Biophysical Journal, Volume *120*

Supplemental information

Biophysical properties of the isolated spike protein binding helix of human ACE2

Anirban Das, Vicky Vishvakarma, Arpan Dey, Simli Dey, Ankur Gupta, Mitradip Das, Krishna Kant Vishwakarma, Debsankar Saha Roy, Swati Yadav, Shubham Kesarwani, Ravindra Venkatramani, and Sudipta Maiti

Supplementary Information for

Biophysical Properties of the Isolated Spike Protein Binding Helix of Human ACE2

Anirban Das[#$], Vicky Vishvakarma[#$], Arpan Dey[#], Simli Dey[#], Ankur Gupta[#], Mitradip Das[#], Krishna Kant Vishwakarma[#], Debsankar Saha Roy[#], Swati Yadav[^], Shubham Kesarwani[§], Ravindra Venkatramani[#], and Sudipta Maiti[#*]

[#]Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India

[^]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bellary Road, Bangalore 560065, Karnataka, India

[§]Centre for Cardiovascular Biology and Disease (CCBD), Institute of Stem Cell Science and Regenerative Medicine (inStem), Gandhi Krishi Vigyan Kendra Campus, Bangalore 560065, Karnataka, India

[$]Equal contributions

*Email: maiti@tifr.res.in

This PDF file includes:

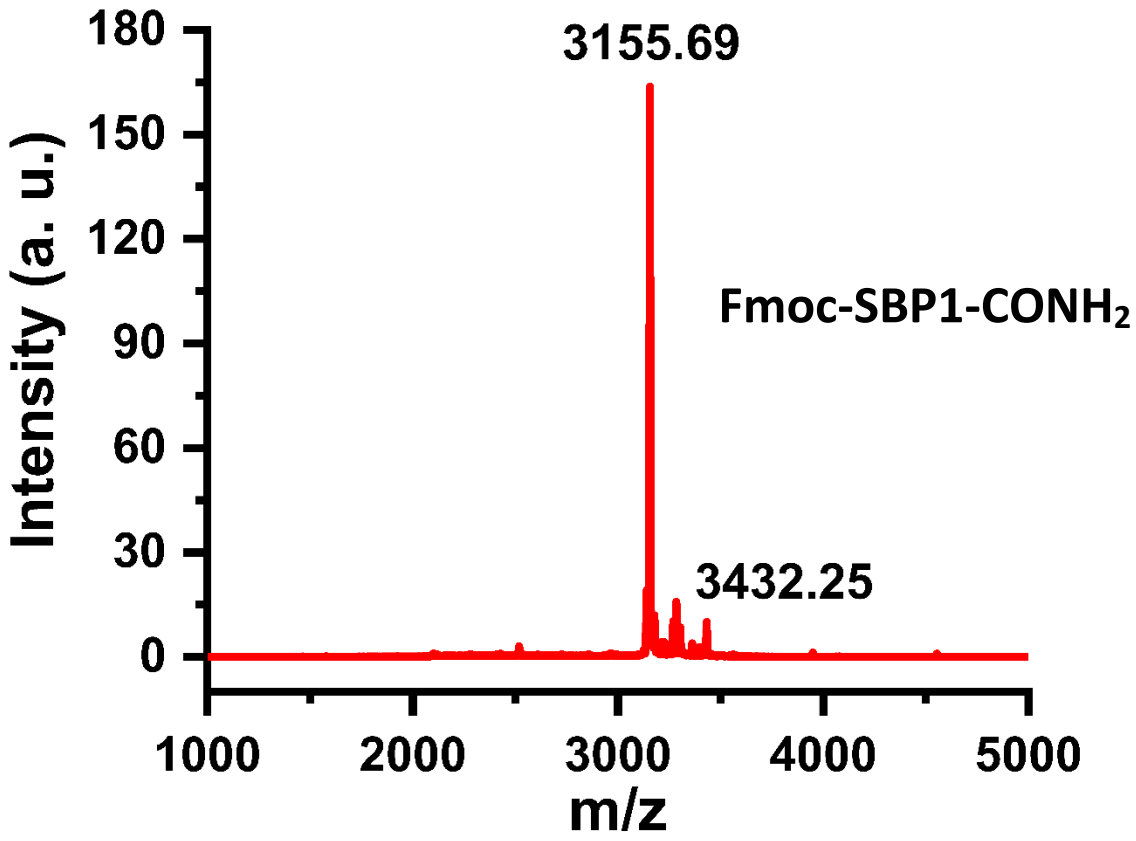


Figure S1. Matrix-assisted laser desorption/ionization - time-of-flight (MALDI-TOF) mass spectrum (in reflector positive mode) of the N-terminal Fmoc-protected SBP1-$CONH_2$. The peptide was dissolved in 1:1 acetonitrile-water mixture containing 0.1% trifluoroacetic acid and used for the MS analysis.

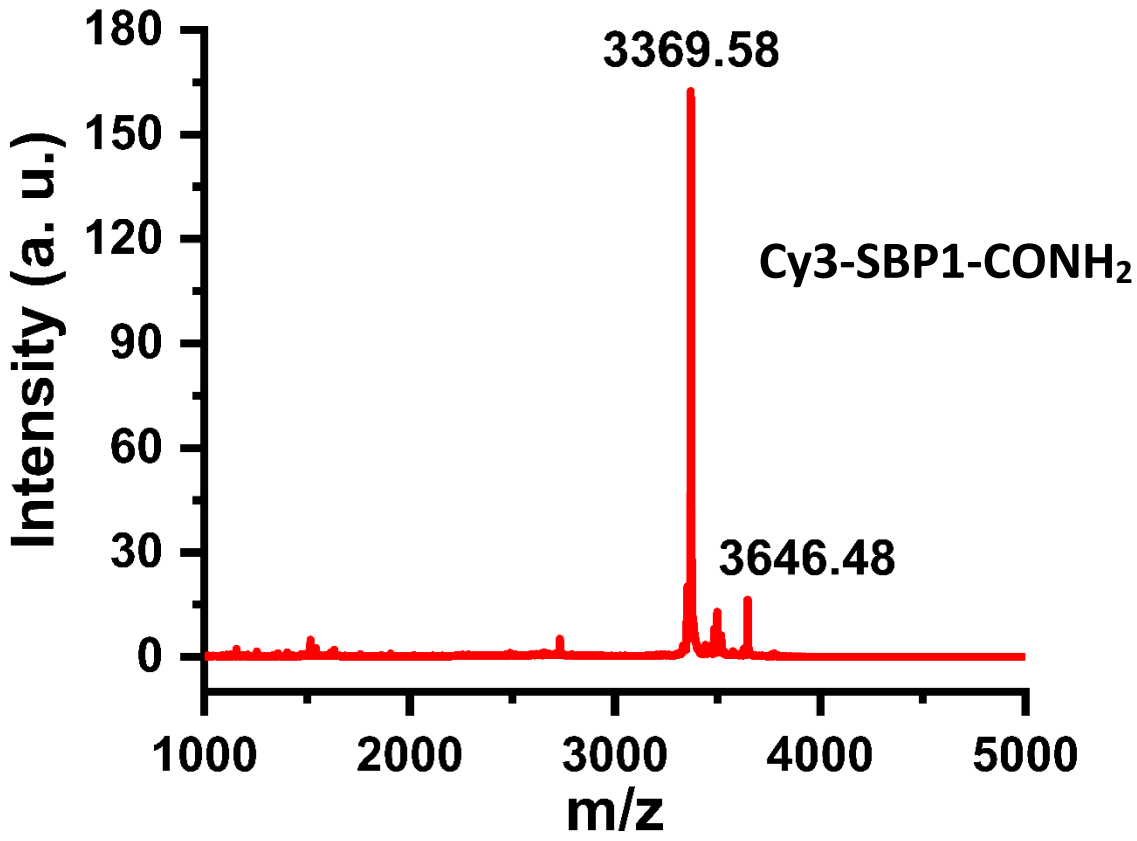
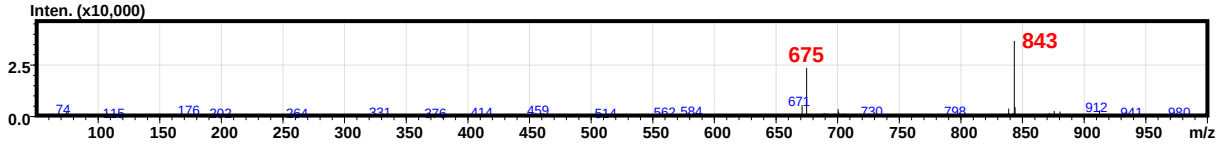


