structure modelling of HLA alleles and selected CTL & HTL epitopes (Swiss-model & PEPstrMOD)
Molecular interaction analysis of selected overlapping epitopes with HLA allele (AutoDock Tool 4.2, AutoDock Vina, & Gromacs 5.1.4)
Molecular interaction analysis of selected CTL epitopes with TAP transporter (AutoDock Vina)
Design and characterisation of CTL & HTL Multi-Epitope Vaccines
Interferon gama inducing epitope prediction from designed MEVs ("IFNepitope")
MEVs allergenicity and antigenicity prediction (AlgPred & VaxiJen)
Physicochemical analysis of MEVs (ProtParam)
3D structure modelling and refinement of CTL & HTL MEVs (RaptorX, GalaxyWEB)
In-silico validation of refined CTL & HTL MEV models (RAMPAGE)
Discontinuous B-cell epitope prediction from CTL & HTL MEVs (IEDB "ElliPro")
Molecular interaction and complex stability study of CTL & HTL MEVs with TLR2 & TLR4
(PatchDock & Gromacs 5.1.4)
In-silico analysis of cDNA of CTL & HTL MEVs for cloning and expression in mammalian cell line (Human)
(Codon Usage Wrangler Tool & GenScript Rare Codon Analysis Tool)

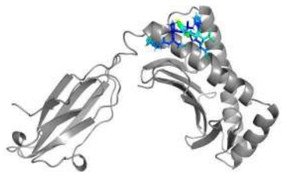


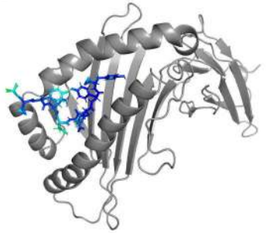

A
1. AEFFSFFRTF-B4403
2. AENPVFTVF-B4402
3. AENPVFTVF-B4403
4. AVYNNEFYY-A1101
5. AVYNNEFYY-A3002
6. KYLSDLLFVF-A2301
7. KYLSDLLFVF-A2402
8. LDPVVTDVVY-B3501
9. RLFNMYRSY-A3002
10. RLFNMYRSY-A3201
11. RLFNMYRSY-B1501
12. VSDAKMLSY-A0101
RMS fluctuation (nm)
ATOM
B
1. ALSKYLSDLLFVFGP-DRB3-0101
2. DLALSKYLSDLLFVF-DRB3-0101
3. DNIHLLAEFFSFFRT-DPA1-0103
4. DNIHLLAEFFSFFRT-DPB1-0201
5. GNVCLVSDAKMLSYA-DRB1-0301
6. IDNIHLLAEFFSFFR-DPA1-0103
7. IDNIHLLAEFFSFFR-DPB1-0201
8. IHLLAEFFSFFRTFG-DPA1-0103
9. IHLLAEFFSFFRTFG-DPB1-0201
10. KYLSDLLFVFGPNLQ-DRB3-0101
11. LALSKYLSDLLFVFG-DRB3-0101
12. LSKYLSDLLFVFGPN-DRB3-0101
13. NGNVCLVSDAKMLSY-DRB1-0301
14. NIDNIHLLAEFFSFF-DPA1-0103
15. NIDNIHLLAEFFSFF-DPB1-0201
16. NIHLLAEFFSFFRTF-DPA1-0103
17. NIHLLAEFFSFFRTF-DPB1-0201
18. NNGNVCLVSDAKMLS-DRB1-0301
19. NVCLVSDAKMLSYAP-DRB1-0301
20. SKYLSDLLFVFGPNL-DRB3-0101
21. VCLVSDAKMLSYAPE-DRB1-0301
22. YLSDLLFVFGPNLQD-DRB3-0101
RMS fluctuation (nm)
ATOM

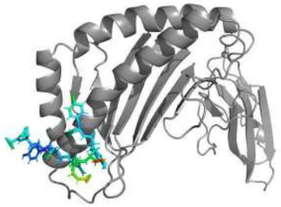
A. B-factor of CTL epitope & HLA I allele complesxes

