Article

Hot-spots and their contribution to the self-assembly of the viral capsid: *in-vitro* validation

Alejandra Gabriela Valdez-Lara [1], Mariana Andrade-Medina [1], Josué Alejandro Alemán-Vilis [1], Aldo Adrián Pérez-Montoya [1], Nayely Pineda-Aguilar [2], Eduardo Martínez-Guerra [2], Paul Gaytán [3] and Mauricio Carrillo-Tripp [1,*]

[1] Biomolecular Diversity Laboratory, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Unidad Monterrey, Vía del Conocimiento 201, Parque PIIT, C.P. 66600, Apodaca, Nuevo León, México

[2] Centro de Investigación en Materiales Avanzados Unidad Monterrey, Parque PIIT, C.P. 66600, Apodaca, Nuevo León, México

[3] Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Col. Chamilpa, Cuernavaca, Morelos 62210, México

* Correspondence: mauricio.carrillo@cinvestav.mx (M.C.-T.)

Received: date; Accepted: date; Published: date

Abstract: The viral capsid is a macromolecular complex formed by self-assembled proteins which, in many cases, are biopolymers with an identical amino acid sequence. Specific protein-protein interactions drive the capsid self-assembly process. However, it is believed that only a small set of interface residues significantly contribute to the formation of the capsid, the so-called "hot-spots". Here, we investigate the effect of point-mutations on previously predicted hot-spots of the icosahedral Cowpea Chlorotic Mottle Virus. We characterize the formation and stability of virus-like particles *in-vitro* by thermal shift assays. Our results show that a single mutation on residue E176 or V189 significantly perturbs the quaternary protein-protein interaction, preventing the formation of the capsid. Furthermore, these mutations do not destroy the tertiary fold of the capsid protein. Our findings give evidence of the *in-silico* hot-spot prediction accuracy. As a whole, our methodology provides a guide to the rational development of molecules that could inhibit virus formation.

Keywords: molecular biophysics, molecular biology, hot-spot prediction, ccmv

1. Introduction

Two essential molecular components constitute a virus particle. The genetic material in the form of a single or double-stranded RNA or DNA molecule encodes all the genes necessary for the infection. Second, a sophisticated protein shell, known as the viral capsid, whose function is to transport and protect the genetic material from the harsh environment. The capsid is assembled from multiple copies of one or a few chemically distinct capsid proteins (CPs) which pack the viral genome and relevant accessory proteins *in-vivo*. The CP-CP interactions are critical for the formation, stability, and integrity of the viral capsid. The extent of interactions and the type of resulting capsids are encoded in the CP as they yield uniform capsids in geometry and size. Hence, analysis of protein-protein interactions in viral capsids within and across members of virus families will shed light on the fundamental principles of the molecular mechanisms involved in the process of protein self-assembly.

Some viral CPs spontaneously form empty capsids when expressed in heterologous expression systems and assembled in the right conditions *in-vitro* [1–4]. These empty capsids are known as virus-like particles (VLPs). In recent years, VLPs have gained attention because of their potential use in biotechnology and medicine. For example, the inhibition of protein-protein interactions (PPI) is now a reality for an increasing number of cases [5]. This strategy has seen solid advances toward near-future therapeutic applications. Currently, one approach in this direction is the search, or rational design, of small molecules or peptides that inhibit specific PPIs. However, there are several problems involved in this process. One of them is that, as opposed to an enzyme's active site, large interfaces stabilize PPIs.

Nonetheless, it is reasonable to think that a small molecule interacting at the position of the so-called "hot-spots" will compete with the binding of the associated subunit without necessarily covering the whole surface of interaction. Hence, knowing the accurate location and physicochemical nature of the hot-spots is key to develop inhibitory strategies that target PPIs.

According to the hot-spot hypothesis, from all the residues forming the protein-protein interface, only a few of them contribute significantly to the thermodynamic stability of the complex. Following our structural-conservation-based definition, the identification of such residues in viral capsids can be obtained in three steps. After gathering the capsid structural data from members of a given family from VIPERdb [6], i) select all the interface residues in each member, ii) identify residues conserved in sequence through a multiple alignment, and iii) identify residues conserved in quaternary structure using CapsidMaps [7]. The intersection of these three sets of residues are the hot-spots of that family.

In this work, we conduct an experimental verification of our hot-spot prediction for the icosahedral Cowpea Chlorotic Mottle Virus (CCMV), a member of the *Bromoviridae* family. We report the details of the conservation-based prediction pipeline and *in-silico* validation results in a companion article. In particular, we identified six hot-spots. Two of them, residues P99 and F120 belong to an interface related by the 3- and 5-fold symmetry axes. Residues E176, R179, P188, and V189 belong to the interface related by a 2-fold symmetry axis. Here, we investigate the effect on the thermal stability of the CCMV capsid when we introduce mutations at the hot-spot locations of the 2-fold related interface. We show that some of those mutations prevent the formation of the capsid while maintaining the CP tertiary fold.

2. Results

We made point mutations to the CCMV CP in four hot-spots (E176, R179, P188, and V189) and two randomly-selected non-conserved interface residues (experimental controls E77 and V133). All of them belong to the interface formed around a 2-fold symmetry axis. Figure 1 shows their location in the tertiary structure in one of the 180 CPs forming the capsid and all the interface residues in a 2-fold related dimer interface.

2.1. CCMV CPs are readily produced in a heterologous system

We corroborated correct in-frame cloning of all variants by sequencing the pET19b plasmid. Then, we ran Agarose gel electrophoresis resulting from subsequent PCR amplification of the CP wild-type and the six point-mutation genes using specific forward and reverse primers which amplified a 577 bp final product. After cell transformation, CCMV CP variants were individually expressed in *E. coli*. We verified the presence of the protein in each fraction of the purification process by monitoring its migration behavior in an SDS-PAGE gel at 12%. Most of the recombinant protein was retained after passing the bacterial clarified by the affinity column. The CP+His-TAG was eluted in 1.5 mL fractions after washing with small amounts of imidazole. Thus, other residual *E. coli* proteins that got trapped in the column were eluted. The CP wild-type (WT) has a molecular weight of ≈20 kDa [3] and a molecular extinction coefficient of 23,590 M^{-1} cm^{-1} [8]. However, the CP+His-TAG variants migrate to a molecular weight of ≈24 kDa in the SDS-PAGE gel due to the inclusion of the histidine tag and all residues corresponding to the recognition site of the PreScission Protease® restriction enzyme. We have previously shown that the CP without the histidine tag migrates to the correct position. Furthermore, even with the histidine tag, CCMV VLPs are formed, i. e., the tag does not affect the self-assembly process [4]. The final concentration of the purified CP was ≈2 mg/mL (determined by Bradford protein assay) per 1 L of induced cell culture. We obtained all CCMV CP variants with high purity levels in most eluates, i.e., no other contaminating protein is detected. As a final verification, Western blotting with antibody against THE^{TM} His Tag Antibody was carried out for all variants. We summarize these results in Figure 2.