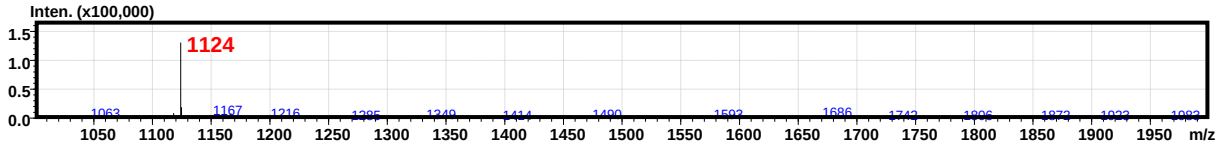
Figure S2. MALDI-TOF mass spectrum (in reflector positive mode) of the N-terminal Cyanine 3-labelled SBP1-$CONH_2$ (Cy3-SBP1-$CONH_2$). The peak at m/z = 3369.58 (highest intensity peak) corresponds to dye labelled SBP1 peptide. This indicates that the dye labelling was complete (~100%).

(A) Spectrum View

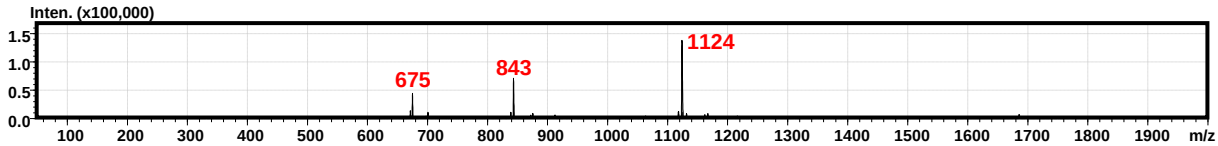
Event#: 1 **Scan(E+)** Ret. Time: [0.560->1.760] - [0.000->0.560] Scan#: [57->177] - [1->57]



Event#: 2 **Scan(E+)** Ret. Time: [0.565->1.765] - [0.005->0.565] Scan#: [59->179] - [3->59]



Event#: 3 **Profile(E+)** Ret. Time: [0.601->1.801] - [0.041->0.601] Scan#: [63->183] - [7->63]



(B) Contour View

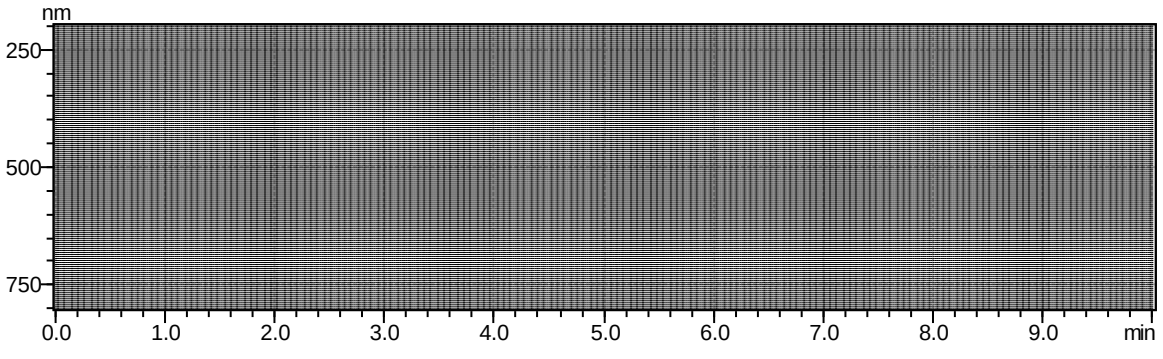


Figure S3. Electrospray ionization mass spectrometric (ESIMS) analysis of the Cy3-SBP1-$CONH_2$ peptide. The lyophilized peptide was dissolved in 1:1 acetonitrile-water mixture containing 0.1% trifluoroacetic acid and directly injected in the MS for the analysis. (A) Shows the ESI-MS spectra of the peptide in the positive mode (in the spectrum view), Event#: (1) from m/z = 0 to 1000, (2) m/z = 1000 to 2000 and (3) Profile mode: from m/z = 0 to 2000, with the $[M + 3H]^{3+}$, $[M + 4H]^{4+}$ and $[M + 5H]^{5+}$ peaks (highlighted in red) having the maximum intensities. (B) Shows the elution profile in the contour view confirming the presence of the Cy3-labelled peptide.