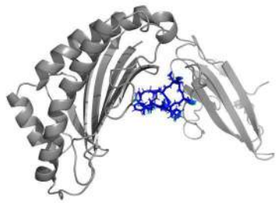

1. AEFFSFFRTF-B4403

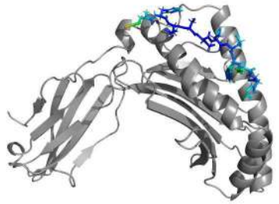

2. AENPVFTVF-B4402

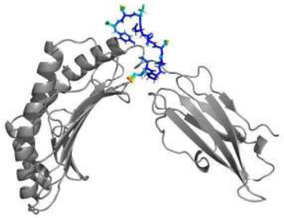

3. AENPVFTVF-B4403

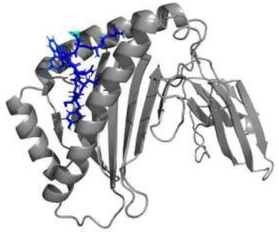

4. AVYNNEFYY-A1101

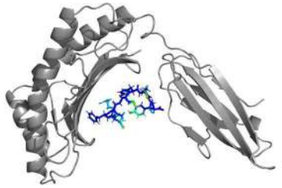

5. AVYNNEFYY-A3002

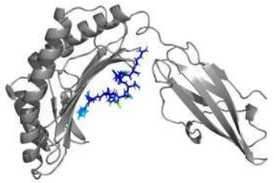

6. KYLSDLLFVF-A2301

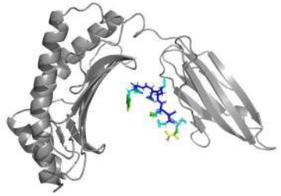

7. KYLSDLLFVF-A2402


8. LDPVVTDVVY-B3501

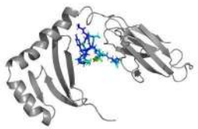

9. RLFNMYRSY-A3002

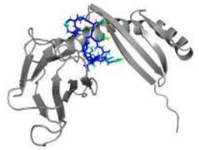

10. RLFNMYRSY-A3201

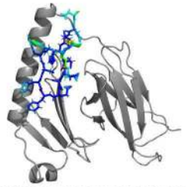

11. RLFNMYRSY-B1501

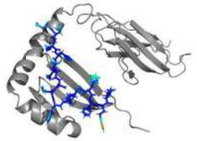

12. VSDAKMLSY-A0101

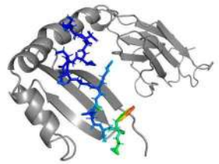
B. B-factor of HTL epitope & HLA II allele complesxes

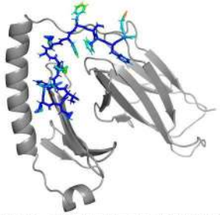

1. ALSKYLSDLLFVFGP-DRB3-0101

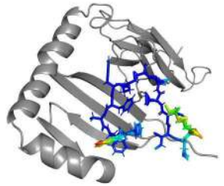

2. DLALSKYLSDLLFVF-DRB3-0101

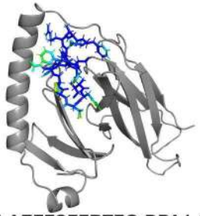

3. DNIHLLAEFFSFFRT-DPA1-0103

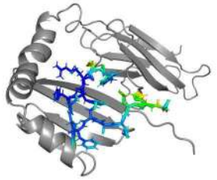

4. DNIHLLAEFFSFFRT-DPB1-0201

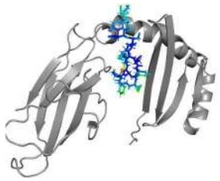

5. GNVCLVSDAKMLSYA-DRB1-0301

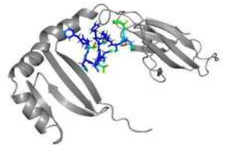

6. IDNIHLLAEFFSFFR-DPA1-0103


7. IDNIHLLAEFFSFFR-DPB1-0201

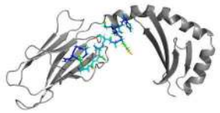

8.IHLLAEFFSFFRTFG-DPA1-0103

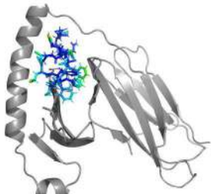

9. IHLLAEFFSFFRTFG-DPB1-0201

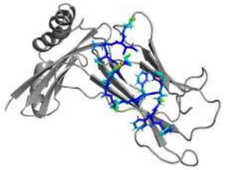

10. KYLSDLLFVFGPNLQ-DRB3-0101

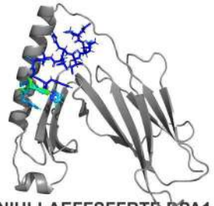

11. LALSKYLSDLLFVFG-DRB3-0101

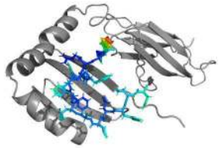

12. LSKYLSDLLFVFGPN-DRB3-0101

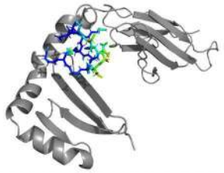

13. NGNVCLVSDAKMLSY-DRB1-0301

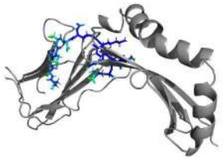

14. NIDNIHLLAEFFSFF-DPA1-0103

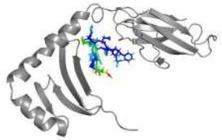

15. NIDNIHLLAEFFSFF-DPB1-0201

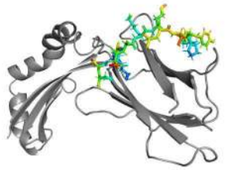

16. NIHLLAEFFSFFRTF-DPA1-0103

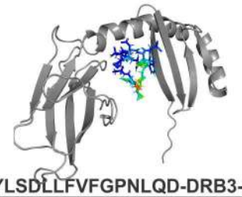

17. NIHLLAEFFSFFRTF-DPB1-0201


18. NNGNVCLVSDAKMLS-DRB1-0301


19. NVCLVSDAKMLSYAP-DRB1-0301


20. SKYLSDLLFVFGPNL-DRB3-0101


21. VCLVSDAKMLSYAPE-DRB1-0301


22. YLSDLLFVFGPNLQD-DRB3-0101