PNAS
www.pnas.org

Supplementary Information for

Crystal structure of tomato spotted wilt virus G_N reveals a unique dimerization and an evolutionary link to animal infecting viruses

Yoav Bahat[a], Joel Alter[a] and Moshe Dessau[a,1]

[a]Azrieli Faculty of Medicine, Bar-Ilan University, Safed, Israel
[1]To whom correspondence should be addressed: E-mail: moshe.dessau@biu.ac.il. 8 Henrietta Szold St. , B 104, Safed 1311502, Israel. Phone: +972 (72) 264-4905, Fax: +972 (4) 6229256.

Corresponding author: Moshe Dessau

Email: moshe.dessau@biu.ac.il

This PDF file includes:

Figures S1 to S5
Tables S1
SI References

www.pnas.org/cgi/doi/10.1073/pnas.2004657117

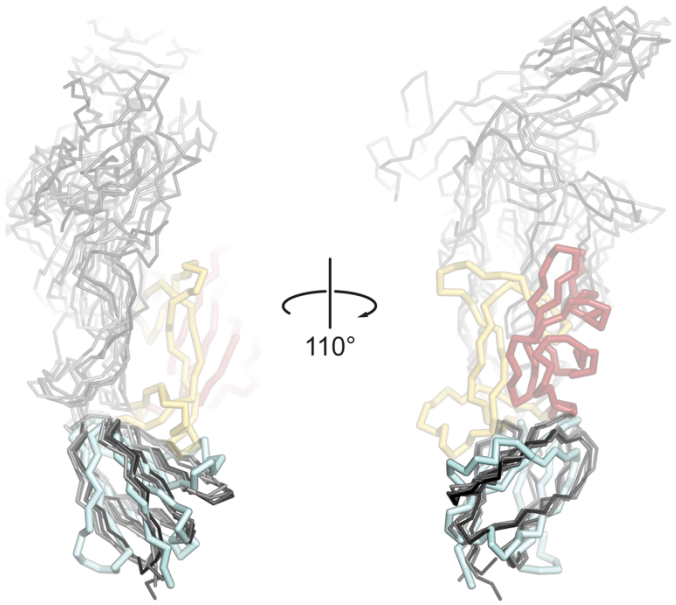

Fig. S1. Superposition of TSWV with alphavirus E2. TSWV G_N structure was superimposed onto various structures of alphavirus E2; CHIKV (5ANY), Sinbis virus (3MUU), Eastern Equine Encephalitis Virus (6MX4) and Venezuelan Equine Encephalitis Virus (3J0C). The average RMAD was 2.9 Å over ~70 residues of the CTD.

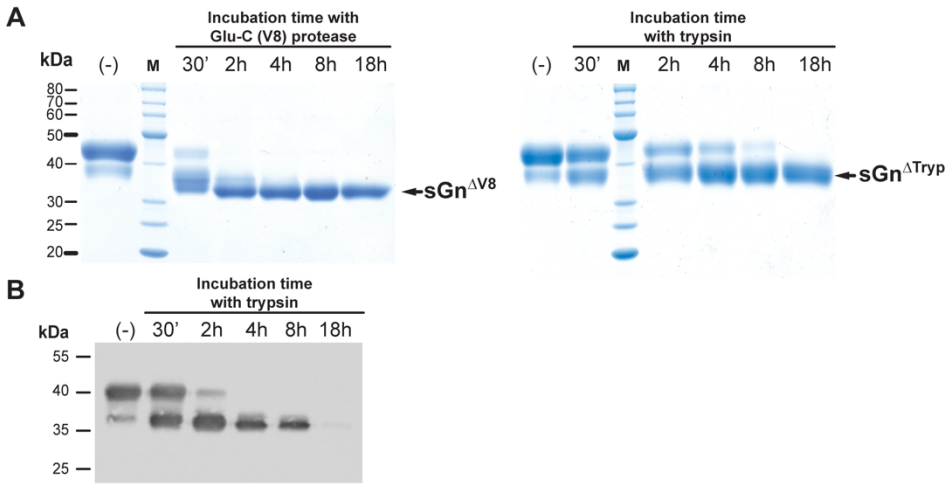

Fig. S2. Limited proteolysis of sG_N. Purified sG_N was incubated on ice with either Glu-C (V8) protease or Trypsin at a protein to protease ratio of 150:1 or 1000:1, respectively. Aliquots were taken from the reaction at the indicated times, SDS sample buffer was added, and samples were boiled and analyzed later by SDS-PAGE and Coomassie blue staining. The digestion products are noted on right side of the relevant gel.