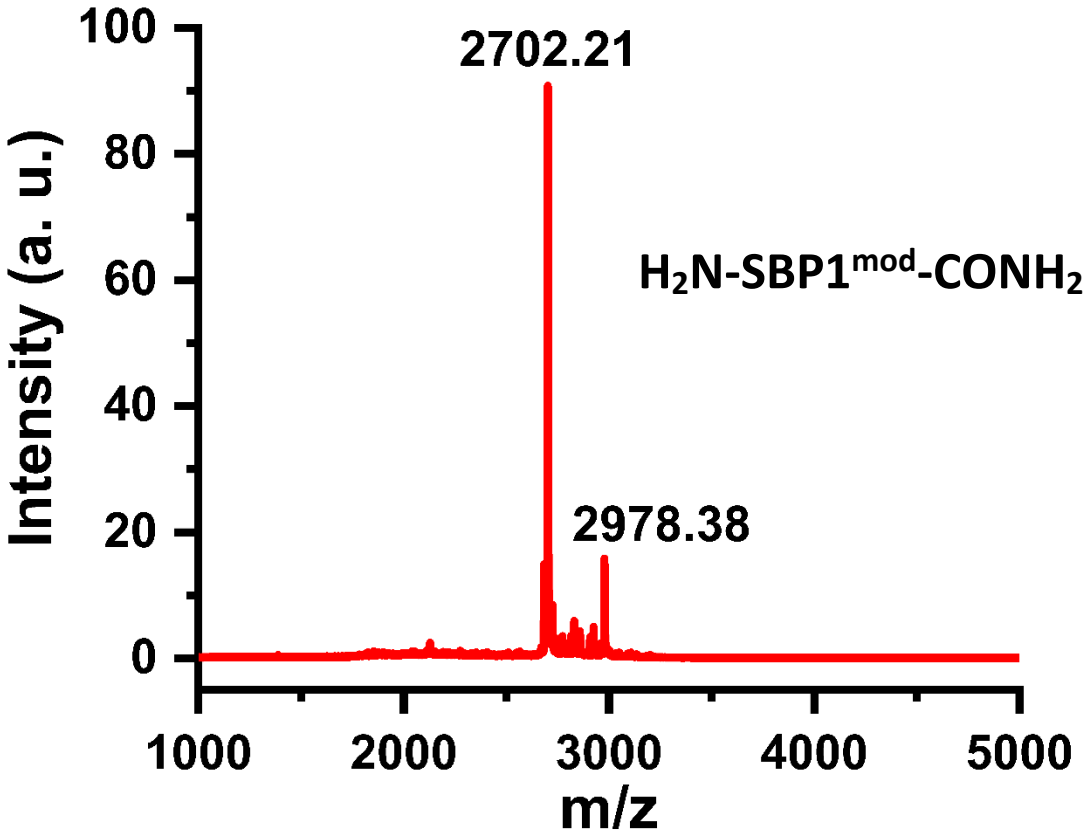


Figure S4. MALDI-TOF mass spectrum (in reflector positive mode) of the $SBP1^{mod}$-$CONH_2$ (H_2N-$SBP1^{mod}$-$CONH_2$, modified SBP1) peptide. The peak at m/z = 2702.21 (highest intensity peak) corresponds to $SBP1^{mod}$-$CONH_2$ peptide.

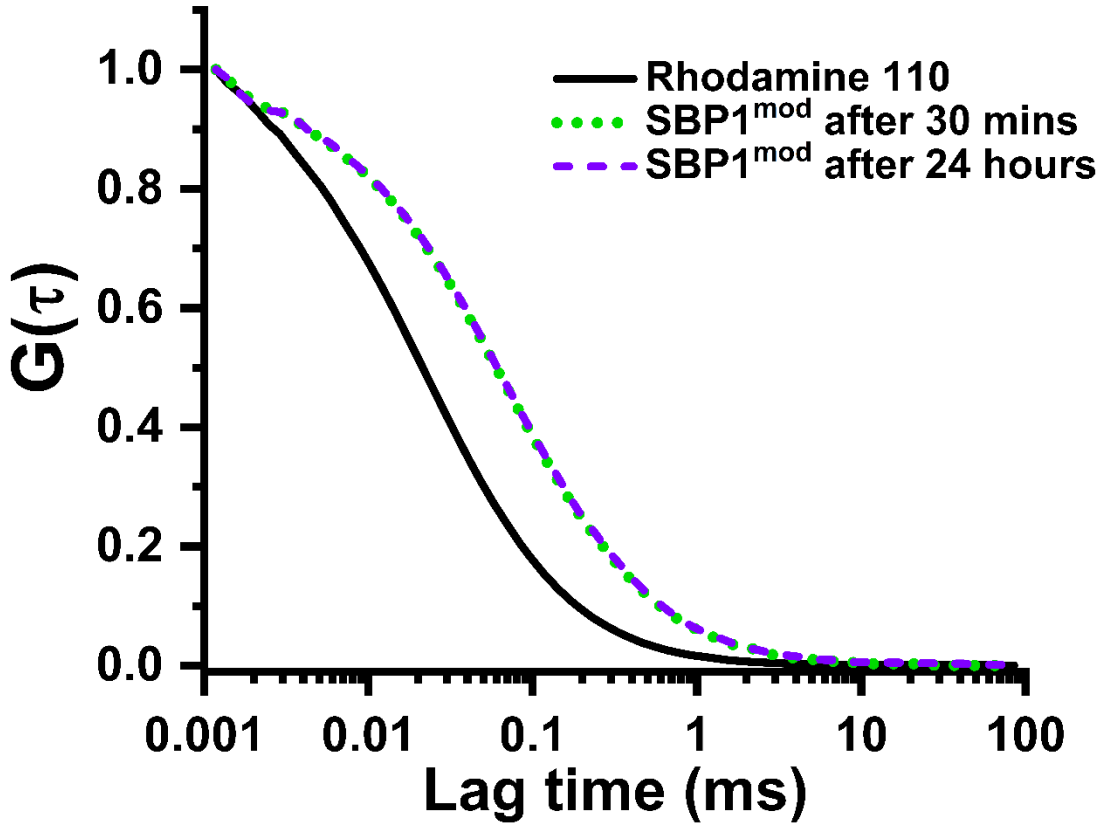


Figure S5. Normalized fluorescence autocorrelation data obtained from free rhodamine 110 in solution (black, solid line), 140 nM Rh110-labelled SBP1^{mod} peptide after 30 minutes (green, dotted line) and 24 hours (violet, dashed line) of incubation.

Photon Counting Histogram (PCH) Analysis:

We have carried out a PCH analysis of the FCS data. The experimental PCH data is obtained using a MATLAB program from time trace measurement in Time Tagged Time-Resolved (TTTR) mode using PicoHarp 300 software (Picoquant). The PCH data is fitted with the following equation for multiple independent species[1] using a code written in Python.

$$\pi(k; \overline{N}_1, \overline{N}_2, \epsilon_1, \epsilon_2) = \pi(k; \overline{N}_1, \epsilon_1) * \pi(k; \overline{N}_2, \epsilon_2)$$

Where the function $\pi(k; \overline{N}, \epsilon)$ describes the probability of observing *k* photon counts in an open system for a single species particle solution with $\overline{N}_1$, $\overline{N}_2$ representing the average number of particles corresponding to species 1 and 2 respectively and ϵ_1, ϵ_2 representing brightness of molecule of species 1 and 2 respectively.

While a solution containing monomers and dimers [Fig. S6 (A)] does fit well. A solution containing only monomers do not fit well [Fig. S6 (B)]. PCH fits including larger number of components, given the quality of the data, does not provide unique parameter values. However, this analysis shows that the solution cannot have just monomers.

