

RESEARCH ARTICLE

How Computational Epitope Mapping Identifies the Interactions between Nanoparticles Derived from Papaya Mosaic Virus Capsid Proteins and Immune System

Mahbobeh Zamani-Babgohari¹, Kathleen L. Hefferon^{1,*}, Tsu Huang¹ and Mounir G. AbouHaidar¹

¹Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, ON, Canada

Abstract: Background: Nanoparticles derived from plant viruses possess fascinating structures, versatile functions and safe properties, rendering them valuable for a variety of applications. Papaya mosaic Virus-Like Particles (VLPs) are nanoparticles that contain a repetitive number of virus capsid proteins (PMV-CP) and are considered to be promising platforms for vaccine design. Previous studies have reported the antigenicity of PMV nanoparticles in mammalian systems.

Materials and Methods: As experiments that concern vaccine development require careful design and can be time consuming, computational experiments are of particular importance. Therefore, prior to expressing PMV-CP in *E. coli* and producing nanoparticles, we performed an *in silico* analysis of the virus particles using software programs based on a series of sophisticated algorithms and modeling networks as useful tools for vaccine design. A computational study of PMV-CP in the context of the immune system reaction allowed us to clarify particle structure and other unknown features prior to their introduction *in vitro*.

Results: The results illustrated that the produced nanoparticles can trigger an immune response in the absence of fusion with any foreign antigen.

Conclusion: Based on the *in silico* analyses, the empty capsid protein was determined to be recognised by different B and T cells, as well as cells which carry MHC epitopes.

ARTICLE HISTORY

Received: August 02, 2018

Revised: May 09, 2019

Accepted: May 09, 2019

DOI:

10.2174/1389202920666190527080230

Keywords: PMV-CP, virus nanoparticles, antigenic epitopes, immune response, vaccine design, *in silico* analysis.

1. INTRODUCTION

Virus infections which lack effective treatments remain a serious threat to public health. Currently, vaccination offers the most potent strategy for long term protection against virus infections. Conventional vaccines presented as either attenuated or inactivated forms of the virus are the most effective available treatments for many diseases [1]. However, these vaccines are not efficacious against aggressive viral infections such as Human Immunodeficiency Virus (HIV), Hepatitis Virus C (HCV), Dengue viruses, *etc.* due to specific characteristics of the virus to escape immunity or infect people's immune system, which opens new insights into vaccine design. Furthermore, the live attenuated/inactivated viruses might become reactivated and replicate in patients resulting in severe infections. New developments in vaccine research are emerging to address these safety issues. Plant virus nanoparticles represent one of the recent promising approaches in this field. These particles, also called Virus-Like Particles (VLPs), have received much attention as platforms for vaccine development.

In the absence of the virus genome, structural proteins represent the sole subunits of virus-like particles and enable

the VLPs to mimic antigenic conformation of the original virus. This infers that plant VLPs are safe tools that can be used to elicit an immune response [2]. The ability of VLPs to self assemble and move systemically is the other advantage which distinguishes VLPs from many synthetic nanoparticles. Different host systems (such as *E. coli*, insect cells, yeast and plants) might be used for the production of VLPs as recombinant proteins and consequently, for the assembly of VLPs. The functionality and stability of these VLPs are crucial characteristics required for further analyses and vary depending on the expression system as well as the purification method. In some cases, for example, the inability of purified VLPs to remain stable at temperatures such as 37°C constitutes another challenge [3]. Regardless, plant derived VLPs such as Tobacco Mosaic Virus (TMV), Cowpea Mosaic Virus (CMV), Papaya Mosaic Virus (PMV) and Potato Virus X (PVX) have been explored for vaccine production against infectious diseases and cancer cells [4, 5]. Even though plant virus nanoparticles do not usually result in adverse reactions, according to preclinical tests, they can be highly immunogenic [4].

The procedure for developing a vaccine against an infectious disease using VLPs requires an understanding of the different phenomena that are involved in pathogen-host interactions as well as the molecular mechanisms involved in the life cycle of the pathogen. To date, there are quite a number of VLPs produced, but not all of them are applicable

*Address correspondence to this author at the Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, ON, Canada; Tel: 607 280 0094; E-mail: Kathleen.hefferon@utoronto.ca

as vaccine targets [6]. Peptide fusion, conjugation, encapsulation and electrostatic interaction are the major strategies to load VLPs. These methods, on the other hand, have potential drawbacks. For example, the peptide should be amenable to conjugation and fusion, thereby permitting correct VLP assembly and function [1]. Upon introduction to the immune system, VLPs are controlled by different factors to effectively stimulate an immune response. These include the diameter of VLPs allowing the dendritic cells to efficiently uptake VLPs, induction of the immune response, as well as activation of B and T cells and their downstream effects for mediating a reactive response [5, 7]. As one of the plant derived VLPs, PMV nanoparticles have been employed with both fusion and conjugation methods to trigger a cytotoxic immune response that affected MHC class I and human T cells [3, 8]. In these studies, the modified VLPs were identified as effective vaccine platforms against Influenzae and Hepatitis C viruses [5, 9].

To generate an appropriate vaccine target, the construction and examination of the designed platforms are conducted by careful experimental approaches and take a prolonged period of time to be evaluated and approved as sufficient for testing in clinical trials. Alternatively, a computational biology approach that makes use of various techniques such as algorithms, modeling of protein structures, prediction of immunogenicity and epitope binding sites, offers robust, cost-effective and rapid methodologies for assessing experimental endeavours in advance of the wet lab experimental approach [10]. Therefore, in this study, we aimed to generate PMV nanoparticles in a bacterial system and test the immunogenicity of the generated VLPs using bioinformatics analyses. Using this approach, possible drawbacks and further complications can be avoided. In addition to this, computational analysis allows us to study different aspects of viral antigenicity within the human system and therefore opens the potential for comparative studies with other vaccine platforms.

2. MATERIALS AND METHODS

2.1. Cloning of Papaya Mosaic Virus Capsid Protein (PMV-CP)

The genome sequence of Papaya mosaic virus had been cloned into pUC18 by Ikegami [11]. On the PMV-CP sequence, the second ATG was considered to be the start codon (amplifying amino acids 6 to 215) and the following forward and reverse primers were used respectively 5'-CATATGGTTGATTCTAAGAAAAGTCAA-3' and 5'-CCTGCTCATGGCACC ACCATCCACTAGTCTTAA-GCACCATCACCATCACCATTAGTAAGGATCC-3'. The amplified region was cloned into pET29b using *Bam* HI and *Nde* I restriction sites (Fig. 1A). The insertion was confirmed by restriction digestion and polymerase chain reaction, followed by sequencing analysis.