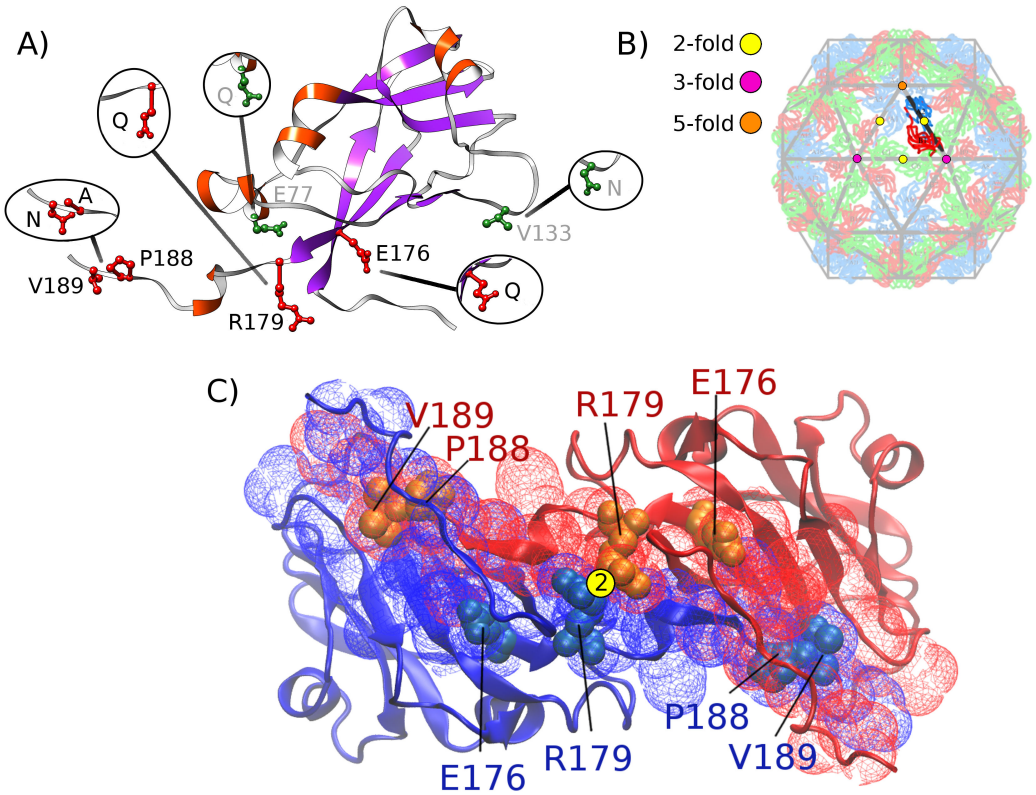

Figure 1. CCMV CP point mutations and dimer interface. (**a**) The 2-fold related hot-spots (red side-chains) and two non-conserved interface residues (green side-chains). Their location in the tertiary structure of the CP is shown indicating the corresponding mutations made. The secondary structure is color-coded (α-helix in orange, β-sheet in purple. (**b**) Capsid quaternary structure (VIPERdb ID 1cwp), showing the location of the symmetry folds on the icosahedral asymmetric unit. A 2-fold related dimer is highlighted (A2 in blue and B1 in red). (**c**) The same dimer, showing all the residues forming the large 2-fold related interface (mesh) and the location of the four hot-spots per subunit (spheres).

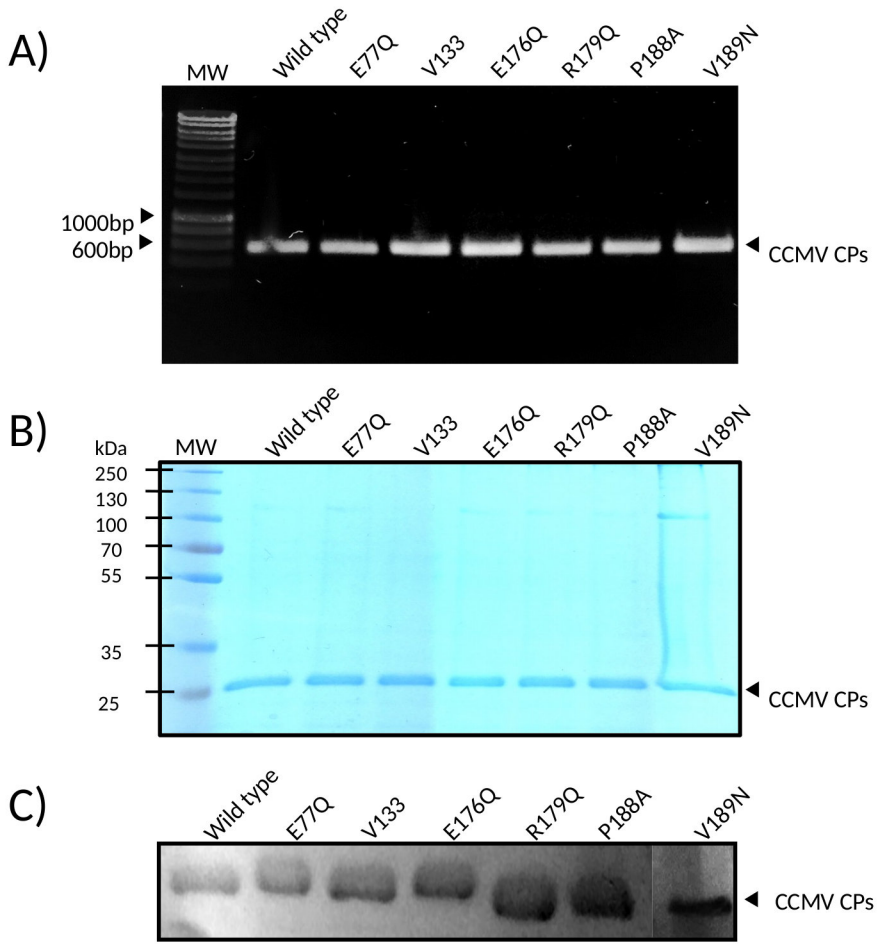

Figure 2. CP gene amplification, protein separation, and verification by western blot. (**a**) Agarose gel electrophoresis of seven PCR products resulting from subsequent amplification of the CP wild-type and six point-mutation genes using specific forward and reverse primers, which amplified a 572 bp final product. (**b**) Purification of CCMV CP wild-type and six point-mutation followed by SDS-PAGE analysis. Visualized with Coomassie blue staining. (**c**) Western blotting with antibody against THE^{TM} His Tag Antibody, mAb, Mouse (subtype IgG1).

2.2. CCMV VLPs self-assemble in pH=4.8

The correct formation of virus-like particles of wild-type CCMV was confirmed by tungsten negative-stain on a scanning electron microscope (SEM). After expression and purification, we placed the protein in assembly buffer (pH=4.8) on a copper grid with a formvar/carbon-coated mesh. We found a high density of VLPs having the characteristic 30 nm diameter (Figure 3).