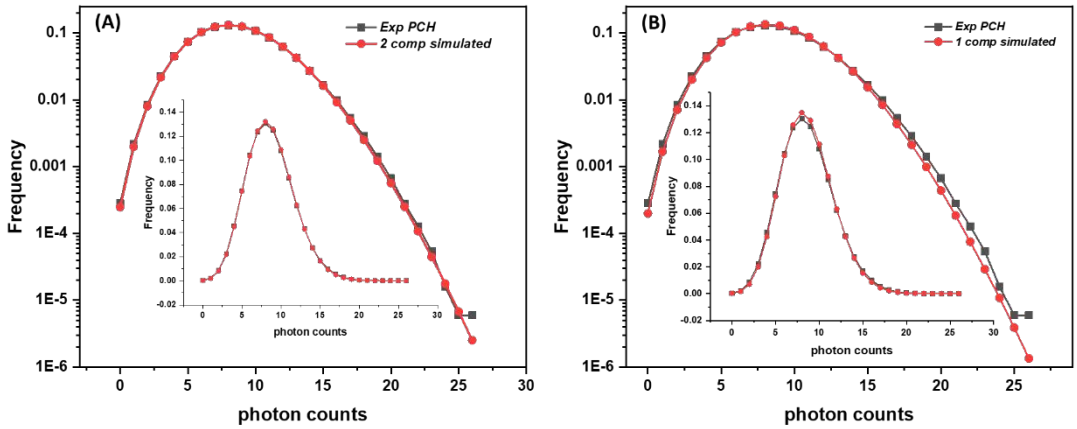


Figure S6. Photon counting histogram for SBP1^{mod} at 140 nM. (A)The histogram is fitted with 2 species having different brightness (ϵ) values (ϵ_1 = 0.12, ϵ_2 = 0.24, for the monomer and the dimer) with 1:2.3 ratio for $\overline{N}_1$ and $\overline{N}_2$ respectively, using the theoretical PCH function $\pi(k; \overline{N}_1, \overline{N}_2, \epsilon_1, \epsilon_2)$[1].The inset displays the same data in linear scale for comparison. (B) The histogram is fitted with 1 species using the theoretical PCH function $\pi(k; \overline{N}, \epsilon)$[1] having ϵ = 0.20 and $\overline{N}$ = 95. The inset displays the same data in linear scale for comparison.

Dynamic Light Scattering (DLS) for Monitoring Aggregation of SBP1 and SBP1^{mod}:

It is well-known that adding a fluorescence label can potentially alter the properties of any peptide. We have now carried out DLS measurements of both SBP1 and SBP1^{mod} (unlabelled) to clarify whether the observed oligomerization is an effect of the labelling or is somehow altered upon dye labelling. Fig. S7 shows the correlation traces of SBP1 and SBP1^{mod} at 1 µM [Panels (A) and (B)] and at 10 µM [Panels (C) and (D), respectively] concentrations. We observe multiple traces which show the presence of larger aggregates in the solution. We cannot access concentration below this for such a small peptide, due to the lack of sensitivity of DLS (compared to FCS). We conclude that while the fluorescent labels may change some properties, the peptide has a tendency to aggregate even without any fluorescent labels.

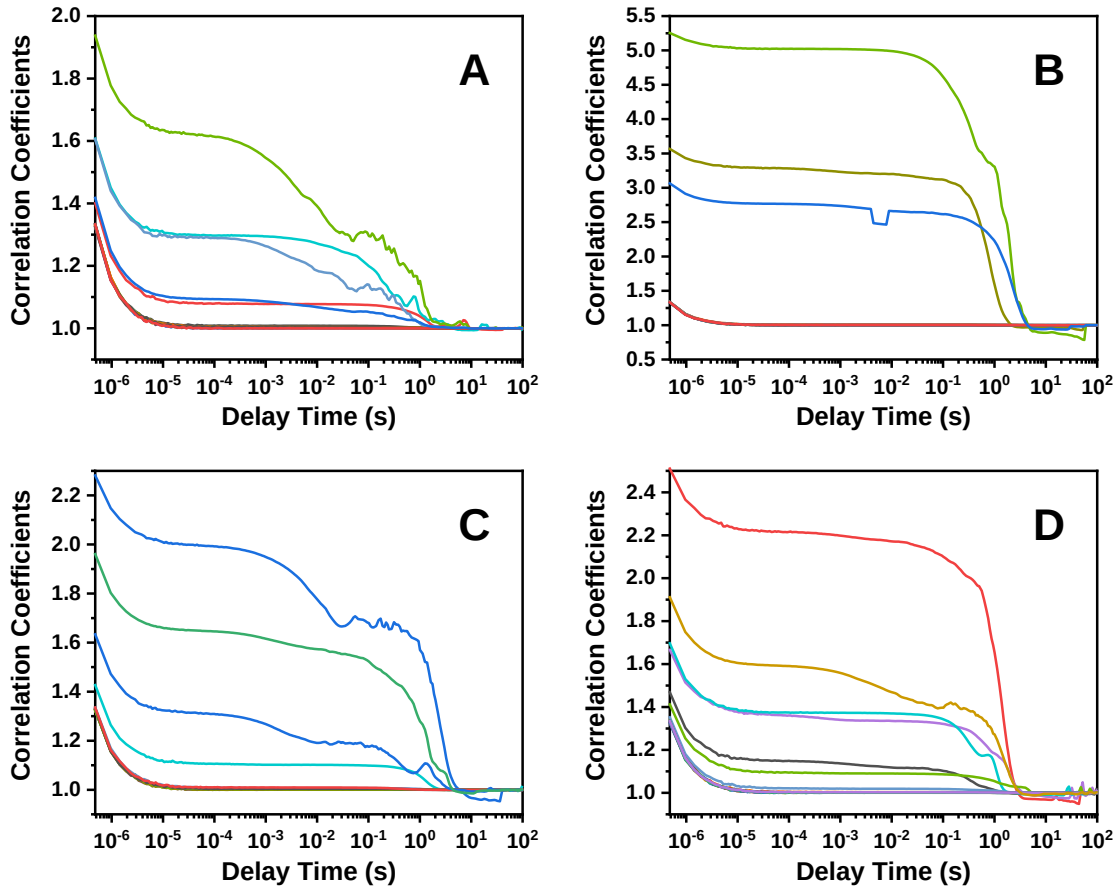
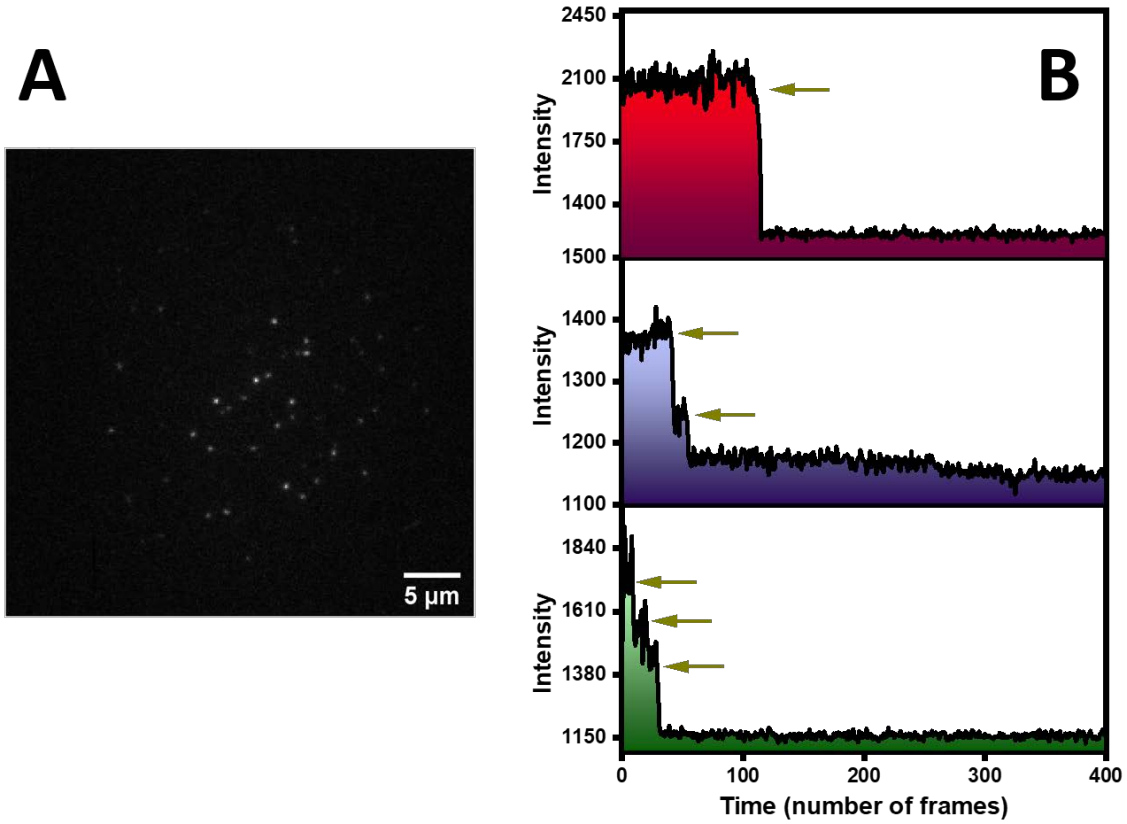