2.2. Expression of Papaya Mosaic Virus Capsid Protein (PMV-CP) in *E. coli*

To express PMV-CP protein, the *E. coli* strain BL21 codon plus was transformed with the pET29b plasmid containing the PMV-CP. The transformed cells were grown overnight in LB media (1% tryptone, 0.5% yeast extract and 1%

NaCl containing 50 µg/ml kanamycin and 50 µg/ml chloramphenicol). The next day, a 50 µL of this preculture was inoculated to 1 mL of the fresh LB medium containing the same antibiotics and grown to reach $OD_{600} = 0.5$. The protein was induced by adding 1mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) and overnight culture. Thereafter, the cultures were centrifuged for five minutes and the pellet was resuspended in Tris-HCl, pH 7.0 and 5X SDS loading dye containing 2-Mercaptoethanol. The protein extracts were boiled for five minutes and centrifuged at 4°C for another 10 minutes. PMV-CP are expected to remain soluble and present within the supernatant of the solution.

Electrophoresis was carried out on a 12 % SDS-polyacrylamide gel. The fractionated proteins were transferred onto nitrocellulose membranes. The membranes were pre-incubated with blocking solution [3% (w/v) skim milk in TBS containing 0.05% Tween-20] for 1 h, incubated with mouse anti-PMV IgG (diluted in blocking solution to 1: 3,200), and probed with peroxidase-conjugated anti-mouse IgG (diluted in blocking solution to 1: 1,000). Proteins were detected using an enhanced chemiluminescent (ECL) western blotting detection reagent.

2.3. Purification of VLPs

To purify the VLPs, the bacterial pellet resulting from the overnight induction was resuspended in 50mM Tris-HCl, pH 7.0. The solution was then passed through a French press twice for proper cell disruption. Afterwards, the resulting solution was centrifuged at 4°C for 30 min. The supernatant was used for ultra-centrifugation at 35,000 rpm for 3.5 hours at 4°C in a Beckman 50.2 TI rotor. The resulting pellet was resuspended in Tris-HCL buffer at 4°C for further use.

2.4. Electron Microscopy

To visualize the assembled VLPs using a transmission electron microscope, the grids were washed with Cytochrome C solution. The diluted proteins were then absorbed on carbon-coated formvar grids and stained with 1 % uranyl acetate at room temperature. A Hitachi HT7700 transmission electron microscope was used for observation of the grids and visualization of the VLPs.

2.5. Computational Techniques

The protein sequence of PMV-CP (accession no, D00240.1) was obtained from a previous report by AbouHaidar in 1988 [11] and used for the following bioinformatic analyses to detect the immunogenicity of these VLPs in human immune system.

2.5.1. Prediction and B Cell Epitopes

Bcepred was conducted for recognition of B cell epitopes of PMV nanoparticles http://tools.immuneepitope.org/tools/ElliPro/iedb_input. ElliPro was also utilized for prediction of discontinuous B cell epitopes. The server provides linear and discontinuous antibody epitope prediction based upon the 3D structure of the protein. ElliPro links predicted epitopes with a score, defined as a PI (Protrusion Index) value averaged over epitope residues. For this method, the protein's 3D structure is predicted using ellipsoids, where an ellipsoid with $PI = 0.9$ includes 90% of the protein residues [12]. Pre-

diction of 3D structure of the PMV nanoparticles by I-TASSER was also conducted with a higher accuracy.

2.5.2. Identification of T Cell Epitopes

T cell epitope prediction was conducted using a CTLPred server that employs a direct method for CTL epitope prediction. The database can be accessed at <http://crdd.osdd.net/cgi-bin/ctlpred/ctlpred.pl>. Machine learning techniques SVM and ANN are used to develop a CTL epitope prediction. This technology also allows for consensus and combined prediction.

2.5.3. MHC Class I and II Epitope Prediction

EpiJen v10 at <http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm> was used for MHC class I epitope prediction. In MHCpred, Allele specific Quantitative Structure Activity Relationship (QSAR) models were generated using Partial Least Squares (PLS). The server is based on additive algorithms. Three different interactions are involved including the interactions between single amino acids and the binding site, between adjacent and any other amino acids and their impact on binding. In addition, the algorithm considers modelling of different stages such as proteasome cleavage, TAP binding and MHC binding as the significant factors in epitope recognition [13].

Prediction of MHC class II binding epitopes responding to PMV nanoparticles was employed at http://crdd.osdd.net/tmp/propred/_using quantitative matrices. The distinct characteristics of this server are the graphical view plots, a profile describing the score, the profile threshold and best scoring subsequence. The peaks of score distribution associate with the binding affinity of peptide. In fact, the binding strength correlates with immunogenicity [14].

2.5.4. Analyses of Protein Structure and Disorders

The secondary and tertiary structures of the generated PMV nanoparticles were analysed using the PSIPRED method at http://npsaprab.i.bcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4. The method utilizes two feed-forward neural networks analysing the output of PSI BLAST. The server accuracy based on three state range (Q3) is estimated at 81.6 % [15]. The 3D structure was developed at https://zhanglab.ccmb.med.umich.edu/cgi-bin/itasser_submit.cgi [16]. The disordered protein structure was tested at <http://iupred.enzim.hu/pred.php> which is a publicly available database and identifies the abnormal regions lying within the protein structure.

2.5.5. Prediction of Allergenic Residues

The possibility of acquiring an allergic reaction of PMV nanoparticles was tested by the ALGPred server at http://crdd.osdd.net/cgi-bin/algpred/algpred_main_temp.pl. The server employs the analysis according to the physicochemical parameters of the target protein.

2.6. Validation of Data

Different computational servers predicted B and T cell epitopes, MHC binding and intrinsically disordered regions. The programs generally use CTLPred, ElliPro, EpiJen and PSIPRED databases. CTLPred database consists of 7000

epitopes and classifies more than 2000 epitopes and non-epitopes where the accuracy is based on cross-validation (LOOCV) at a cutoff score. In this method, sensitivity and specificity are almost equal. ElliPro offers the best AUC value of 0.732 [12].

3. RESULTS

3.1. Production of PMV Nanoparticles as Designed Vaccine Platforms

Papaya mosaic virus nanoparticles were expressed in *E. coli*, strain BL21 (DE3) codon plus. The structure of the expressed protein (PMV nanoparticles) was compared with the PMV virus extracted from infected plants where the appearance of the viral nanoparticles looked largely similar. However, comparing with the wild type virus extracted from infected plants, the PMV-CP expressed in *E. coli* resulted in shorter-sized particles (Fig. 1C and D).

Papaya mosaic virus is a filamentous virus with an average size of 500-540 nm in length and 13-14.2 nm in diameter. PMV nanoparticles are composed of 1400 capsid proteins [17, 18]. SDS-PAGE analysis illustrated that the high yield of recombinant protein. The molecular weight of PMV capsid protein was estimated to be between 21 to 26 kDa. The protein was observed at the expected molecular size even though other studies have reported a higher molecular weight for PMV CP of 33 kDa [5]. In our case, the inclusion of a six histidine-tag did not result in an increase of molecular weight which may be due to differences among expression vectors or bacterial strains (Fig. 1A).