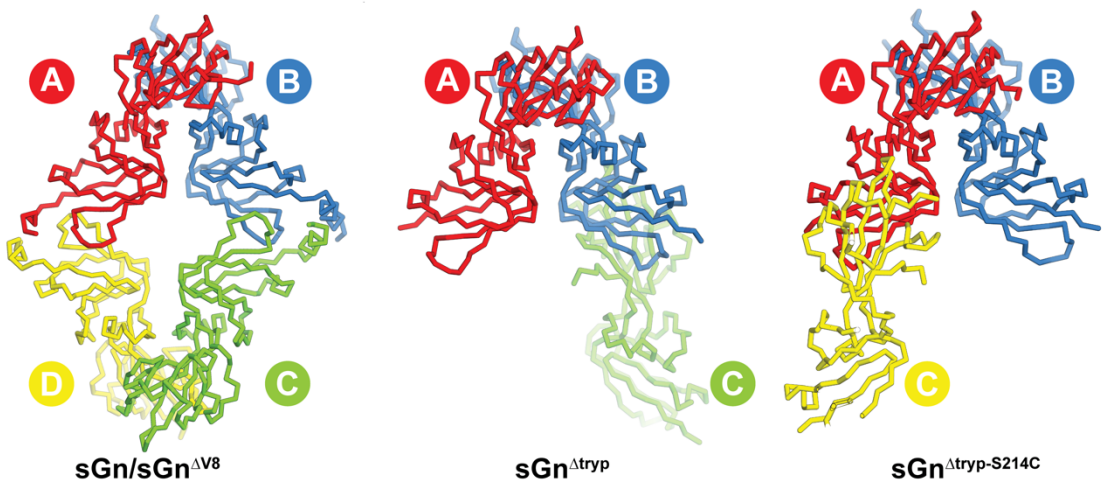

Fig. S3. Asymmetric unit (ASU) organization in the different crystal forms. Ribbon representation of the packing of the different G_N constructs. Each chain is noted by a different color with the same labeling as in Figure 4.