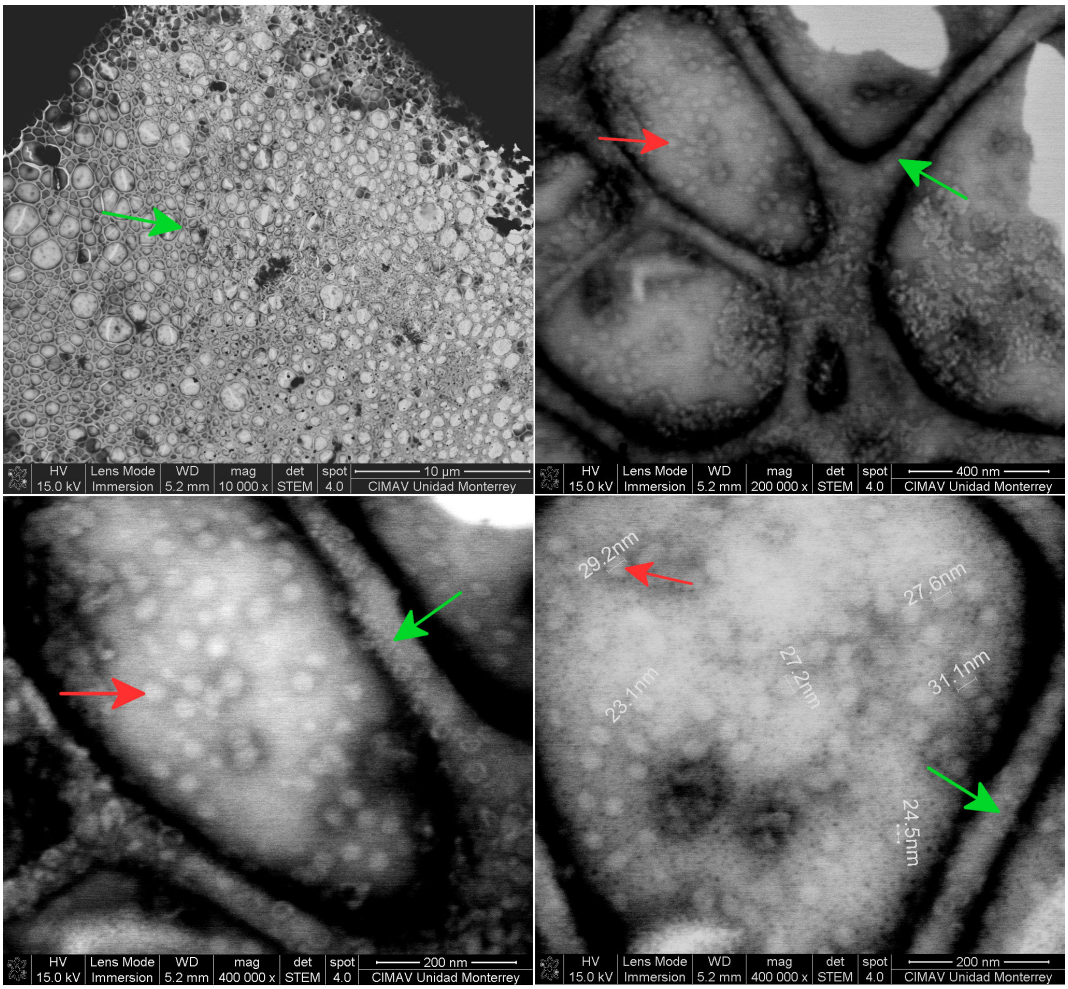

Figure 3. Virus-like particles of wild-type CCMV saw on a scanning electron microscope (SEM) by negative-stain. The sample was placed on a 300 copper grid with a formvar/carbon-coated mesh (green arrows). A high density of VLPs with a characteristic diameter of ≈30 nm are visible at magnifications of 200,000x and 400,000x (red arrows).

2.3. Hot-spots significantly contribute to the VLP self-assembly

We used the TSA-Tm tool to analyze the raw data produced by the thermal shift assays. If the analyzed protein does not form oligomers of any kind, we will only find one distinctive peak on the first derivative of the fluorescence vs. temperature profile. On the other hand, when a protein self-assembles into oligomers, VLPs in this case, we should find a peak corresponding to the disassembly, Tm_1, and a second peak corresponding to the thermal denaturation of the protein, Tm_2. We show an illustration of the second case in Figure 4.

In order to characterize the reliability of the TSA methodology, we compared it with previously reported results by differential scanning calorimetry (DSC) of Lysozyme [9]. We conducted a

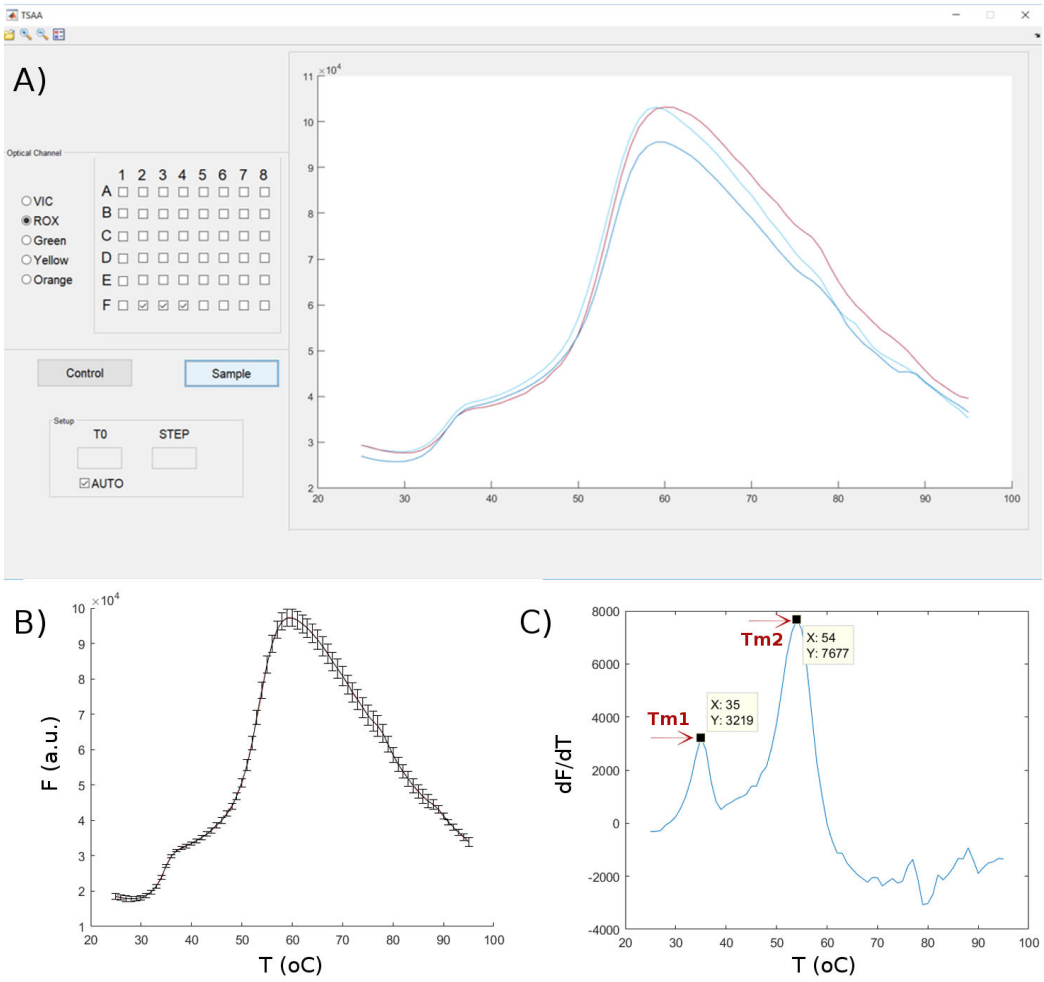

Figure 4. TSA-Tm analysis tool. A) The user can select the optical channel and the set of wells to analyze and visualize the results through a graphical interface. Then, B) subtraction of the blank (control) and averaging of repetitions, as well as C) calculation of dF/dT is automatic. The melting temperature values, Tm_i, correspond to the peaks found in the first derivative of the fluorescence vs. temperature profile.

thorough analysis in a pH range of 1 to 12 and measured the Lysozyme denaturation temperature behavior (Figure 5). We describe the specific conditions used for each pH value in Table 1. We fitted a third-order orthogonal polynomial model to the TSA data. The coefficients found were $Y = 34.1464646 + 20.75628X - 3.17248X^2 + 0.13604X^3$, with an R^2=0.9251 and all the coefficient terms being statistically significant with a *p-value*< 0.001. Both DSC and TSA Tm_2 values for Lysozyme fall within a 95% confidence interval of the model, demonstrating that both methods can detect the same trends in protein thermostability.