3.2. Analysis of Physicochemical and Structural Characteristics of Generated PMV Nanoparticles

The molecular weight of the generated PMV capsid protein was estimated to be 23.85 kDa and the theoretical pI was calculated to be 7.83. The protein contains 15 negative and 16 positively charged residues. They were estimated to be soluble with the probability of 0.56 upon overexpression in the bacterial host. The extinction coefficient of the protein was estimated to be $17,085 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nm. The half-life was predicted to be 30 hours in mammalian reticulocytes and above 20 and 10 hours in yeast and *E. coli*, respectively. The instability index was computed as 42.29, and as a result, PMV CP was considered to be an unstable protein. The aliphatic index was 72.41 and the grand average of hydropathicity (GRAVY) was calculated to be -0.220 (Fig. 1B).

The PSIPRED server revealed that the proteins contain 32.71% alpha helix, 50.47% random coil and 16.82% extended strand. The schematic picture of the secondary structure is shown in Fig. (2).

The tertiary structure of PMV capsid was developed using I-TASSER (Fig. 2C). Based on the estimated C-score (1.72), predicted structure poses high confidence within the defined ranges (-5, 2) of the program. The higher score associates with improved model confidence. The program predicted the tertiary structure of the PMV-CP ligand binding sites (Fig. 2C and D). I-TASSER also predicted the residue D64 as the active site residue.

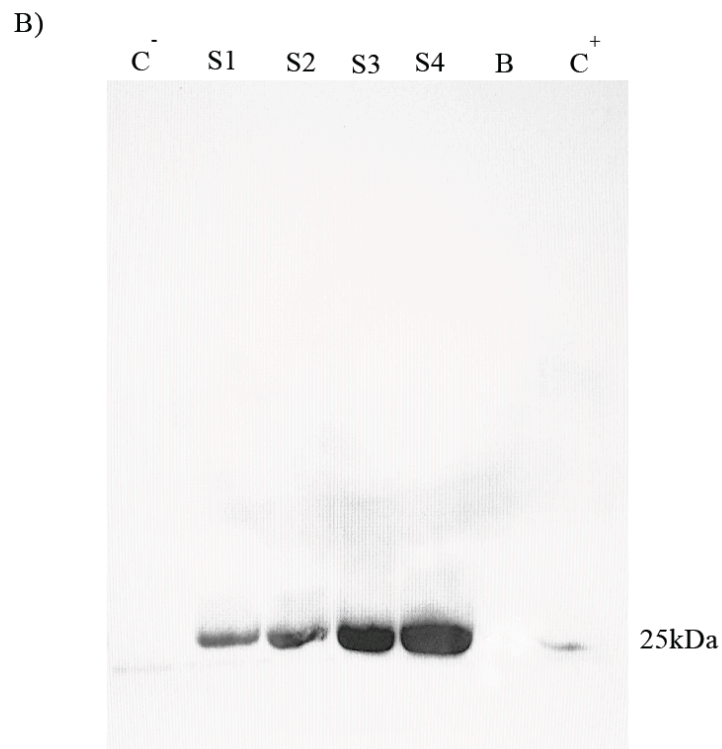
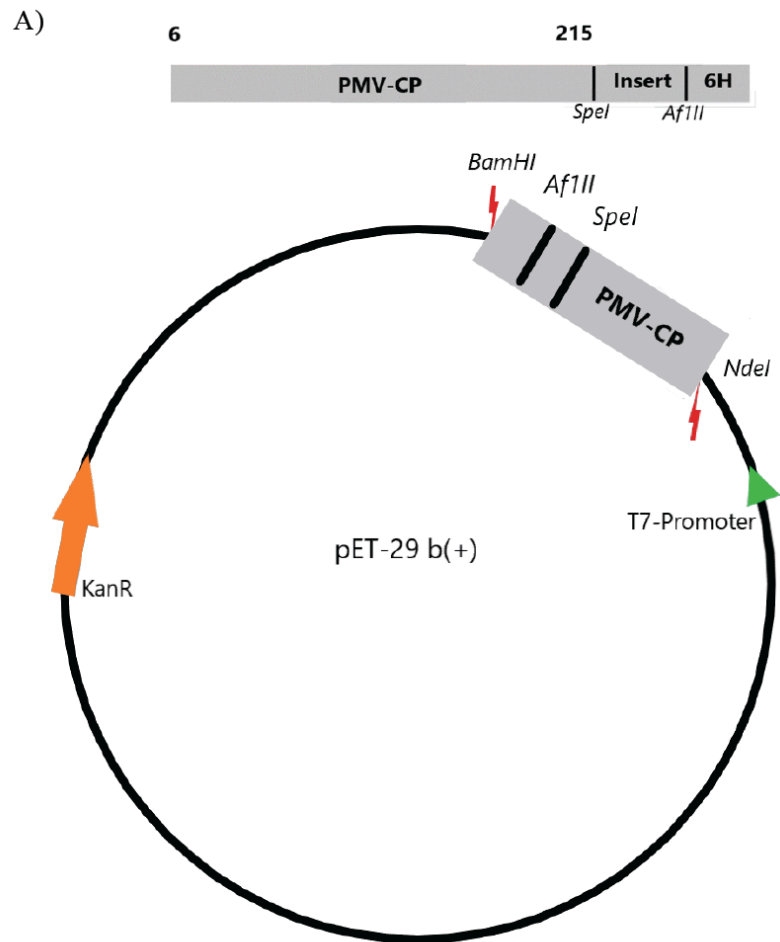


Fig. (1). contd....