Figure S7. Correlation curves obtained from the dynamic light scattering (DLS) measurements of SBP1 and SBP1^{mod} at concentrations of 1 µM [Panels (A) and (B)] and at 10 µM [Panels (C) and (D)], respectively.



115

116 **Figure S8.** (A) TIRF image of freshly prepared Cy3-labelled SBP1 (1 nM) on glass coverslip.
117 Scale bar is 5 µm. (B) Time traces observed during the photobleaching of individual SBP1
118 oligomers.

119

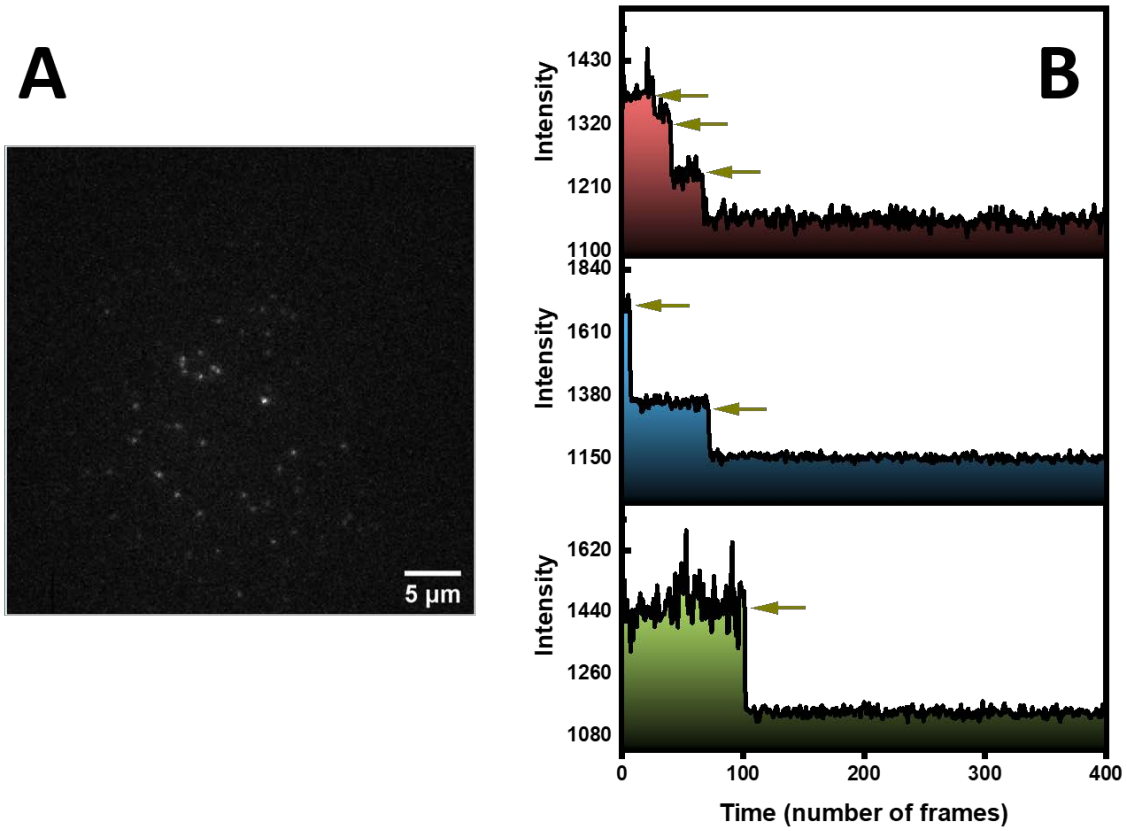
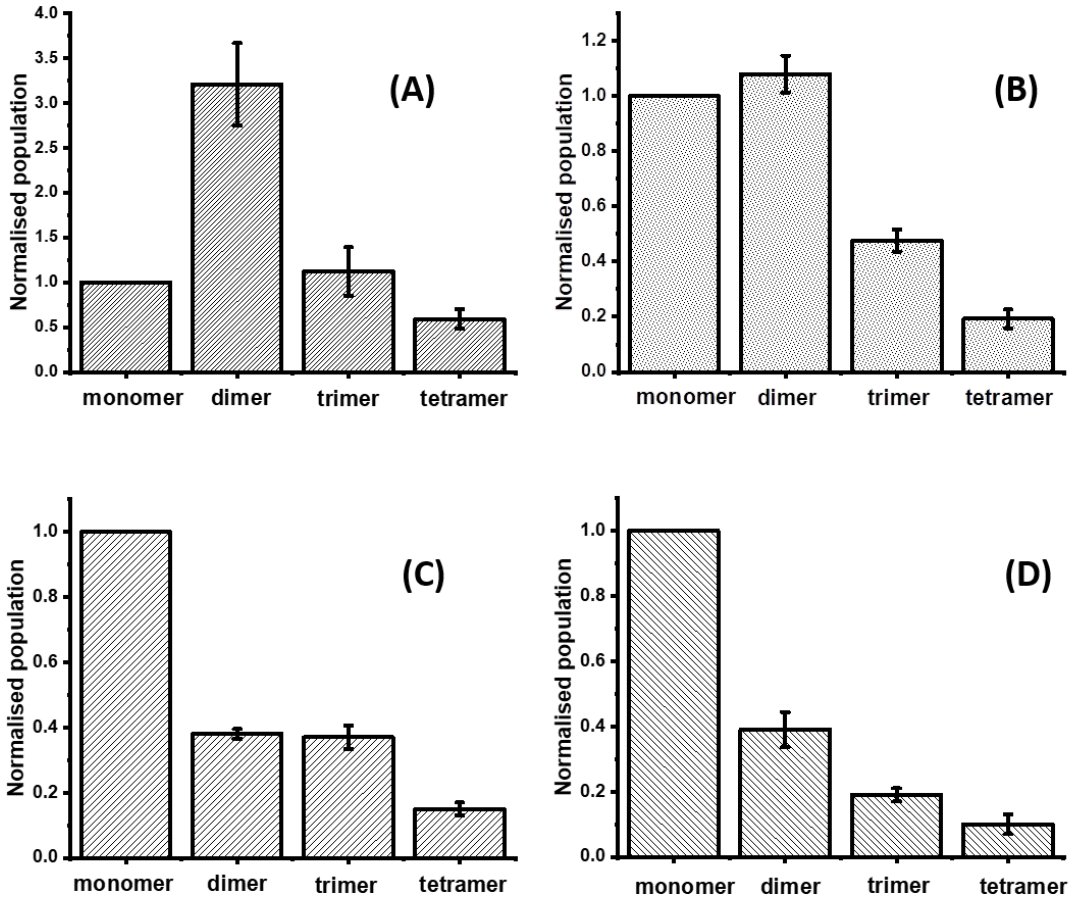