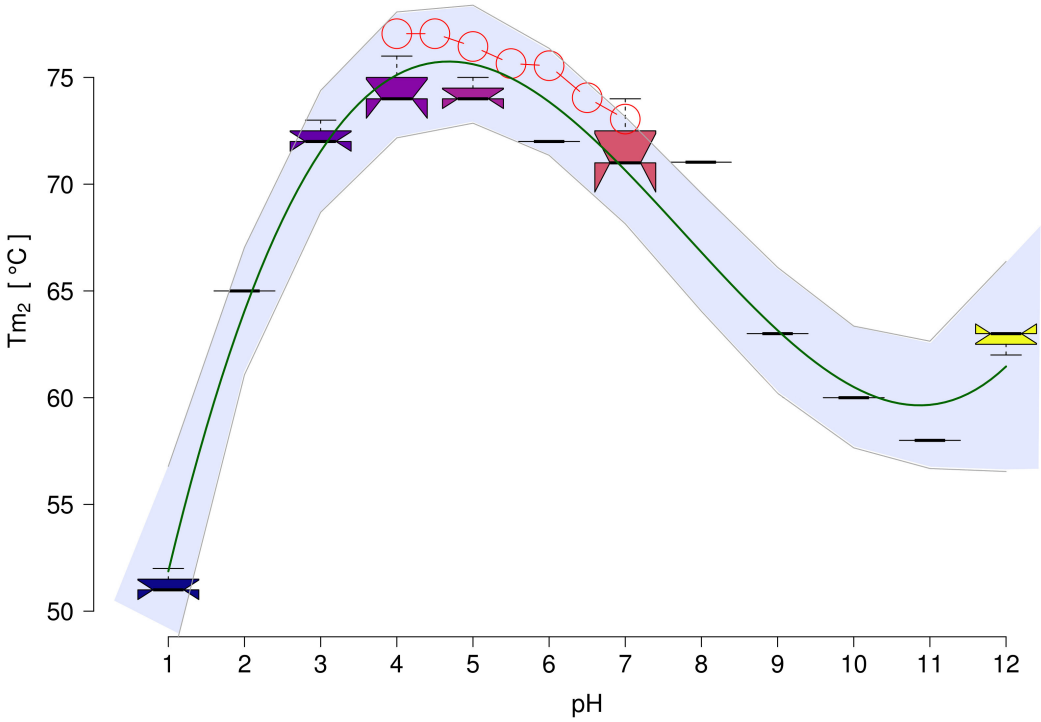

Figure 5. Denaturation temperature, Tm_2, of Lysozyme as a function of pH value (boxplots). Previously reported values measured by differential scanning calorimetry (DSC) are shown with red circles [9]. A third-order orthogonal polynomial model was fitted to the TSA data (green line; R^2=0.9251 with all the coefficient terms being statistically significant with a *p-value*< 0.001). The shaded region indicates 95% confidence intervals.

We conducted a series of thermal shift assays to characterize the CCMV VLP self-assembly and the thermal stability of the CP. We included two different conditions in the experimental design; VLP assembly at pH=4.8 and disassembly at pH=7.5. We analyzed seven CP variants; the wild-type (WT), two non-conserved interface residues (control group), and the four hot-spots. We found melting temperatures for capsid disassembly, Tm_1, and CP denaturation, Tm_2. Every assay included an independent measurement of Lysozyme as a positive control. A large number of repetitions in several cases allowed for useful statistics. Most results were significant to the WT, summarized in Figure 6 and Table 2.

As expected, none of the seven variants formed VLPs at pH=7.5. However, mutations do affect the stability of the tertiary structure of the CP. Of the control group, mutant E77Q stabilizes, and mutant V133N destabilizes the fold of the protein to some degree. Hot-spot mutants E176Q and R179Q increase the stability, mutant P188 does not cause a detectable effect although its experimental variability increases, and mutant V189N decreases the thermal stability of the CP.

Lowering the pH value from 7.5 to 4.8 stabilizes the tertiary structure, as seen in the case of Lysozyme. This observation is true in general, except for hot-spot mutant E176Q, which has a large experimental Tm_2 variability (probably due to a small statistical sampling). The value of pH=4.8 might

* All solutions prepared in 25 mL of water

pH	buffer	preparation *
1	KCl/HCl	0.93g/0.77mL
2	KCl/HCl	0.186g/0.043mL
3	Gly/HCl	0.18g/0.048mL
4	AcONa/AcOH	0.225g/0.54mL
5	AcONa/AcOH	0.7g/ 0.21mL
6	Na_2HPO_4/NaH_2PO_4	0.46g/1.48g
7	Na_2HPO_4/NaH_2PO_4	1.94g/0.73g
8	TrizmaHCl/TrizmaBase	1.081g/0.693g
9	$NaCO_3/NaHCO_3$	0.115g/0.917g
10	$NaCO_3/NaHCO_3$	0.143g/0.097g
11	$NaHCO_3$[0.05M]/NaOH[0.1M]	16.6mL/7.56mL
12	KCl[0.2M]/NaOH[0.2M]	20mL/4.8mL

Table 1. Buffer preparation to measure protein melting temperatures by the thermal shift assay method in a wide range of pH values.