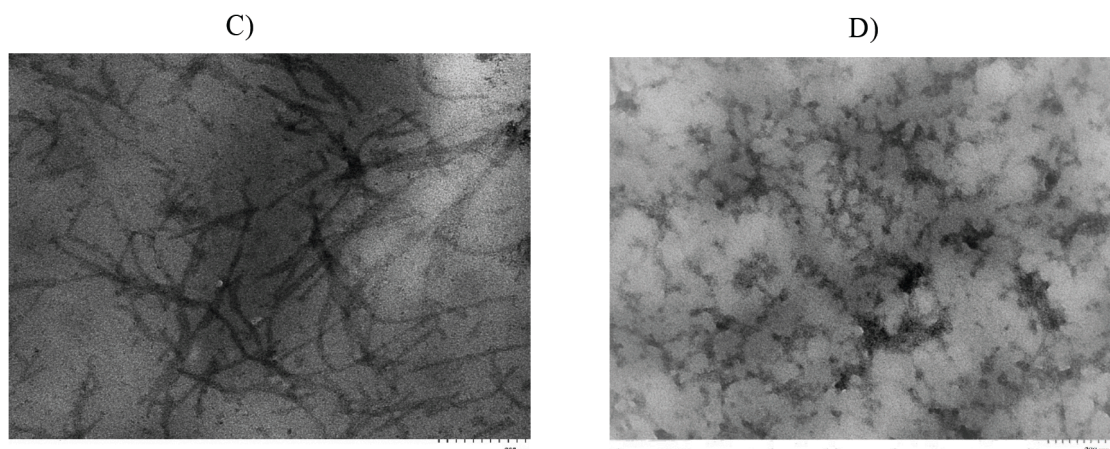


Fig. (1). Production and identification of recombinant PMV-CP. **(A)** A schematic diagram of the designed construct of PMV-CP (amino acids 6-215) inserted into the pET-29b. A 6x His tag was also fused to the C-terminus of the capsid protein. The foreign epitopes, shown as insert on the diagram, can be inserted into C-terminus of the CP using the designed restriction sites (*SpeI* and *AflIII*) **(B)** Western blot experiment confirms the expression of PMV-CP protein in *E. coli*. The production of PMV-CP was confirmed by Western blot experiment using anti-PMV antibody. The letters C, C, S1, S2, S3, S4 and B respectively correspond to positive control (PMV virus), negative controls (untransformed vector pET29b), and four samples from different dilutions and blank. **(C)** PMV like particles isolated from infected plants were visualized with electron microscopy (positive control). **(D)** virus like particles were isolated by ultracentrifugation of PMV-CP recombinant proteins and were identified using electron microscopy. Scale bars represent 200 nm.

3.3. Identification of Disorder Structures Within the Capsid Protein

Intrinsically Disordered Proteins (IDPs) do not have a stable structure under native conditions. IUPred2 and ANCHOR2 determined the disordered binding regions of PMV-CP. The server can also identify protein regions with no stable structure based upon the redox state of their environment [19]. The results illustrated that there is a disordered binding region in C-terminus end of PMV capsid where the ANCHORED2 exhibited a peak between position 180-190 (Fig. 3). The protein also consists of a few disordered regions within both the N and C-terminal ends between positions 10-30 and 160-214 (Fig. 3A and B).

3.4. Prediction of Protein Allergenicity

Prediction of allergenic proteins is a significant strategy especially in the context of therapeutics and engineered pharmaceuticals [20]. Allergenicity analysis of the protein was tested using the ALGPred program. The results indicate that the generated PMV capsids are non-allergenic.

3.5. B and T-cell as well as MHCs Epitope Prediction

The continuous and discontinuous B-cell epitopes were predicted by the ElliPro server using default parameters. The server identifies antibody epitopes of input protein and predicts the protein-antibody complex as a 3D structure. In PMV-CP protein sequence, the epitopes associated with different protein chains were selected by the program (Tables 1 and 2). The 3D structure of the nanoparticles with corresponding epitopes was also determined (Fig. 4). In total, 11 epitopes were selected using this program.

The CTLPred program identified three epitopes as the final T-cell predicted epitopes (Table 3). The MHCs are linked to T-cell epitopes which is a crucial step for T-cell's activation [21]. For this reason, putative MHCs binding

epitopes were also predicted using the EpiJen program (Tables 4 and 5).

It should be mentioned that if a protein functions as an antigen, an epitope either corresponds to a small peptide derived from the original protein antigen (continuous epitope) or a patch of atoms on the surface of the protein (discontinuous epitope) [13].

4. DISCUSSION

The availability of protein databanks with vast amounts of data facilitates the determination of epitopes for designing highly effective, structure-based vaccines [22]. To develop a rational vaccine platform, understanding the structural and biochemical properties of antigens is also crucial. A computational approach creates a clear view of antigen-antibody complexes and helps to uncover the molecular mechanism of interactions between host and pathogen, leading to a more efficient outcome. *In silico* techniques for advanced vaccine design are time and cost effective; they can also uncover targets where conventional methods cannot [23]. Here, we took advantage of computational analyses to predict the immune response to PMV-CP nanoparticles.

Papaya mosaic virus nanoparticles that have been fused or conjugated with antigenic peptides have triggered immunogenic responses in mammals [5, 10, 24, 25]. Over the past decade, various virus nanoparticles have been generated against different diseases [6]. Previous reports, however, focused mostly on the antigenicity of the fused epitopes for achieving the highest mimicry of vaccine platforms. It is important to note that achieving an efficient virus nanoparticle for vaccine production requires a safe expression system in which large quantities of the desired protein can be easily produced [6]. Even though physiochemical and structural properties of virus nanoparticles highly affect their efficacy, the significance of a well-designed platform should never be neglected [25]. We, therefore, examined the immunogenicity

A)

```

1 M S T P N T A F P A I T Q E Q M S S I K V D P T S N L L P S Q E Q L K S V S T L M V A A K V P A A S 50
51 V T T V A L E L V N F C Y D N G S S A Y T V T G P S S I P E I S L A Q L A S I V K A S G T S L R K F 100
101 C R Y F A P I I W N L R T D K M A P A N W E A S G Y K P S A K F A A F D F F D G V E N P A A M Q P P 150
151 S G L I R S P T Q E E R I A N A T N K Q V H L F Q A A A Q D N N F T S N S A F I T K G Q I S G S T P 200
201 T I Q F L P P P E T S T L K
    
```

KEY Annotations Helix Sheet Disordered Disordered protein binding Dompred Boundary DomSSEA Boundary

Annotations M L E E A D

B)

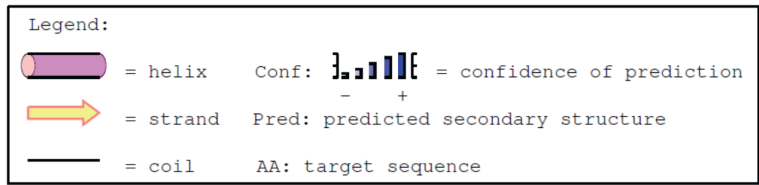
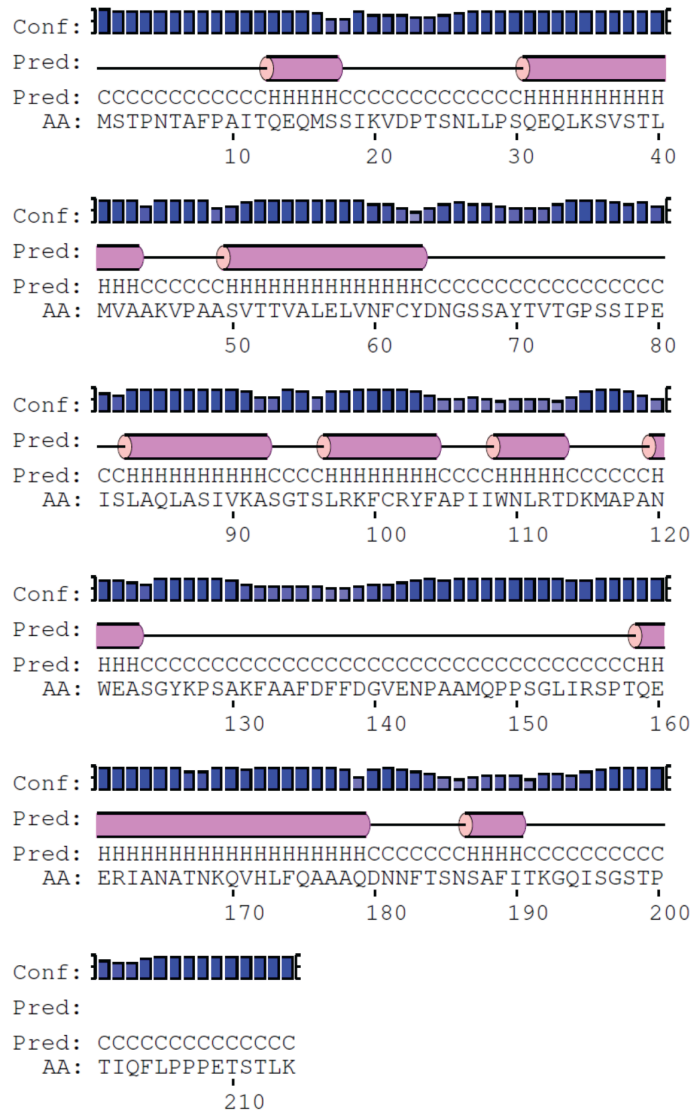


Fig. (2). contd....

