Supplementary Data to:

Investigating vulnerability of the conserved SARS-CoV-2 spike's heptad repeat 2 as target for fusion inhibitors using chimeric miniproteins

Daniel Polo-Megías [a], Mario Cano-Muñoz [a], Alberto G. Berruezo [a,1], Géraldine Laumond [b], Christiane Moog [c,d], and Francisco Conejero-Lara [a,*]

[a] Departamento de Química Física, Instituto de Biotecnología y Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente (UEQ), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

[b] Laboratoire d'ImmunoRhumatologie Moléculaire, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR_S 1109, Institut Thématique Interdisciplinaire (ITI) de Médecine de Précision de Strasbourg, Transplantex NG, Faculté de Médecine, Fédération Hospitalo-Universitaire OMICARE, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, F-67000 Strasbourg, France

[c] Vaccine Research Institute (VRI), F-94000 Créteil, France

*Corresponding author: Francisco Conejero-Lara, Departamento de Química Física, Facultad de Ciencias, Campus Fuentenueva, Universidad de Granada, 18071 Granada, Spain. Tel. +34 958242371; Fax. +34 958 242747, e-mail: conejero@ugr.es.

[1] Present address: Instituto Biofisika (CSIC, UPV/EHU), Fundación Biofísica Bizkaia/Biofisika Bizkaia Fundazioa (FBB), 48940, Leioa, Spain

Table S1-A. Mutations at the HR1 and HR2 regions for the major SARS-CoV-2 variants of concern (VOCs). Mutations are highlighted in red.

HR1

Variant (Seq. reference)	913	950
Wuhan-Hu-1 (NC_045512.2)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Alpha B.1.1.7(MZ344997.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Beta B.1.351(MW598419.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Gamma P1(MW642250.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Delta B.1.617.2(MZ009823.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ**N**	
Omicron BA.1(OL672836.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Omicron BA.2(OM371884.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Omicron BA.4(ON373214.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	
Omicron E.G.5.1(OP790748.1)	TQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQD	

	951	988
Wuhan-Hu-1 (NC_045512.2)	VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE	
Alpha B.1.1.7(MZ344997.1)	VVNQNAQALNTLVKQLSSNFGAISSVLNDIL**A**RLDKVE	
Beta B.1.351(MW598419.1)	VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE	
Gamma P1(MW642250.1)	VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE	
Delta B.1.617.2(MZ009823.1)	VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE	
Omicron BA.1(OL672836.1)	VVN**H**NAQALNTLVKQLSS**K**FGAISSVLNDI**F**SRLDKVE	
Omicron BA.2(OM371884.1)	VVN**H**NAQALNTLVKQLSS**K**FGAISSVLNDILSRLDKVE	
Omicron BA.4(ON373214.1)	VVN**H**NAQALNTLVKQLSS**K**FGAISSVLNDILSRLDKVE	
Omicron E.G.5.1(OP790748.1)	VVN**H**NAQALNTLVKQLSS**K**FGAISSVLNDILSRLDKVE	

HR2

	1164	1202
Wuhan-Hu-1 (NC_045512.2)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Alpha B.1.1.7(MZ344997.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Beta B.1.351(MW598419.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Gamma P1(MW642250.1)	VDLGDISGINAS**F**VNIQKEIDRLNEVAKNLNESLIDLQE	
Delta B.1.617.2(MZ009823.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Omicron BA.1(OL672836.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Omicron BA.2(OM371884.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Omicron BA.4(ON373214.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	
Omicron E.G.5.1(OP790748.1)	VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE	

Table S1-B. Sequence alignment of the Spikes's HR1 and HR2 regions for human coronaviruses. Alignment has been carried out with Clustal Omega at the EMBL's European Bioinformatics Institute (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) [1]

```
NCBI Ref.:          NC_045512.2         SARS-COV-2 (WUHAM)
GenBank:            AAP13441.1          SARS-COV
NCBI Ref.:          YP_009047204.1      MERS-COV
GenBank:            AMK59677.1          HCOV-OC43
GenBank:            AXT92561.1          HCOV-HKU1
GenBank:            AAS58177.1          HCOV-NL63
GenBank:            QNT54842.1          HCOV-229E
```