Figure S9. (A) TIRF image of Cy3-labelled SBP1 (1 nM) after 24 hours of incubation. Scale bar is 5 µm. (B) Time traces observed during the photobleaching of individual SBP1 oligomers.

Comparison of the extent of oligomerization of SBP1 and SBP1^{mod} from Single-Molecule Photobleaching (smPB) Experiment:

We have checked the decrease in the Cy3-SBP1 peptide concentration as a result of sticking to the wall of the tube by measuring the Cy3 fluorescence in the solution. We see that the concentration of a 1 nM solution, after 24 hrs, decreases to 720 pM. We then measured the stoichiometry of the oligomers of a freshly prepared solution having the same 720 pM concentration. We also repeated the same experiments with Cy3-SBP1^{mod}. In this case, the concentration reduces to 733 pM after 24 hrs.



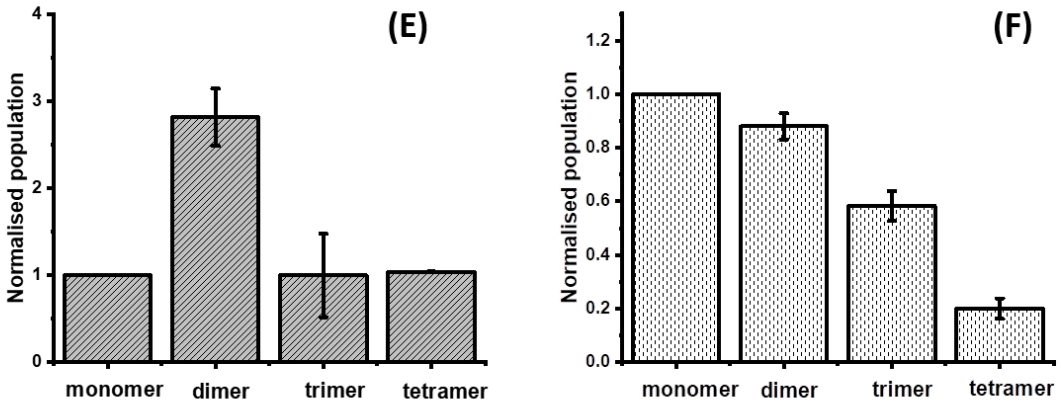


Figure S10. Pre-bleach corrected oligomer distributions of (A) a fresh 1 nM solution of Cy3-SBP1 peptide, (C) after 24 hrs rotation with a final concentration 720 pM, (E) a fresh 720 pM peptide solution. Similarly, for Cy3-SBP1^{mod} (modified SBP1), (B) oligomeric distribution of a fresh 1 nM solution, (D) after 24 hrs rotation with a final concentration of 733 pM, (F) a fresh 733 pM solution. For SBP1 fresh N = 485 points (3 sets), SBP1 after 24 hrs N = 373 points (3 sets), SBP1 720 pM N = 428 points (3 sets), SBP1^{mod} fresh N = 392 points (3 sets), SBP1^{mod} after 24 hrs N = 305 points (3 sets), SBP1^{mod} 733 pM N = 354 points (3 sets). For each sets 6 ROIs (region of interest) are taken. Average and standard error of the mean (plotted error) are calculated.

Figure S10 (A) represents the SBP1 pre-bleach corrected oligomeric distribution of a fresh 1 nM stock. This, after 24 hrs, changes to a monomer heavy distribution [Figure S10 (C)]. The concentration in Figure S10 (C) is 720 pM. A fresh SBP1 stock of the same concentration [Figure S10 (E)], on the other hand, shows a very similar distribution as that in Figure S10 (A). This suggests that monomerization was not due to a mismatch in concentration. Data shown in Figure S10 (A) and S10 (C) are fresh repeats of those shown in Figure 4 (C) and 4 (D) respectively. We observe a slight variation in the relatively ratios of the population of different oligomers, however, the data does not change qualitatively.

A similar observation is seen in the case of Cy3-SBP1^{mod} also. Figure S10 (B) represents the oligomeric distribution of the fresh 1 nM stock of Cy3-SBP1^{mod}. The oligomerization status shows that it is less aggregation prone than SBP1. After 24 hrs, the oligomer changes to a monomer heavy distribution as shown in Figure S10 (D). The concentration was measured to be 733 pM. A fresh stock solution of the same concentration [Figure S10 (F)] showed a very similar oligomer distribution as in Figure S10 (B).

Hence, for both the systems, we see that the observed monomerization is a time dependent phenomenon, and is not caused by a mismatch in concentration between the samples.