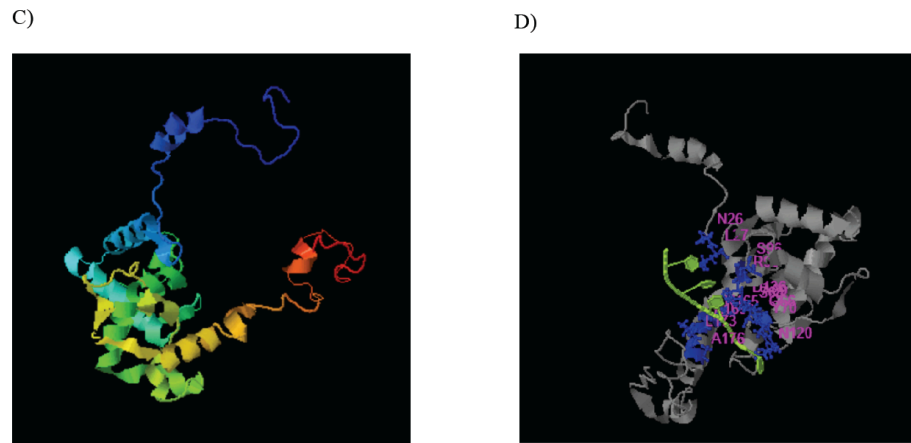


Fig. (2). The secondary and 3D structures of engineered PMV-CP. **(A)** Different features on the sequence are color coded as shown below the protein structure. **(B)** PISPRED predicted the graphical structure of PMV-CP recombinant protein. **(C)** The 3D structure of the protein using ITASSER server. The blue color represents N-terminus and red shows C-terminus of the protein. The server uses different colors based on amino acids properties where the bright colors show the polar amino acids and non-polar residues are in dark shades (*i.e.* the dark blue shows phenylalanine and tyrosine, light blue indicate lysine and arginine, bright blue represents asparagine and glutamine, yellow indicates cysteine and methionine, green associates with valine, leucine and isoleucine, orange stands for serine and threonine and red represents aspartate and glutamate). **(D)** Ligand binding sites predicted by ITASSER program. The blue ball and sticks represent binding residues where pink colored residues at positions 26, 27, 66, 68, 69, 70, 96, 98, 120, 136, 165, 169, 173 and 176 possessing higher C-score were selected by the program. The green color indicates the binding ligands. The remaining residues which are not involved in the indicted ligands binding are in gray color. (The color version of the figure is available in the electronic copy of the article).

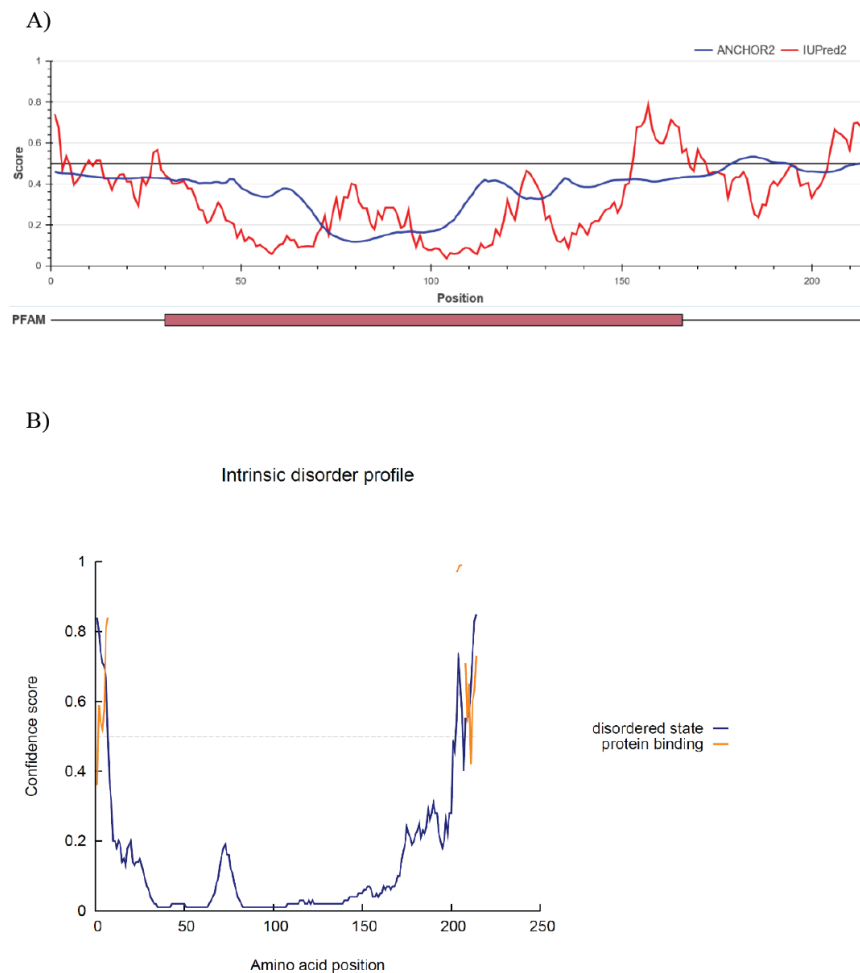


Fig. (3). Identification of disordered protein regions using different programs. **(A)** IPRED program predicted the disordered regions of PMV-CP by IUPred2 and ANCHOR2. Predicted disorders are observed at the N and C-terminus of the protein where the score is higher. **(B)** Intrinsic disorder profile of the protein using PSIPRED predicted N and C-terminus ends of the protein as disordered sites where the disordered state includes scores above the confidence score (0.5). The orange color indicates disordered regions predicted as binding residues. (The color version of the figure is available in the electronic copy of the article).

Table 1. Prediction of linear B-cell epitopes.