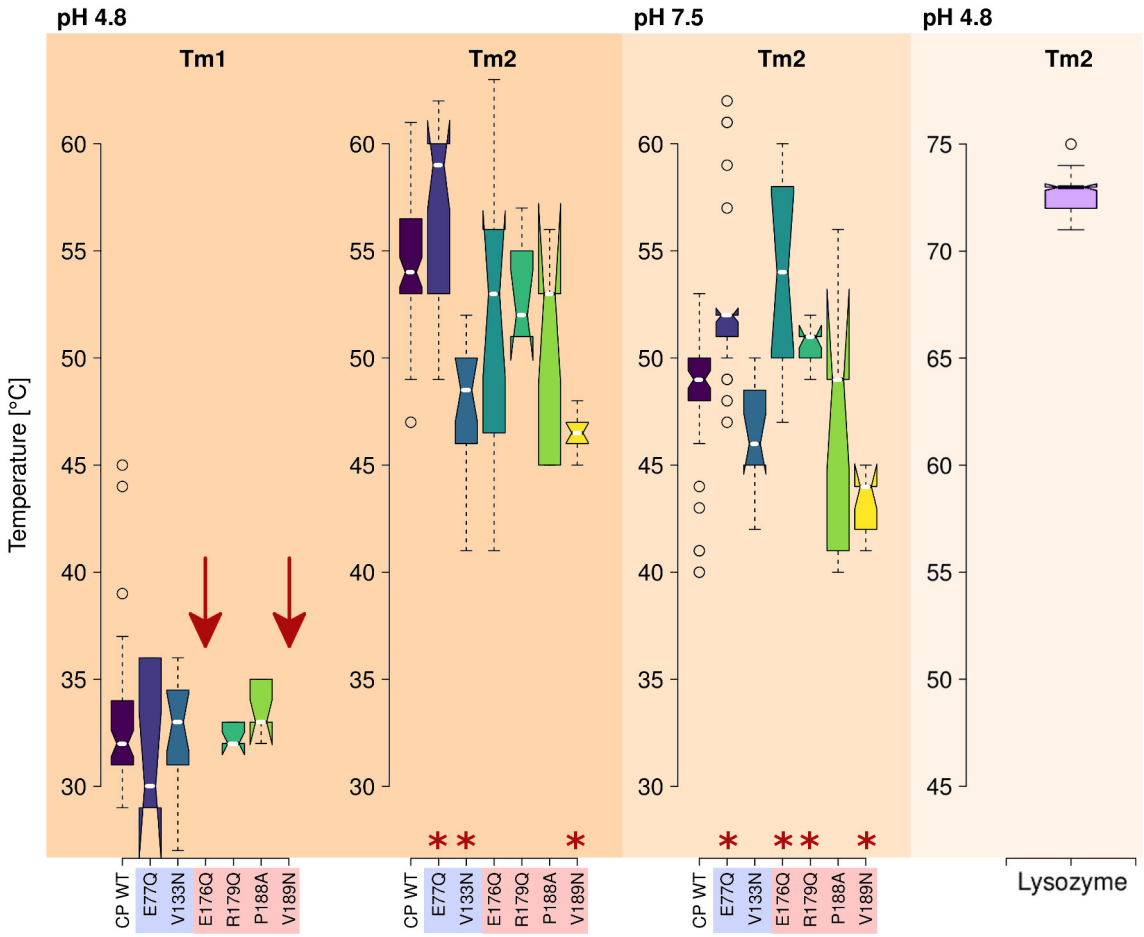

Figure 6. VLP self-assembly and thermal stability of the CCMV CP. Statistical analysis (boxplots) of two different conditions (capsid assembly at pH=4.8 and disassembly at pH=7.5) for the wild-type (WT), two non-conserved interface residues, and the four hot-spots on the CP. Melting temperatures were found for capsid disassembly (Tm_1) and CP denaturation (Tm_2). In the case of hot-spot mutations E176Q and V189N, no VLPs were formed (red arrows). All TSA experiments included the Lysozyme as a positive control. * *p-value*< 0.05 on a t-test with respect to WT.

[a] Control group
[b] Hot-spots
[*] t-test significant to wild-type

	pH 4.8						pH 7.5		
	Tm_1 [°C]			Tm_2 [°C]			Tm_2 [°C]		
Variant	n	avg±sdv	p-value	n	avg±sdv	p-value	n	avg±sdv	p-value
Wild-type	61	33.0±3.1		64	54.4±3.0		61	48.5±2.8	
E077Q[a]	11	31.5±3.0	0.234	28	57.0±4.1	*0.004	23	52.7±4.2	*1.2E-4
V133N[a]	16	32.2±2.9	0.359	18	47.9±3.0	*1.4E-8	15	46.3±2.8	0.014
E176Q[b]	15	No VLP	na	15	51.8±6.5	0.154	12	53.9±4.6	*0.002
R179Q[b]	9	32.3±0.5	0.156	9	52.8±2.3	0.083	9	50.7±0.9	*3.5E-5
P188A[b]	9	33.4±1.2	0.396	9	50.7±4.4	0.035	9	46.9±5.4	0.398
V189N[b]	12	No VLP	na	12	46.5±1.0	*2.2E-16	9	43.3±1.3	*1.0E-8

Table 2. Thermal shift assay melting temperature (Tm_i) values. Two different pH values were tested; 4.8 VLP self-assembly conditions and 7.5 VLP disassembly conditions. Tm_1 is the VLP melting temperature, and Tm_2 is the protein unfolding temperature.

be close to where the protein fold is the most stable. Some patterns remain from the higher pH value. Of the control group, mutant E77Q stabilizes, and mutant V133N destabilizes the fold of the protein on a higher degree than pH=7.5. On the other hand, hot-spot mutants E176Q, R179Q, and P188A do not affect the thermal stability (95% confidence intervals fall within the Tm_2 value of the wild-type). In the case of hot-spot mutant V189N, there is a significant destabilization of the tertiary CP structure.

Based on the consistent detection of a peak corresponding to Tm_1, we found the formation of VLPs at pH=4.8 in the case of the WT (as we had previously reported [4]), the two mutants in the control group, and the two hot-spot mutants R179Q and P188A. None of these four mutants Tm_1 significantly differ from the characteristic VLP disassembly temperature of WT CCMV (95% confidence intervals fall within the Tm_1 value of the WT). Therefore, the quaternary structure and thermal stability of the VLP is not affected by these mutations. However, hot-spot mutants E176Q and V189N prevented the formation of VLPs without disrupting the native fold of the CP.

3. Discussion

In this study, we provide experimental evidence for the validity and prediction power of the structural-conservation-based criteria for theoretical hot-spot identification. Our findings show that a single interface residue can play a significant role in the protein self-assembly mechanism, supporting the general hot-spot hypothesis [10]. Noteworthy, the hot-spots we identified in the CCMV 2-fold related interfaces are significantly relevant only at the quaternary level since the mutation to remove the physicochemical contribution to the intra- and inter-molecular interactions of those residues do not destroy the tertiary protein fold but do inhibit VLP formation. This property does not hold for non-conserved interface residues, even when they have seemingly equivalent physicochemical contributions as the hot-spots. Hence, our *in-vitro* results are concomitant with the *in-silico* free energy study reported in our companion work.

By our definition, hot-spots are located in the inter-subunit interfaces and are conserved both in sequence and quaternary structure within all members of a virus family. In the case of the CCMV, member of the *Bromoviridae* family, we predicted six hot-spots. Two of them are located in 3- and 5-fold related interfaces. The other four were studied here. It was a reasonable expectation that the point-mutation on any of them would either prevent the formation of VLPs or significantly decrease the thermal stability of the particle. However, mutations on hot-spots R179 and P188 showed no distinction from the WT, even though that our *in-silico* calculations predicted a similar contribution to the dimer binding free energy as that from hot-spot V189. When we compare the interactions involved with each one of the hot-spots (Table 3), we do not find any pattern that could help us explain this.