Section S1: Theoretical calculations of hydrodynamic radius and comparison with the experimental results

We have calculated the expected hydrodynamic radius of the SBP1 peptide given its length (and aspect ratio) in the crystal structure. Detailed calculations are shown below:

The axial ratio or aspect ratio p = (a/b), where a = end to end distance (C-alpha, measured) = 3.3 nm and b = diameter (including side chain, from ideal helix assumption) = 1.2 nm. Hence, p = 32.88/12 = 2.74. For a prolate spheroid of this aspect ratio, the Perrin factor is ~1.1[2]. Hence the expected effective hydrodynamic radius, $R_h = (\frac{3V}{4\pi})^{(\frac{1}{3})} \times f_P \sim 1.06\ nm$

From our FCS measurements, we observe that SBP1 has a hydrodynamic radius of 3.1 ± 0.2 nm, and the SBP1^{mod} peptide has a radius of 2.6 ± 0.1 nm. Both of these values are larger than that calculated for the helical monomer (above), so it suggests oligomerization. After 24 hours of incubation, both the peptides have a very similar hydrodynamic radii of ~2.5 nm. If we assume a quasi-spherical shape for the oligomer, then the initial aggregation state (average) of SBP1 is ~25-mers and that of SBP1^{mod} is ~15-mers, while after 24 hours, the average stoichiometry becomes ~13-mers. This of course assumes that the peptide remains in the helical state, which is not true. If the peptide is more disordered, the actual number of monomers in an oligomer would be smaller, as we show below.

A random coil is likely to be smaller than the long helix, so formally it might be suggested that the slow change in radius represents a loss of secondary structure by hydrophobic solvation of the molecule. However, as we see below, it turns out not to be true.

For a Gaussian random coil peptide in a good solvent, an estimate for the hydrodynamic radius (R_h) of random coil peptides is provided by Wilkins *et al.*[3]. For unfolded proteins, $R_h = 2.21 \times N^{0.57}$), which yields a value of 1.32 nm. Therefore, a random coil would actually have a higher R_h than a compact cylinder (~1.06 nm, as shown before), as far as SBP1 is concerned. Therefore, the slow reduction of the radius is unlikely to represent a loss of secondary structure.

This supports our inference that the initial solutions as well as 24-hours incubated solutions of both SBP1 and SBP1^{mod} (at 140 nM) contain oligomeric species. If we calculate on the basis of the random coil radius (taken to be 1.32 nm), the average oligomer size of 2.5 nm corresponds to about 7-mers. At lower concentration, the R_h for SBP1 initially becomes ~1.7 nm, which corresponds to dimers on an average, and after incubation, it becomes ~1.4 nm (which corresponds to a nearly monomeric state) [Fig. 3B and 3D]. This is not far off from the values measured in our single molecule photobleaching measurements, where we incubated the SBP1 peptide at a concentration of 3 nM for 70 minutes.

Section S2: Computational methods

In order to examine the binding of the SBP1 in the helical as well as in the non-helical (unwound) form, we docked the structure with the RBD of the S-protein. The helical conformation of the peptide and RBD were both modelled using the crystal structure of ACE2 bound to the RBD, present in RCSB database with PDB ID: 6M0J[4] (Fig. 1A). For docking, Haddock 2.4 (July 2020) with CNS 1.3 was used[5,6]. The results were analyzed using in-house scripts and visualizations were carried out using VMD[7].

In order to define the binding interface for docking studies, the crystal structure of RBD bound to ACE2 complex (PDB ID: 6M0J[4]) was used to include all residues within 10Å of the ACE2 N-terminal segment 21 to 44. All the residues were considered as 'active' residues for docking purposes. In order to model the SBP1, we extracted residues 21 to 44 of ACE2 from the same complex (PDB ID: 6M0J[4]) and mutated the 44th residue to lysine using CHARMM-GUI[8]. The whole α-helical SBP1 peptide was considered to also be 'active' during docking. The whole peptide was considered to be semi-flexible although strict conditions were provided to preserve the secondary structure of SBP1 to be α-helical. In docking the helical form of SBP1 with RBD, 6000 structures were initially selected for rigid body docking, with the best 400 (based on binding energy scores) chosen for final full flexible refinement and further analysis. All other docking options were set as per default values defined in Haddock (see Table S1).

In order to carry out the docking studies with the 'unwound' form of SBP1, we generated extended SBP1 conformations using a combination of MD simulations and simulated annealing protocols. These computations were carried out using NAMD software version 2.13[9] and the CHARMM36 force field[10] for the peptide. For modelling SBP1, the coordinates of residue 21 to 44 were extracted from the crystal structure of RBD-ACE2 complex (PDB ID: 6M0J) and the residue 44 was mutated to lysine using CHARMM-GUI. Then, we heated the protein in vacuum from 0 to 700 K at a rate of 6K/ps and then evolved the system for ∼2 ns with a time step of 1 fs. After that the protein was cooled from 700 to 300 K at a rate of 6K/ps and then evolved again for ∼2 ns. The cooling step was repeated 5 times to generate different non-helical conformation of SBP1. The 5 non-helical structures of SBP1 thus obtained were solvated in a rectangular TIP3P water box with a solvent padding of 12 Å around the protein and neutralized by adding 4 sodium ions. The Particle mesh Ewald (PME) method for electrostatics with a cutoff value of 12 Å for long-range non-bonded interactions was used. Langevin dynamics and the Nose-Hoover Langevin piston barostat method are used to implement temperature and pressure control respectively. The solvated system was first minimized for 1000 steps and then gradually heated to 300 K while keeping the protein non-hydrogen atoms fixed. Finally, an unconstrained 1 ns NPT run was performed to equilibrate the system at a pressure of 1 bar and temperature of 300 K. This is followed by 10 ns production runs under NVT conditions with a time step of 2 fs. The structure of the peptides

obtained at the end of 5 independent NVT runs are taken as input SBP1 conformations for docking to RBD to check the binding energies. For docking of the non-helical (unwound) SBP1 structures, the definition of the binding interface remained the same as given above for the helical conformation). However, in order to account for the greater flexibility of the non-helical conformation during the process of docking, 10000 structures were generated from rigid body docking, out of which the 1000 best (based on binding energy scores) structures were chosen for fully flexible optimizations and further analysis. In this case, the number of steps used during optimization including heating and dynamics stages were four times longer than the default values. In all cases, cluster analysis was performed using root mean square deviation (RMSD) based clustering with a cutoff of 5Å. A complete list of parameters which differ for docking helical and non-helical conformers along with Haddock defaults is given in Table S1. All other setting/parameters used were as per Haddock defaults.