No.	Chain	Start	End	Peptide	Residues	Score
1	A	11	28	IAFPAITQEQMSSIKVDP	18	0.883
2	A	122	137	MAPANWEASGYKPSAK	16	0.751
3	A	72	101	SSAYTTVTGPSSIPEISLAQLASIVKASGT	30	0.652
1	B	11	26	IAFPAITQEQMSSIKV	16	0.88
2	B	120	141	DKMAPANWEASGYKPSAKFAAF	22	0.718
3	B	72	94	SSAYTTVTGPSSIPEISLAQLAS	23	0.696

Table 2. Prediction of discontinuous B-cell epitopes.

No.	Residues	Residue	Score
1	A:I11, A:A12, A:F13, A:P14, A:A15, A:I16, A:T17, A:Q18, A:E19, A:Q20, A:M21, A:S22, A:S23, A:I24, A:K25, A:V26, A:D27, A:P28	18	0.883
2	B:I11, B:A12, B:F13, B:P14, B:A15, B:I16, B:T17, B:Q18, B:E19, B:Q20, B:M21, B:S22, B:S23, B:I24, B:K25, B:V26	16	0.88
3	B:F66, B:D69, B:N70, B:S72, B:S73, B:A74, B:Y75, B:T76, B:T77, B:V78, B:T79, B:G80, B:P81, B:S82, B:S83, B:I84, B:P85, B:E86, B:I87, B:S88, B:L89, B:A90, B:Q91, B:A93, B:S94, B:K97, B:D120, B:K121, B:M122, B:A123, B:P124, B:A125, B:N126, B:W127, B:E128, B:A129, B:S130, B:G131, B:Y132, B:K133, B:P134, B:S135, B:A136, B:K137, B:F138, B:A139, B:A140, B:F141, B:P163, B:T164, B:Q165, B:E166, B:E167, B:R168, B:I169, B:A170, B:N171, B:A172, B:T173, B:N174	60	0.705
4	A:T44, A:L45, A:V47, A:A48, A:A49, A:K50, A:V51, A:P52, A:F66, A:D69, A:N70, A:S72, A:S73, A:A74, A:Y75, A:T76, A:T77, A:V78, A:T79, A:G80, A:P81, A:S82, A:S83, A:I84, A:P85, A:E86, A:I87, A:S88, A:L89, A:A90, A:Q91, A:A93, A:S94, A:I95, A:K97, A:A98, A:M122, A:A123, A:P124, A:A125, A:N126, A:W127, A:E128, A:A129, A:S130, A:G131, A:Y132, A:K133, A:P134, A:S135, A:A136, A:K137, A:A140, A:F141, A:Q165, A:E166, A:I169, A:A170, A:N171, A:A172, A:T173, A:N174	62	0.691
5	B:L45, B:V47, B:A48, B:A49, B:K50, B:V51	6	0.637

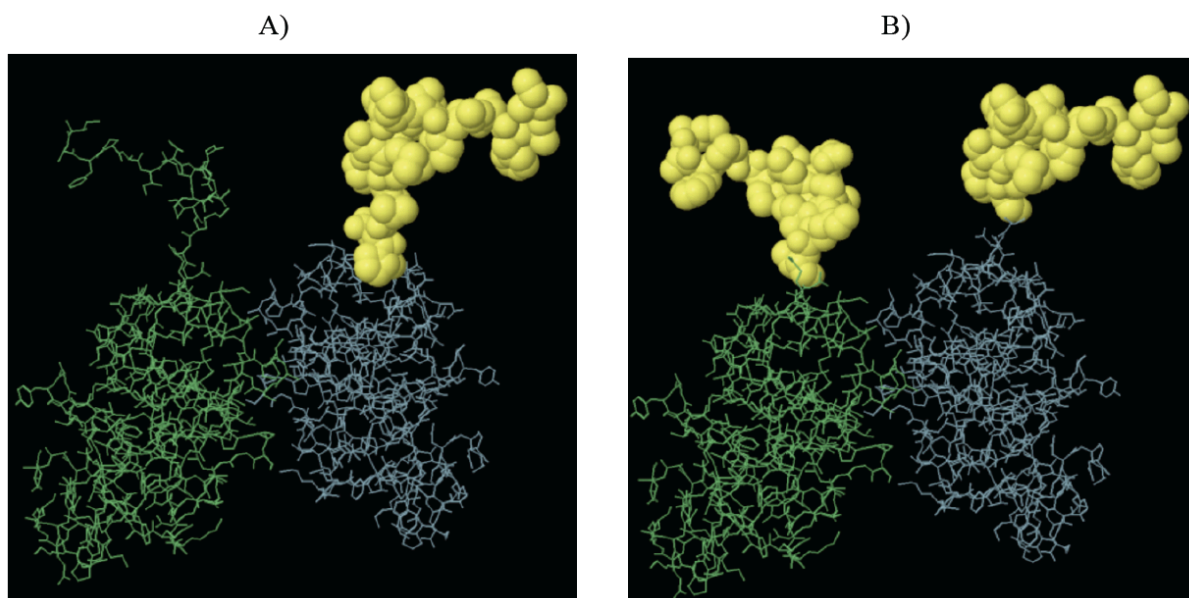


Fig. (4). B-cell prediction epitopes. Ellipro software program predicted continuous and discontinuous B-cell epitopes of the recombinant PMV-CP. (A) The 3D structure of continuous (linear) and (B) discontinuous B-cell epitopes on the surface of the capsid protein. The yellow color indicates the residues identified by B-cell epitopes. Discontinuous epitopes are more in number as they include a patch of atoms on the protein surface. Continuous epitopes are longer in length but less in number as they are small peptides derived from original proteins. (The color version of the figure is available in the electronic copy of the article).

of PMV capsids using immunoinformatic programs whereby sophisticated algorithms and processors analysed the input data. Papaya mosaic virus capsid protein (PMV-CP 6-215 aa), the subunit of PMV nanoparticles, was employed to predict structures, B and T-cell responses and MHC binding epitopes. We designed insertion sites at the C-terminus of the capsid protein wherein any foreign peptide can be inserted and tested for antigenicity using the present approach prior to further lab experiments.

Table 3. CTLPred predicate the T-cell epitopes.

Peptide Rank	Start Position	Sequence	Score
1	128	PSAKFAAFD	1.000
2	11	ITQEQMSSI	0.990
3	155	RSPTQEERI	0.990

Structural analyses of the designed PMV-CP indicated that the accuracy of the predicted 3D structure was reliable, based on I-TASSER scores, C score= 1.72, TM= 0.96± 0.05 and RMSD= 2.2± 1.7 Å. According to the program standards, any TM score greater than 0.5 is considered to be a proper topology prediction. The secondary structure revealed few disordered residues at the N and C-terminus of the capsid protein, which might correspond to the protein binding sites. In line with this result, crystallography analysis of the Papaya mosaic capsid protein has reported that the five residues at the N-terminus of PMV-CP, as well as at residue 175 at the C-terminus, were disordered [26]. Disordered protein regions usually contain short motifs which are considered significant for protein function [27]. PSIPRED predicted the domain boundaries using PSI-BLAST aligned termini algorithm where three peaks at positions 82, 112 and 125 were observed. The server only selected the peak located at position 112 as the putative boundary domain.