HR1

```
SARS-CoV-2          TQNVLYENQKLIANQFNSAIGKIQDSLSSTA--------------SALGKLQD
SARS-CoV            TQNVLYENQKQIANQFNKAISQIQESLTTTS--------------TALGKLQD
MERS-COV            TQQVLSENQKLIANKFNQALGAMQTGFTTTN--------------EAFQKVQD
HCOV-OC43           TMDVLSQNQKLIANAFNNALDAIQEGFDATN--------------SALVKIQA
HCOV-HKU1           TMDVLNKNQKLIANAFNKALLSIQNGFTATN--------------SALAKIQS
HCOV-NL63           QTDVLQENQKILAASFNKAINNIVASFSSVNDAITQTAEAIHTVTIALNKIQD
HCOV-229E           QTDVLQENQRILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQD
                      :** :**: :*  **.*:  :  .:  .               *: *:*

SARS-CoV-2          VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE
SARS-CoV            VVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVE
MERS-CoV            AVNNNAQALSKLASELSNTFGAISASIGDIIQRLDVLE
HCOV-OC43           VVNANAEALNNLLQQLSNRFGAISSSLQEILSRLDALE
HCOV-HKU1           VVNVNAQALNSLLQQLFNKFGAISSSLQEILSRLDNLE
HCOV-NL63           VVNQQGSALNHLTSQLRHNFQAISNSIQAIYDRLDSIQ
HCOV-229E           VVNQQGNSLNHLTSQLRQNFQAISSSIQAIYDRLDIIQ
                    .** :..:*. * .:*   * ***  :  * .*** ::
```

HR2

```
SARS-CoV-2          -VDLGDISGINASVVNIQKEID--------------RLNEVAKNLNESLIDLQE
SARS-CoV            -VDLGDISGINASVVNIQKEID--------------RLNEVAKNLNESLIDLQE
MERS-CoV            -PNFGSLTQINTTLLDLTYEML--------------SLQQVVKALNESYIDLKE
HCOV-OC43           -PDLS-LDYINVTLLDLQVEMN--------------RLQEAIKVLNQSYINLKD
HCOV-HKU1           -PNLTFNSHINATFLDLYYEMN--------------VIQESIKSLNSSFINLKE
HCOV-NL63           KPNFD-LTPFNLTYLNLSSELKQLEAKTASLFQTTVELQGLIDQINSTYVDLKL
HCOV-229E           VPDLV-VEQYNQTILNLTSEISTLENKSAELNYTVQKLQTLIDNINSTLVDLKW
                      ::      * : :::  *:                ::   . :*.: ::*:
```

Table S2. Amino acid sequences of the shortened CoVS-HR1 miniproteins.

Protein variant	Sequence
L3B	DVLYENQKLI ANKFNSAIGK IQDSLSSTAS ELGKLQDEVN QNAQDLNTLV KQLSSNFGRI SSELNDILSR LDKGEPAKDL RSDIDNLESK IAGFNSSLQK VLTNLAQKNQ NVEDKLKGLE SRTSSLEKQI KGIASNFQNE ILKQREYLVN KGSGNVLYEN QKLIENQFNS AIGKIQDSLS STKSALGKLK DVVNQNKQAL NTLVKQLSSN FGAISSVLND IKSRLDKVE
N2A	DVLYENQKLI ANKFNSAIGK IQDSLSSTAS ELGKLQDEVN KGESKKNVED KLKGLESRTS SLEKQIKGIA SNFQNEILKQ REYLVNKGSG NVLYENQKLI ENQFNSAIGK IQDSLSSTKS ALGKLKDVVN QN
N2B	DVLYENQKLI ANKFNSAIGK IQDSLSSTAS ELGKLQDEVN QNAQDLNTGK SGENLAQKNQ NVEDKLKGLE SRTSSLEKQI KGIASNFQNE ILKQREYLVN KGSGNVLYEN QKLIENQFNS AIGKIQDSLS STKSALGKLK DVVNQNKQAL NT
C2A	SELGKLQDEV NQNAQDLNTL VKQLSSNFGR ISSELNDILS RLDKGEPAKD LRSDIDNLES KIAGFNSSLQ KVLTNLAQKN QNVEDKLKKA GSDVKKLKDV VNQNKQALNT LVKQLSSNFG AISSVLNDIK SRLDKVEW
L3C	DVLYENQKLI ANKFNSAIKK IQDSLSSTAS ELKKLQDEVN QNAQDLNTLV KQLSSNFKRI SSELNDILSR LDKGEPAKDL RSDIDNLESK IAKFNSSLQK VLTNLAQKNQ NVEDKLKTLE SRTSSLEKQI KKIASNFQNE ILKQREYLVN KGSGNVLYEN QKLIENQFNS AIKKIQDSLS STKSALKKLK DVVNQNKQAL NTLVKQLSSN FSAISSVLND IKSRLDKVE
N2C	DVLYENQKLI ANKFNSAIKK IQDSLSSTAS ELKKLQDEVN KGESKKNVED KLKTLESRTS SLEKQIKKIA SNFQNEILKQ REYLVNKGSG NVLYENQKLI ENQFNSAIKK IQDSLSSTKS ALKKLKDVVN QN
N2D	DVLYENQKLI ANKFNSAIKK IQDSLSSTAS ELKKLQDEVN QNAQDLNTGK SGENLAQKNQ NVEDKLKTLE SRTSSLEKQI KKIASNFQNE ILKQREYLVN KGSGNVLYEN QKLIENQFNS AIKKIQDSLS STKSALKKLK DVVNQNKQAL NT
C2C	SELKKLQDEV NQNAQDLNTL VKQLSSNFKR ISSELNDILS RLDKGEPAKD LRSDIDNLES KIAKFNSSLQ KVLTNLAQKN QNVEDKLKKA GSDVKKLKDV VNQNKQALNT LVKQLSSNFS AISSVLNDIK SRLDKVEW

All proteins contain an N-terminal methionine and a C-terminal tail of sequence GGGGSHHHHHH. A C-terminal tryptophan added to confer UV absorption is highlighted in green for C2A and C2C proteins. Loop residues connecting the helices are highlighted in yellow. Mutated amino acids to substitute glycine residues are highlighted in magenta. Other amino acid substitutions made to engineer stabilizing salt bridges in the trimeric coiled-coil structure are colored in red (negative charge) or blue (positive charge).