Parameter	Default	Helix	Non-Helical
ssdihed	none	alphabeta	none
structures_0	6000	6000	10000
structures_1	400	400	1000
anastruc_1	400	400	1000
waterrefine	400	400	1000
waterheatsteps	100	100	500
watersteps	1250	1250	5000
watercoolsteps	500	500	2000
clust_meth	FCC	RMSD	RMSD
clust_cutoff	0.6	5	5

Table S1. Docking options used in Haddock for docking of helical and non-helical states of SBP1 with RBD of S-protein. Haddock default values are provided as a reference.

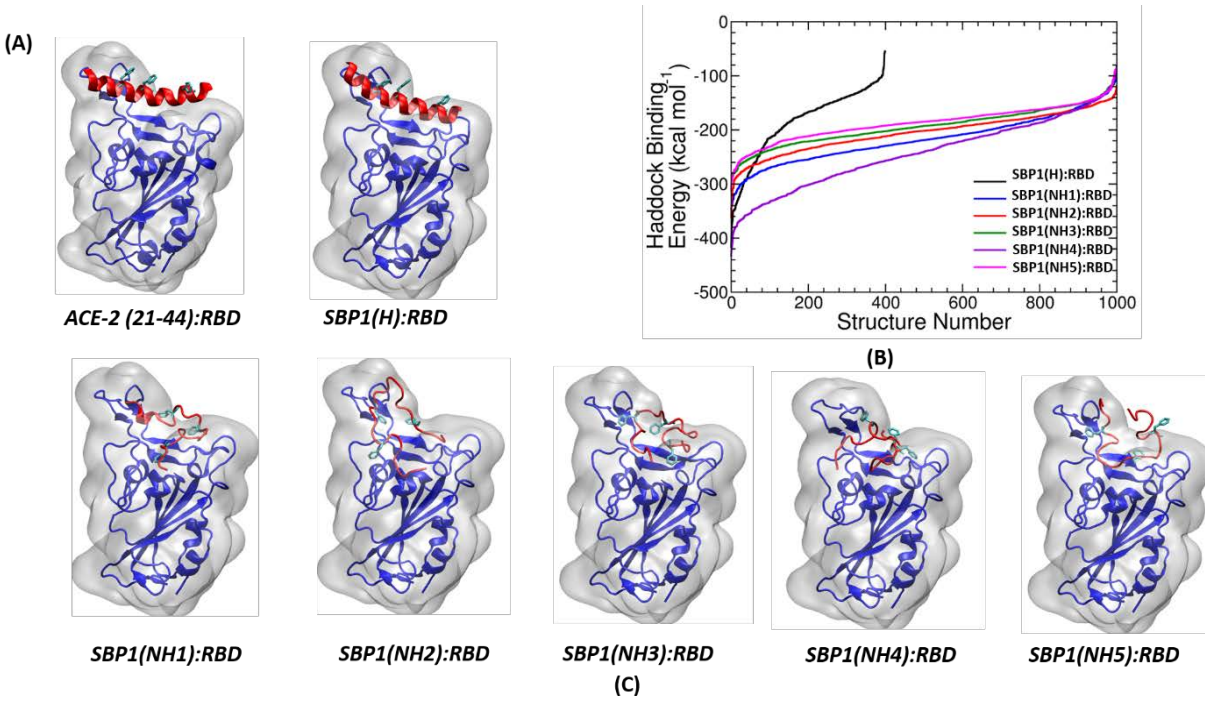


Figure S11. (A) Comparison of the binding mode for the N-terminal segment 23-41 of ACE2 and the helical conformation of SBP1 with RBD (surface and secondary structure representation). (B) Binding energies for the 1000 best RBD docked structures of helical and non-helical SBP1 peptides. (C) Binding modes of the 5 non-helical SBP1 structures derived from MD and simulated annealing with RBD.

No	SBP1 Structure	Binding Energy Scores
1	Helix	-416.888 kcal mol^{-1}
2	Non-helical 1	-346.787 kcal mol^{-1}
3	Non-helical 2	-361.570 kcal mol^{-1}
4	Non-helical 3	-311.241 kcal mol^{-1}
5	Non-helical 4	-432.882 kcal mol^{-1}
6	Non-helical 5	-309.539 kcal mol^{-1}
7	Non-helical Average	-352.404 kcal mol^{-1}

Table S2. Haddock binding energy scores of the best docking poses of SBP1 (helical and 5 non-helical conformations) with RBD of S-protein (PDB ID: 6M0J[4]).

SBP1(H):RBD		*SBP1(NH1):RBD*		*SBP1(NH2):RBD*		*SBP1(NH3):RBD*		*SBP1(NH4):RBD*		*SBP1(NH5):RBD*	
RBD	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**
Y505	E37	Q493	K26	Y449	Q42	K417	E23	S494	Q42	Q493	T27
K417	D30	Y489	E22	R408	E23	G502	S43	Q493	S43	Y489	Y41
N487	I21	N487	E22	D405	K26	Q493	T27	N487	Q24	N487	Q42
N487	Q24	K417	E23			E484	K31	K417	E23	K417	E35
Q498	Q42							Q409	E37	N501	K31
E484	K31							K417	E37	K417	E35
S477	Q24							Y505	F32		
F497	Q42							E484	S43		

Table S3. H-bonding interactions of helical and non-helical SBP1 with the spike RBD. The table shows the residues of SBP1 and RBD which form H-bonds with each other in SBP1:RBD complexes.

SBP1(H):RBD		*SBP1(NH1):RBD*		*SBP1(NH2):RBD*		*SBP1(NH3):RBD*		*SBP1(NH4):RBD*		*SBP1(NH5):RBD*	
RBD	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**	**RBD**	**SBP1**
K417	D30	K417	E23	D405	K26	E484	K31	K417	E23	K417	E35
E484	K31	D405	K44	R408	E23						
R403	E37										

Table S4. Salt-bridge interactions of helical and non-helical SBP1 with the spike RBD. The table shows the residues of SBP1 and RBD which form salt-bridges with each other in SBP1:RBD complexes.

SI References:

(1) Chen, Y.; Müller, J. D.; So, P. T. C.; Gratton, E. The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy. *Biophys. J.* **1999**, *77* (1), 553–567. https://doi.org/https://doi.org/10.1016/S0006-3495(99)76912-2.

(2) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*; Biophysical chemistry; W. H. Freeman, 1980.

(3) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **1999**, *38* (50), 16424–16431. https://doi.org/10.1021/bi991765q.

(4) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581* (7807), 215–220. https://doi.org/10.1038/s41586-020-2180-5.

(5) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731–1737. https://doi.org/10.1021/ja026939x.

(6) van Zundert, G. C. P.; Rodrigues, J. P. G. L. M.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428* (4), 720–725. https://doi.org/https://doi.org/10.1016/j.jmb.2015.09.014.

(7) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(8) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User

Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865. https://doi.org/https://doi.org/10.1002/jcc.20945.

(9) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 44130. https://doi.org/10.1063/5.0014475.

(10) Huang, J.; MacKerell Jr, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145. https://doi.org/https://doi.org/10.1002/jcc.23354.