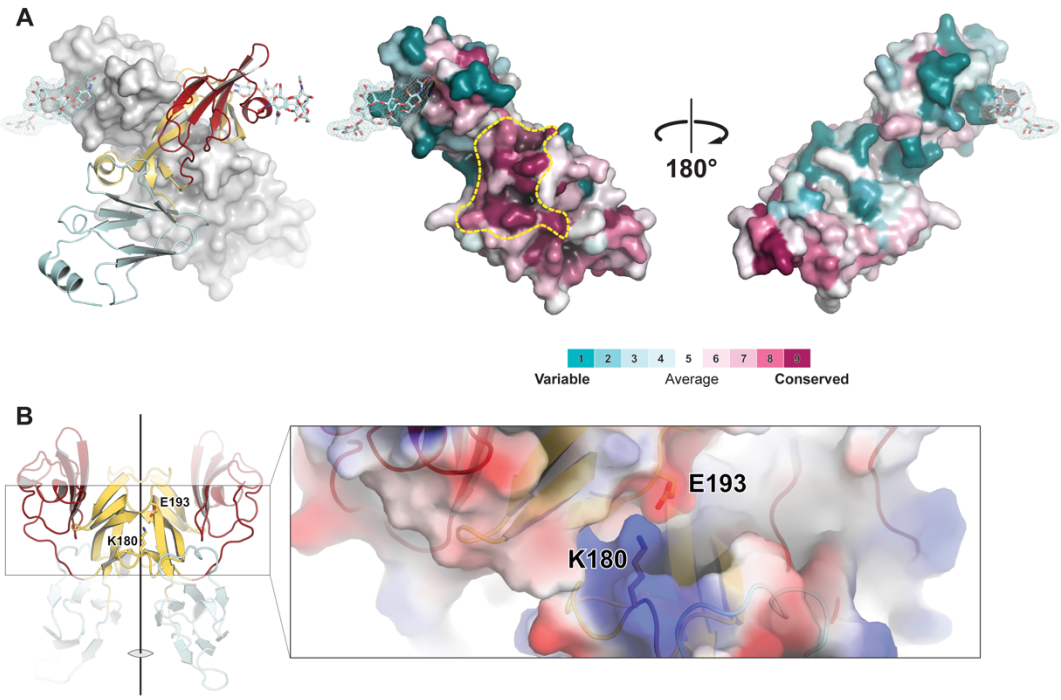

Fig. S4. Conservation analysis and electrostatic surface potential complementarity in the PD-PD dimer interface. (A) CONSURF analysis (1, 2). On the left, a view of G_N non-covalent dimer for orientation. Conservation scores are as per legend. See materials and methods for GeneBank numbers of the aligned sequences. The interface that participates in the dimerization is highlighted by the dashed yellow line. (B) Cartoon representation of the non-covalent dimer (left) in a view perpendicular to the two-fold axis. Residues that participate in the electrostatic interaction are shown as sticks. On the right is a surface representation of the electrostatic molecular surface potential at the dimer interface, calculated using the APBS plugin in PyMOL (The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.)

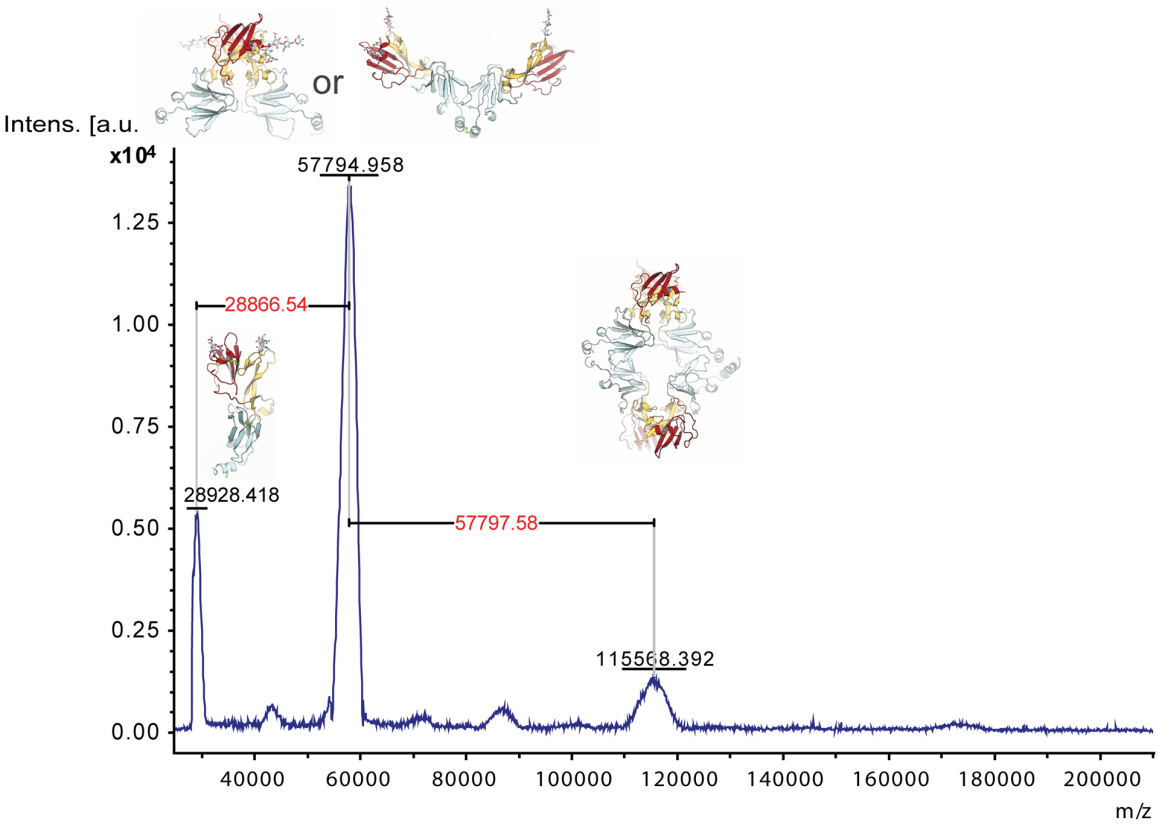

Fig. S5. MALDI-TOF analysis of $sG_N^{\Delta V8}$ dissolved crystals. The peaks in the figure correspond to the indicated masses in black. Red masses correspond to the mass differences between the peaks. Peaks were assigned to either monomers, dimers or tetramers.