Table 4. Predicted MHCs (class I).

Starting Position	Peptide	Predicted -LOGIC50 (M)	Predicted IC50 Value (NM)
60	ALELVNFCY	9.773	0.17
101	TSLRKFCRY	7.527	29.72
172	ATNKQVHLF	7.004	99.08
56	VTTVALELV	6.768	170.61
124	PANWEASGY	6.693	202.77
189	FTSNSAFIT	6.675	211.35
182	AAAQDNNFT	6.636	231.21
100	GTSLRKFCR	6.515	305.49
25	KVDPTSPLL	6.456	349.95
76	TTVTGPSSI	6.383	414
49	AKVPAASVT	6.344	452.9

The physiochemical analysis of PMV-CP revealed that the protein is unstable, which may in turn lead to instability of the virus nanoparticles. In fact, protein stability is one of the main issues regarding vaccine production, as the nanoparticles require enough time to trigger an antigenic response. Otherwise, virus nanoparticles will be degraded before having a chance to interact with the immune system. It has been reported that a PMV construct harbouring a fusion peptide at the C-terminus was not stable at temperatures above 30°C [24]. This observed issue might simply be related to the structure of capsid protein which along with the fusion epitope reduces stability. Changing the site of fusion, epitope size and the expression strategy may result in improving nano-particle stability [9]. The wet lab experiment indicated that PMV-CP nanoparticles were stable at 4°C at least for several weeks.

Identification of PMV-CP associated B-cell epitopes by ElliPro server containing a clustering algorithm, MODELLER program and Jmol viewer resulted in six continuous and five discontinuous B-cell epitopes based on the PI value. It should be mentioned that an epitope can be continuous or discontinuous [12]. Both of these epitope formats were predicted in this study. The Jmol viewer allowed for the visualization of B-cell epitopes on the 3D structure of PMV nanoparticles.

One of the most significant factors in subunit vaccine design are peptides which stimulate cytotoxic T lymphocytes; these were identified using an artificial neural network and SVM technology. The predicted epitopes are not necessarily capable of binding to MHCs. MHC epitopes were predicted using EpiJen and MHCpred programs. The algorithm considers proteasome cleavage and TAP binding, allowing a more reliable MHC binding mechanism. MHCs are known for their role in the immune response through binding to antigenic peptide segments which finally activates specific T-cells. Using chemical and immunological techniques, Rioux and colleagues identified the residues of PMV-CP exposed

Table 5. Predicted epitopes of MHCs (class II).

Rank	Sequence	Position	Score	Highest Score (%)
1	IVKASGTSL	94	1.39	23.17
2	FCRYFAPII	105	0.37	6.17
3	FTSNSAFIT	188	0.34	5.67
4	MVAAKVPA	45	0.07	1.17
5	LRKFCRYFA	102	-0.02	0
6	LKSVSTLMV	38	-0.35	0
7	YTTVTGPSS	74	-0.58	0
8	LMVAAKVPA	44	-0.7	0
9	VHLFQAAAQ	176	-0.72	0
10	YKPSAKFAA	131	-1	0
11	LFQAAAQDN	178	-1.2	0
12	FDFFDGVEN	140	-1.22	0
13	VALELVNFC	58	-1.4	0
14	IANATNKQV	168	-1.5	0
15	ISLAQLASI	86	-1.6	0
16	LAQLASIVK	88	-1.8	0
17	FCYDNGSSA	65	-2.1	0
18	FDGVENPAA	143	-2.18	0
19	VPAASVTV	50	-2.21	0
20	VNFCYDNGS	63	-2.3	0

on the capsid surface [18]. Comparison of the reported regions with our predicted immunogenic epitopes supports the present results. For instance, B-cell epitopes located at positions 122-137, 120-141; T-cell predicted epitopes covering the region 128-136; MHC class I with predicted epitopes at 172-180, 124-132, 182-191, 189-197 and MHC class II with the predicted region at 188-196 containing some residues in common with previous studies [3]. However, there are new epitopes predicted in this study for the first time. Rioux mentioned that it is unlikely that they uncovered all of the surface exposed regions because of limitations regarding the immunoblot approach [3]. In fact, identification of such antigenic fragments is a challenge in the context of conventional strategies in which overlapping fragments that are required to be synthesized and examined for immunogenicity is both time consuming and costly. On the other hand, scanning software can identify B-cell, T-cell and MHC epitopes [28]. This method is very useful for developing vaccine subunits. PMV nanoparticles fused with antigenic peptides have been reported by Denis and colleagues [5] to be internalized by APCs (antigen presenting cells) associated with MHC molecules. In the current study, the predicted MHCs belonged to both MHC classes, indicating that PMV nanoparticles are capable of being identified by T-helper cells ($CD4^+$). Plant virus nanoparticles might also interact with other proteins on

the cell surface or internal cell proteins opening a novel avenue for research. Cowpea Mosaic Virus (CPMV), for example, was shown to interact with vimentin protein on the surface of endothelial cells introducing vimentin as a vascular endothelial marker [29]. CPMV was also used for *in situ* vaccination where activation of the innate immune system activated phagocytes as a combination method of therapy with CD47 blocking antibody in cancer therapy [30]. Nanoparticles derived from plant viruses have already been used in different fields as vaccine platforms, imaging devices and so forth. Computational analyses of these nanoparticles prior to any wet lab experiment can shorten the research path and result in an innovative and reliable outcome which might take much longer time in the absence of such analyses.

CONCLUSION

One of the primary steps regarding the application of computational biology for vaccine design includes the recognition of the immunogenic epitopes from the non-immunogenic ones. In this study, bioinformatics programs were used as an alternative approach along with experimental techniques to improve the design platform. Identification of antigenic sites recognized by the host immune system provides us with useful information for engineering the anti-

gen where necessary. These results predicted the immunogenicity of empty nanoparticles and permitted comparison with wet lab-based experiments. Indeed, using computational approaches enabled us to predict specific epitopes associated with the certain class of immune cells, thus shedding light on future studies uncovering the underlying mechanisms of the immune response of virus nanoparticles for vaccine development.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Denis Leclerc at Laval University for kindly providing us with the antibody needed for the experiment. We also thank Dr. Srividhya Venkataraman and Dr. Nasir Mahmood at the University of Toronto for their technical support and advise in conducting experiments.