[a] Control group
[b] Hot-spots
[c] Statistics for the whole 2-fold related interface are: min=3.7, max=6.8, avg=5.3, stdev=0.8

M_1	M_2	com-com[c]	Interaction type	
GLU 176[b]	PHE 186	5.4	Acid	Non-Polar
ARG 179[b]	ARG 179[b]	5.3	Basic	Basic
	ASP 183	4.7		Acid
	ASP 184	6.4		Acid
PRO 188[b]	GLY 41	5.2	Non-Polar	Non-Polar
	ALA 43	4.6		Non-Polar
VAL 189[b]	LEU 92	6.3	Non-Polar	Non-Polar
	VAL 133[a]	4.9		Non-Polar
	ALA 135	4.5		Non-Polar
GLU 77[a]	GLU 77[a]	4.1	Acid	Acid
VAL 133[a]	VAL 189[b]	4.9	Non-Polar	Non-Polar
	TYR 190	6.7		Polar

Table 3. Contact table for the 2-fold related hot-spots and two non-conserved interface residues studied in this work. M_1-M_2 represents both A_2-B_1 and B_1-A_2 because of the homodimer symmetry. The distance between the center of mass (com) between the residues in the pair is given in Å along with the physicochemical interaction type.

Hot-spot E176 has only one contact with no apparent important interaction type. We estimated that it is the hot-spot which contributes the most (75% of the binding free energy). Congruent with this finding, mutant E176Q does not form VLPs. A similar mutation, E77Q, does not affect the quaternary structure. The other three hot-spots contribute less but in a similar amount to the stability of the dimer (45% of the binding free energy). However, only the mutant V189N prevents VLP formation. V189 is involved in three hydrophobic contacts. A change from a non-polar to a polar environment will perturb the interface. This effect seems to be significant only in the hot-spot location. A similar mutation, V133N, does not affect the VLP quaternary structure nor thermal stability. On the other hand, mutation R179Q eliminates one repulsive and two attractive electrostatic interactions in the interface. Mutation P188A maintains a hydrophobic environment but increases the com-com distance of two interactions in the interface, which should decrease the stability on the dimer. Nevertheless, these two mutations behave like the control group and the WT. Hence, if not relevant for the formation and stability of the capsid, there must be another evolutionary factor that has fixed them.

Regarding our experimental methodology, the thermal shift assay is a powerful tool to predict protein stabilizing conditions [11,12], optimization of protein purification [13], interacting ligands for drug discovery [14,15], and biochemical properties of capsid proteins [16]. Our results on the wide-range pH dependence of the denaturation temperature of Lysozyme by TSA showed an excellent agreement with DSC results. This agreement is in line with previous reports that have shown a good correlation between the two techniques [17]. The fact that the TSA results are consistently lower than those from DSC was explained by the possibility that the dye used in TSA could be sensitive to temperature changes which locally affect the surface hydrophobicity and protein-solvent interactions, a factor that does not exist in DSC. Despite this, it is now accepted that the TSA method is very reproducible, repeatable, and accurate; significant changes of a few degrees in Tm_i truly reflect a change on the investigated system.

We have shown in this and a companion work that the CCMV residues E176 and V189 largely contribute to the binding affinity of the 2-fold related homodimer. This CP complex is the first step in the capsid self-assembly process [18]. Our mutagenesis results on thermal stability show that the conservation-based hot-spot predictions have indeed a significant contribution to the formation and stability of the protein complex. In view of our findings, a similar study of the thermodynamic

contribution of hot-spots P99 and F120 is in progress. Most likely, their role will be in the formation of intermediate states in the process of capsid assembly, e. g., PODs [19].

4. Materials and Methods

4.1. CCMV CPs point mutation construction

We cloned the nucleotide sequence of the native *cp* gene from CCMV into the pET19b plasmid [4]. Following the adapted PCR protocol previously reported, the amplification of the plasmid template containing the native *cp* gene was run in two separate PCR reaction. These reactions contained 1X High Fidelity buffer, 0.5 μM dNTPs, 0.2 μM of either forward or reverse primer, 10 ng/μL of template plasmid, and 0.5 μL of Phusion DNA polymerase (BioLabs ®) in a 25 μL reaction. The PCR reaction was performed as follows: 1 cycle at 98 °C for 1:30 min, 10 cycles at 98 °C for 15 s, 60.4 °C for 30 s, 72 °C for 3:30 min, and 1 cycle at 72 °C for 5 min. After the first ten cycles, the two single-primer PCR products were combined in one test tube (giving a total volume of 50 μL), and then the amplification was continued with 20 more cycles with the same conditions described above. The reaction was supplemented with 20 units of DpnI (BioLabs ®) and incubated overnight at 37 °C. The restriction enzyme DpnI digests the methylated template plasmid strands, leaving the newly synthesized strands intact. This product was purified and used to transform chemically competent *E. coli* DH5α cells. Plasmids were isolated from the resulting colonies and screened for the desired modification. We visualized PCR products and DNA plasmids by 1 % agarose (SIGMA Cat.#A9539) gel dyed with GelRed (BIOTIUM Cat.#41003). Finally, the positive clones were sequenced to confirm the desired modification and the absence of additional modifications.

4.2. CCMV CPs expression in E. coli

In order to express the CCMV CPs chemically competent cells of *E. coli*, Rosetta 2(DE3) pLysS ® (Novagen, Germany) were transformed with the CCMV point mutation CPs previously generated from native *cp* gene cloned into plasmid pET19b. This strain is resistant to chloramphenicol, and the plasmid pET19 confers resistance to ampicillin, carbenicillin, and related antibiotics. The transformed cells were grown in LB medium agar plate supplemented with carbenicillin (100 μg/mL) and chloramphenicol (34 μg/mL). One colony was selected, and a pre-inoculum was prepared in 100 mL of liquid LB medium supplemented with carbenicillin (100 μg/mL) and chloramphenicol (34 μg/mL). The culture was grown at 37 °C for 16 h and 200 rpm. This overnight culture was transferred to 1 L of fresh LB liquid medium, supplemented with carbenicillin and chloramphenicol, at the same concentration as before, and incubated at 37 °C and 200 rpm of agitation up to an optical density of 0.7 absorbance units. Finally, recombinant expression of the CCMV CP was induced with the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG), at a final concentration of 0.75 mM and incubation for 16 h at 25 °C and 200 rpm. Induced cells were collected by centrifugation at 3,000 $\times$ g at 4 °C for 30 min. The bacterial pellet was stored at -80 °C.

4.3. CCMV CPs purification by affinity chromatography

The pellet was re-suspended in 60 mL of lysis buffer (100 mM Tris-HCl, 500 mM NaCl, pH=6.0) supplemented with lysozyme (1 mg/mL final concentration) and phenylmethylsulfonyl fluoride (PMSF) to a final concentration of 1 mM, and incubated for 1 h at 4 °C with constant stirring. The solution was sonicated for 10 min, using 20 s pulses, at an intensity of 20% (QSONICA model Q500) with a standard 1/2 in diameter probe tip (#4220). Centrifugation at 13,500 x g was done for 30 min at 4 °C. The pellet was discarded, and the soluble fraction was filtered through a 0.45 μm membrane and further loaded into a Pierce ® Cobalt affinity column (HisPur^{TM} Cobalt Chromatography Cartridge) previously equilibrated with ten volumes of lysis buffer. Consecutive washes were made using three different buffers varying the imidazole concentration (100 mM Tris-HCl, 500 mM NaCl and L1 = 15 mM imidazole, pH=7.0; L2 = 20 mM imidazole pH=7.2; L3 = 35 mM imidazole, pH=7.3). Each wash was

done with 20 volumes of the Cobalt affinity column. Recombinant CP was eluted using elution buffer (100 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, pH=8.0) in 2 mL fraction. Half of the recombinant protein was dialyzed for 16 h at 4 °C in 1 L of dialysis buffer (100 mM Tris-HCl, 500 mM NaCl, pH=7.5). To effectively retain the CP, a dialysis membrane of 14 kDa MWCO (MEMBRA-CEL ® MD44) was used. The concentration of the dialyzed protein was estimated using the Quick Start®Bradford Dye Reagent 1X (BIO-RAD Cat. #500-0205). The protein was stored at 4 °C before use.