Table S1. Data collection and refinement statistics

		sG_N-Os	sG_N	$sG_N^{\Delta V8}$	$sG_N^{\Delta Tryp}$	$sG_N^{\Delta Tryp\text{-}S214C}$
PDB ID			6Y9L	6Y9M	6YA0	6YA2
Data collection						
Wavelength		1.14	0.9763	1.03961	0.9795	0.9184
Resolution range[a]		72.64-4.5 (4.66-4.5)	49.02-4.10 (4.25-4.10)	45.6-3.4 (3.52 - 3.4)	46.2- 2.86 (2.96- 2.86)	18.83-2.50 (2.59- 2.50)
Space group		*C222₁*	*C222₁*	*C222₁*	*P2₁*	*P2₁*
Unit cell	*a, b, c (Å)*	167.8, 217.2, 145.2	166.0, 215.9, 146.4	167.2, 218.2, 145.5	68.8, 74.9, 81.3	69.7, 76.0, 71.2
	α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 103.27, 90	90, 106.56, 90
Total reflections		422805 (43456)	133273 (11408)	1399124 (86127)	63863 (6628)	170586 (17809)
Unique reflections		30437 (1596)	20687 (1798)	36907 (3602)	18705 (215)	24421 (704)
Multiplicity		13.9 (14.2)	6.4 (6.2)	37.9 (23.6)	3.4 (3.6)	7.0 (7.2)
Completeness (%)		99.95 (99.94)	98.18 (86.55)	99.74 (98.39)	75.92 (11.55)	82.56 (28.45)
Mean I/sigma(I)		11.48 (1.05)	8.58 (0.36)	18.13 (0.64)	10.51 (0.95)	13.21 (0.97)
Wilson B-factor		40	295.07	160.71	77.68	50.18
R_{merge}[b]		0.1167 (2.61)	0.1069 (4)	0.1526 (6.68)	0.06385 (1.53)	0.1019 (1.80)
R_{meas}		0.1212 (2.71)	0.1166 (4.36)	0.1547 (6.83)	0.07598 (1.80)	0.1101 (1.94)
R_{pim}		0.03261 (0.72)	0.0461 (1.71)	0.02501 (1.39)	0.0408 (0.95)	0.04131 (0.72)
CC1/2		0.999 (0.56)	0.998 (0.23)	1 (0.28)	0.998 (0.49)	0.998 (0.58)
CC*		1 (0.85)	1 (0.61)	1 (0.66)	1 (0.81)	1 (0.86)
Refinement statistics						
Reflections used in refinement			20554 (1795)	36843 (3601)	14242 (215)	20416 (684)
Reflections used for R_{free}[c]			1028 (91)	1840 (171)	1425 (21)	1026 (36)
R_{work}[c]			0.2765 (0.4339)	0.2645 (0.42)	0.2175 (0.3126)	0.2130 (0.2987)
R_{free}[c]			0.3017 (0.4681)	0.2884 (0.4431)	0.2613 (0.4644)	0.2445 (0.3392)
Number of non-hydrogen atoms			6325	6441	4396	4438
macromolecules			6082	6148	4264	4247
ligands			243	286	100	112
solvent			-	7	32	79
Protein residues			801	791	560	551
RMS[d] (bonds)			0.004	0.011	0.006	0.006
RMS[d] (angles)			0.72	1.42	0.76	1.02
Ramachandran favored (%)			88.9	93.5	93.39	97.05
allowed (%)			10.71	6.5	6.42	2.95
outliers (%)			0.39	0	0.18	0
Rotamer outliers (%)			3.76	0.69	5.94	4.17
B-factor						
Average			331.02	159.34	91.45	75.27
macromolecules			326.95	155.75	92.21	73.81
Synchrotron Beam line		ESRF ID29	BESSY II 14.1	ESRF ID23-1	ESRF ID29	BESSY II 14.2

[a] Highest resolution shell is shown in parentheses

[b] $R_{merge} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - \langle I \rangle_{hkl}| / \Sigma_{hkl}\Sigma_i |I_{hkl,i}|$, where I_{hkl} is the intensity of a reflection and $\langle I \rangle_{hkl}$ is the average of all observations of the reflection

[c] R_{free}, R_{work} with 10% of F_{obs} sequestered before refinemen

[d] R.M.S., root mean square

SI References

1. M. Landau *et al.*, ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33**, W299-302 (2005).
2. H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **38**, W529-533 (2010).