REFERENCES

- [1] Rohovie, M.J.; Nagasawa, M.; Swartz, J.R. Virus-like particles: next-generation nanoparticles for targeted therapeutic delivery. *Bioeng. Translational Med.*, **2017**, *2*(1), 43-57.
- [2] Lee-Shanok, R. Construction and preliminary characterization of papaya mosaic virus as an expression vector for the presentation of foreign epitopes. M.Sc. Thesis, University of Toronto: Canada, **2001**.
- [3] Rioux, G.; Babin, C.; Majeau, N.; Leclerc, D. Engineering of papaya mosaic virus (PapMV) nanoparticles through fusion of the HA11 peptide to several putative surface-exposed sites. *PLoS One*, **2012**, *7*(2), e31925.
- [4] Hefferon, K.L. Repurposing plant virus nanoparticles. *Vaccines*, **2018**, *6*(11), 1-10.
- [5] Denis, J.; Majeau, N.; Acosta-Ramirez, E.; Savard, C.; Bedard, M.C.; Simard, S.; Lecours, K.; Bolduc, M.; Pare, C.; Willems, B.; Shoukry, N.; Tessier, P.; Lacasse, P.; Lamarre, A.; Lapointe, R.; Lopez, M.C.; Leclerc, D. Immunogenicity of papaya mosaic virus-like particles fused to a hepatitis C virus epitope: evidence for the critical function of multimerization. *Virology*, **2007**, *363*, 59-68.
- [6] Roy, P.; Noad, R. Virus-like particles as a vaccine delivery system: myths and facts. *Human Vaccines*, **2008**, *4*(1), 5-12.
- [7] Fifis, T.; Gamvrellis, A.; Crimeen-Irwin, B.; Pietersz, G.A.; Li, J.; Mottram, P.L.; McKenzie, I.F.; Plebanski, M. Size-dependent immunogenicity: therapeutic and protective properties of nano-vaccines against tumors. *J. Immunol.*, **2004**, *173*(5), 3148-3154.
- [8] Leclerc, D.; Beauseigle, D.; Denis, J.; Morin, H.; Paré, C.; Lamarre, A.; Lapointe, R. Proteasome-independent major histocompatibility complex class I cross-presentation mediated by papaya mosaic virus-like particles leads to expansion of specific human T cells. *J. Virol.*, **2007**, *81*(3), 1319-1326.
- [9] Thérien, A.; Bédard, M.; Carignan, D.; Rioux, G.; Gauthier-Landry, L.; Laliberté-Gagné, M.É.; Bolduc, M.; Savard, P.; Leclerc, D. A versatile papaya mosaic virus (PapMV) vaccine platform based on sortase-mediated antigen coupling. *J. Nanobiotechnology*, **2017**, *15*(1), 54.
- [10] Saadi, M.; Karkhah, A.; Nouri, H.R. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infect. Genet. Evol.*, **2017**, *51*, 227-234.
- [11] Ikegami, R. *Papaya mosaic potexvirus as an expression vector for foreign peptides*. M.Sc. Thesis, University of Toronto: Canada, 1996.
- [12] Ponomarenko, J.; Bui, H.-H.; Li, W.; Fusseder, N.; Bourne, P.E.; Sette, A.; Peters, B. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, **2008**, *9*(1), 514.
- [13] Doytchinova, I.A.; Guan, P.; Flower, D.R. EpiJen: a server for multistep T cell epitope prediction. *BMC Bioinformatics*, **2006**, *7*(1), 131.
- [14] Singh, H.; Raghava, G. ProPred: prediction of HLA-DR binding sites. *Bioinformatics*, **2001**, *17*(12), 1236-1237.
- [15] McGuffin, L.J.; Bryson, K.; Jones, D.T. The PSIPRED protein structure prediction server. *Bioinformatics*, **2000**, *16*(4), 404-405.
- [16] Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **2008**, *9*(1), 40.
- [17] Tollin, P.; Bancroft, J.B.; Richardson, J.F.; Payne, N.C.; Beveridge, T.J. Diffraction studies of papaya mosaic virus. *Virology*, **1979**, *98*(1), 108-115.
- [18] Rioux, G.; Majeau, N.; Leclerc, D. Mapping the surface-exposed regions of papaya mosaic virus nanoparticles. *FEBS J.*, **2012**, *279*(11), 2004-2011.
- [19] Dosztányi, Z. Prediction of protein disorder based on IUPred. *Protein Sci.*, **2018**, *27*(1), 331-340.
- [20] Saha, S.; Raghava, G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.*, **2006**, *34*(Suppl_2), W202-W209.
- [21] Soria-Guerra, R.E.; Nieto-Gomez, R.; Govea-Alonso, D.O.; Rosales-Mendoza, S. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *J. Biomed. Informatics*, **2015**, *53*, 405-414.
- [22] Liljeroos, L.; Malito, E.; Ferlenghi, I.; Bottomley, M.J. Structural and computational biology in the design of immunogenic vaccine antigens. *J. Immunol Res.*, **2015**, *2015*, 156241.
- [23] Kulp, D.W.; Schief, W.R. Advances in structure-based vaccine design. *Curr. Opin. Virol.*, **2013**, *3*(3), 322-331.
- [24] Carignan, D.; Thérien, A.; Rioux, G.; Paquet, G.; Gagné, M.É.L.; Bolduc, M.; Savard, P.; Leclerc, D. Engineering of the PapMV vaccine platform with a shortened M2e peptide leads to an effective one dose influenza vaccine. *Vaccine*, **2015**, *33*(51), 7245-7253.
- [25] Arcangeli, C.; Circelli, P.; Donini, M.; Aljabali, A.A.; Benvenuto, E.; Lomonosoff, G.P.; Marusic, C. Structure-based design and experimental engineering of a plant virus nanoparticle for the presentation of immunogenic epitopes and as a drug carrier. *J. Biomol. Struct. Dyn.*, **2014**, *32*(4), 630-647.
- [26] Yang, S.; Wang, T.; Bohon, J.; Gagné, M.É.L.; Bolduc, M.; Leclerc, D.; Li, H. Crystal structure of the coat protein of the flexible filamentous papaya mosaic virus. *J. Mol. Biol.*, **2012**, *422*(2), 263-273.
- [27] Cozzetto, D.; Jones, D.T. The contribution of intrinsic disorder prediction to the elucidation of protein function. *Curr. Opin. Struct. Biol.*, **2013**, *23*(3), 467-472.
- [28] De Lalla, C.; Sturniolo, T.; Abbruzzese, L.; Hammer, J.; Sidoli, A.; Sinigaglia, F.; Panina-Bordignon, P. Cutting edge: identification of novel T cell epitopes in Lol p5a by computational prediction. *J.*

- Immunol.*, **1999**, 163(4), 1725-1729.
- [29] Koudelka, K.J.; Destito, G.; Plummer, E.M.; Trauger, S.A.; Siuzdak, G.; Manchester, M. Endothelial targeting of cowpea mosaic virus (CPMV) via surface vimentin. *PLoS Pathogens*, **2012**, 5(5), e1000417.
- [30] Wang, C.; Steinmetz, N.F. CD47 blockade and cowpea mosaic virus nanoparticle *in situ* vaccination triggers phagocytosis and tumor killing. *Adv. Healthcare Mater.*, **2019**, Available from: https://www.researchgate.net/publication/331551015_CD47_Blockade_and_Cowpea_Mosaic_Virus_Nanoparticle_In_Situ_Vaccination_Triggers_Phagocytosis_and_Tumor_Killing.