4.4. Protein separation and CCMV CPs molecular weight signature

We verified the identity of the CCMV CPs by molecular weight characterization. Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis assays were run for each of the different mutant CPs. Samples were mixed 4:1 with Laemmli 4x buffer (250 mM Tris-HCl pH=6.8, 8% SDS, 0.2% bromophenol blue, 40% glycerol and 20% 2-Mercaptoethanol). Total protein was denatured by heating at 97 °C for 5 min, cooled on ice for 5 min and loaded on the polyacrylamide gel at 12%. Each gel was run at 130 V for 1.5 h. Gels were stained using standard Coomassie Blue.

4.5. Verification of the CCMV CPs by Western Blot

Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis was performed as described before. Proteins were transferred to Immun-Blot ® PVDF membrane for protein Blotting (BIO-RAD Cat. #162 − 0177). Briefly, a "sandwich" was prepared with the following successive layers: i) three thicknesses of Mini Trans-Blot ® filter paper (BIO-RAD Cat. #1703932) 7.5 x 10 cm; ii) the SDS-polyacrylamide slab gel (with the stacking gel removed); iii) Immun-Blot ® PVDF membrane; iv) three more thicknesses of filter paper. All components in contact with the slab gel were previously wet in the transfer solution (10 mM Tris-base, 50 mM Glycine, 10% methanol). The blot was run using a Semidry Electroblotter from Thermo Fisher Scientific (Model HEP-1) at 15 V for 1 h using a BIO-RAD PowerPac ® HC power supply. Immediately after the transfer, the PVDF membrane was immersed in a wash buffer solution (100 mM Tris-HCl, 0.9% NaCl, 0.1% Tween20 pH=7.5) and incubated at 25 °C for 15 min, then it was briefly rinsed with two changes of wash buffer. The membrane was blocked for non-specific binding sites by immersion in Block buffer (100 mM Tris-HCl, 0.9% NaCl, 3% non-fat dry milk, pH=7.5) at 25 °C for 1 h. The recombinant protein was detected using the primary antibody THE^{TM} His Tag Antibody, mAb, Mouse (subtype IgG1) GenScript (Cat. #A00186) diluted 1:3000 in buffer (100 mM Tris-HCl, 0.9% NaCl, 0.05% Tween20, 0.5% non-fat dry milk, pH=7.5). The membrane was incubated at 4 °C for 16 h, washed and incubated again with HRP labeled secondary antibody diluted at the same dilution factor and buffer as the primary antibody. The detection was made using SuperSignal®West Dura Extended Duration Substrate (Thermo Scientific Cat. #34075) according to the manufacturer's instructions. The blot was photographed 5 min after the substrate was applied using the Bio-Imaging System (MicroChemi Cat. #70 − 25 − 00).

4.6. in-vitro *assembly of CCMV virus-like particles*

The self-assembly process of the different CPs was induced in aliquots of purified protein placed separately in a dialysis membrane of 14 kDa MWCO (MEMBRA-CEL®MD44) and dialyzed simultaneously at 4 °C for 16 h against 1 L of assembly buffer (0.1 M Sodium acetate buffer, 500 mM NaCl, pH=4.8).

4.7. Electron Microscopy

We confirmed VLP formation by direct observation in a Scanning Electron Microscope (SEM). Samples were taken after the assembly reaction and diluted to 0.15mg/mL. 3 μL of the diluted sample was placed on hydrophilic formvar coated mesh copper grids and incubated for 10min at room temperature to allow for VLPs adherence. Excess sample was removed by absorption with filter paper. 3 μL of phosphotungstic acid (PTA) at 2% was added on the formvar grid for negative staining. The

sample was incubated for 10 min, and excess PTA was removed with filter paper. Grids were examined in a Scanning Electron Microscope Field Emission Model 200 Nova NanoSEM FEI Company.

4.8. Thermal characterization of CCMV virus-like particles

The use of dyes like SYPRO Orange to monitor thermal denaturation of proteins by fluorescence detection provides a fast and inexpensive way for determination of protein stability using qPCR machines [20]. This technique is commonly referred to as thermal shift assay (TSA). Here, we use such a methodology to study the VLPs self-assembly and stability in solution. The VLP-dye system is gradually heated while measuring the fluorescence signal. The methodology enables for the detection of two characteristic melting temperatures; Tm_1 corresponds to the VLP disassembly, while Tm_2 corresponds to the protein thermal unfolding (denaturation). The specific value of Tm_i is defined as $max_i(dF/dT)$.

The protocol of a TSA experiment involves the incubation of CCMV CPs with SYPRO Orange dye in PCR strip tubes. Each reaction contains the following components in a final volume of 30 μL: 1x assembly buffer (0.1 M Sodium acetate buffer, 500 mM NaCl, pH=4.8), or 1x disassembly buffer (100 mM Tris-HCl, 500 mM NaCl, pH=7.5), 7.5 μM of purified CCMV CPs, and 6x SYPRO Orange. The TSA reactions were performed using the StepOne^{TM} Real-Time PCR System (Applied Biosystems Cat. #4376357), with a temperature ramp starting at 25 °C and ending at 95 °C in 1 °C/min increments. The fluorescence signal was read at excitation and emission λ wavelengths of 490 nm and 585 nm respectively. A sample of Lysozyme resuspended in assembly buffer was included in each run as a positive control (Sigma 62970). Raw data sets were exported from the qPCR machine for further analysis using the software we describe below.

4.9. Data analysis

The raw data produced by the readings of the qPCR machine are the fluorescence intensity as a function of cycles or temperature per well. Depending on the plate size and the number of optical channels, 48, 96, or more experiments per run can be made with information from several wavelength readings in a high-throughput fashion. Usually, the qPCR machine control software can export this raw data set. However, further analysis to obtain the melting temperature (Tm) can be a cumbersome process. We developed a Matlab program to calculate such values. In addition to facilitate the analysis by reading and ordering the exported qPCR data in the XLSX formatted file, it provides a graphical interface where the user can choose the optical channel and set of wells to analyze, calculate the average and standard deviation of the fluorescence vs. temperature profile if more than one repetition was included in the experimental design, and numerically calculate dF/dT. The source code and instructions to use the software can be accessed at https://github.com/tripplab/TSA-Tm. Once we obtained the Tm_i values for all the CP variants and conditions, we used the R software environment version 3.6.1 for statistical computing and graphics through RStudio version 1.1.463.

Author Contributions: Conceptualization, M.C.-T.; methodology, M.A.-M and M.C.-T.; software, A.A.P.-M.; validation, A.G.V.-L., J.A.A.-V. and M.A.-M; formal analysis, M.C.-T.; investigation, A.G.V.-L., J.A.A.-V., M.A.-M, N.P.-A., E.M.-G., P.G. and M.C.-T.; resources, N.P.-A., E.M.-G. and M.C.-T.; data curation, M.C.-T.; writing–original draft preparation, M.C.-T.; writing–review and editing, M.C.-T.; visualization, A.G.V.-L., A.A.P.-M., M.A.-M, N.P.-A., E.M.-G. and M.C.-T.; supervision, M.A.-M and M.C.-T.; project administration, M.A.-M and M.C.-T.; funding acquisition, M.C.-T.

Funding: This research was funded by the Consejo Nacional de Ciencia y Tecnología México (CONACYT grants number 132376 and CB2017-2018 A1-S-17041 to M.C.-T.).

Acknowledgments: The authors thank Dr. Bruno A. Escalante Acosta from CINVESTAV Unidad Monterrey for kindly providing access to the qPCR machine used in this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CCMV	Cowpea Chlorotic Mottle Virus
CP	capsid protein
VLP	virus-like particle
PPI	protein-protein interactions
WT	wild-type
COM	center-of-mass

1. Fox, J.M.; Wang, G.; Speir, J.A.; Olson, N.H.; Johnson, J.E.; Baker, T.S.; Young, M.J. Comparison of the Native CCMV Virion within VitroAssembled CCMV Virions by Cryoelectron Microscopy and Image Reconstruction. *Virology* **1998**, *244*, 212 – 218. doi:https://doi.org/10.1006/viro.1998.9107.
2. Sorger, P.; Stockley, P.; Harrison, S. Structure and assembly of turnip crinkle virus: II. Mechanism of reassembly in vitro. *Journal of Molecular Biology* **1986**, *191*, 639 – 658. doi:https://doi.org/10.1016/0022-2836(86)90451-1.
3. Zhao, X.; Fox, J.M.; Olson, N.H.; Baker, T.S.; Young, M.J. In Vitro Assembly of Cowpea Chlorotic Mottle Virus from Coat Protein Expressed in Escherichia coli and in Vitro-Transcribed Viral cDNA. *Virology* **1995**, *207*, 486 – 494. doi:https://doi.org/10.1006/viro.1995.1108.
4. Díaz-Valle, A.; García-Salcedo, Y.M.; Chávez-Calvillo, G.; Silva-Rosales, L.; Carrillo-Tripp, M. Highly efficient strategy for the heterologous expression and purification of soluble Cowpea chlorotic mottle virus capsid protein and in vitro pH-dependent assembly of virus-like particles. *J Virol Methods* **2015**, *225*, 23–9.
5. Bakail, M.; Ochsenbein, F. Targeting protein–protein interactions, a wide open field for drug design. *Comptes Rendus Chimie* **2016**, *19*, 19 – 27. Emerging Chemistry in France, doi:https://doi.org/10.1016/j.crci.2015.12.004.
6. Carrillo-Tripp, M.; Shepherd, C.M.; Borelli, I.A.; Venkataraman, S.; Lander, G.; Natarajan, P.; Johnson, J.E.; Brooks, C.L.; Reddy, V.S. VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res* **2009**, *37*, D436–42.
7. Carrillo-Tripp, M.; Montiel-García, D.J.; Brooks, C.L.; Reddy, V.S. CapsidMaps: protein-protein interaction pattern discovery platform for the structural analysis of virus capsids using Google Maps. *J Struct Biol* **2015**, *190*, 47–55.
8. Lavelle, L.; Gingery, M.; Phillips, M.; Gelbart, W.M.; Knobler, C.M.; Cadena-Nava, R.D.; Vega-Acosta, J.R.; Pinedo-Torres, L.A.; Ruiz-Garcia, J. Phase Diagram of Self-assembled Viral Capsid Protein Polymorphs. *The Journal of Physical Chemistry B* **2009**, *113*, 3813–3819.
9. Efimova, Y.M.; Haemers, S.; Wierczinski, B.; Norde, W.; Well, A.A.v. Stability of globular proteins in H2O and D2O. *Biopolymers* **2007**, *85*, 264–273.
10. Bogan, A.A.; Thorn, K.S. Anatomy of hot spots in protein interfaces11Edited by J. Wells. *Journal of Molecular Biology* **1998**, *280*, 1 – 9. doi:https://doi.org/10.1006/jmbi.1998.1843.
11. Sviben, D.; Bertoša, B.; Hloušek-Kasun, A.; Forcic, D.; Halassy, B.; Brgles, M. Investigation of the thermal shift assay and its power to predict protein and virus stabilizing conditions. *Journal of Pharmaceutical and Biomedical Analysis* **2018**, *161*, 73 – 82. doi:https://doi.org/10.1016/j.jpba.2018.08.017.
12. Bruce, D.; Cardew, E.; Freitag-Pohl, S.; Pohl, E. How to Stabilize Protein: Stability Screens for Thermal Shift Assays and Nano Differential Scanning Fluorimetry in the Virus-X Project. *Journal of Visualized Experiments* **2019**, *144*, e58666.
13. Boivin, S.; Kozak, S.; Meijers, R. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expression and Purification* **2013**, *91*, 192 – 206. doi:https://doi.org/10.1016/j.pep.2013.08.002.
14. Kranz, J.K.; Schalk-Hihi, C. Chapter eleven - Protein Thermal Shifts to Identify Low Molecular Weight Fragments. In *Fragment-Based Drug Design*; Kuo, L.C., Ed.; Academic Press, 2011; Vol. 493, *Methods in Enzymology*, pp. 277 – 298. doi:https://doi.org/10.1016/B978-0-12-381274-2.00011-X.

15. Rupesh, K.R.; Smith, A.; Boehmer, P.E. Ligand induced stabilization of the melting temperature of the HSV-1 single-strand DNA binding protein using the thermal shift assay. *Biochemical and Biophysical Research Communications* **2014**, *454*, 604 – 608. doi:https://doi.org/10.1016/j.bbrc.2014.10.145.
16. Miyazaki, Y.; Miyake, A.; Doi, N.; Koma, T.; Uchiyama, T.; Adachi, A.; Nomaguchi, M. Comparison of Biochemical Properties of HIV-1 and HIV-2 Capsid Proteins. *Frontiers in Microbiology* **2017**, *8*, 1082. doi:10.3389/fmicb.2017.01082.
17. He, F.; Hogan, S.; Latypov, R.F.; Narhi, L.O.; Razinkov, V.I. High throughput thermostability screening of monoclonal antibody formulations. *Journal of Pharmaceutical Sciences* **2010**, *99*, 1707 – 1720. doi:https://doi.org/10.1002/jps.21955.
18. Katen, S.; Zlotnick, A. Chapter 14 The Thermodynamics of Virus Capsid Assembly. In *Biothermodynamics, Part A*; Academic Press, 2009; Vol. 455, *Methods in Enzymology*, pp. 395 – 417. doi:https://doi.org/10.1016/S0076-6879(08)04214-6.
19. Zlotnick, A.; Aldrich, R.; Johnson, J.M.; Ceres, P.; Young, M.J. Mechanism of Capsid Assembly for an Icosahedral Plant Virus. *Virology* **2000**, *277*, 450 – 456. doi:https://doi.org/10.1006/viro.2000.0619.
20. Huynh, K.; Partch, C.L. Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. *Current Protocols in Protein Science* **2015**, *79*, 28.9.1–28.9.14.

Sample Availability: Samples of the transformed *E. coli* with cloned expression plasmid of WT and all variants studied here are available from the authors.