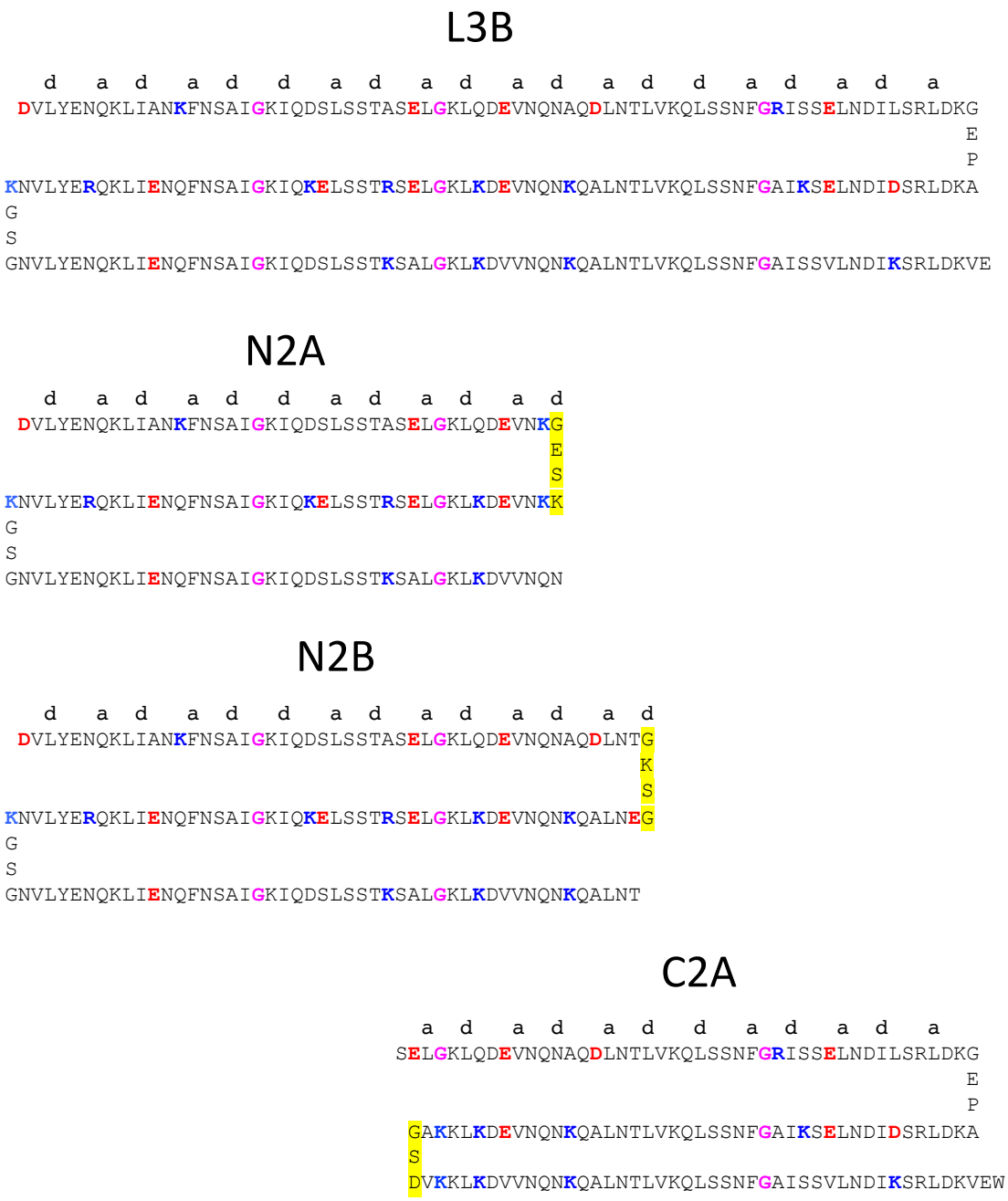


Figure S1: Sequence and topological organization of the CoVS-HR1 miniproteins. The newly built loops connecting each consecutive helix have been highlighted in yellow. Engineered residues to create stabilizing salt bridges at (e) and (g) positions of the heptad repeats are colored in blue and red. Substituted glycine residues in the L3C, N2C, N2D and C2C variants are colored in magenta.

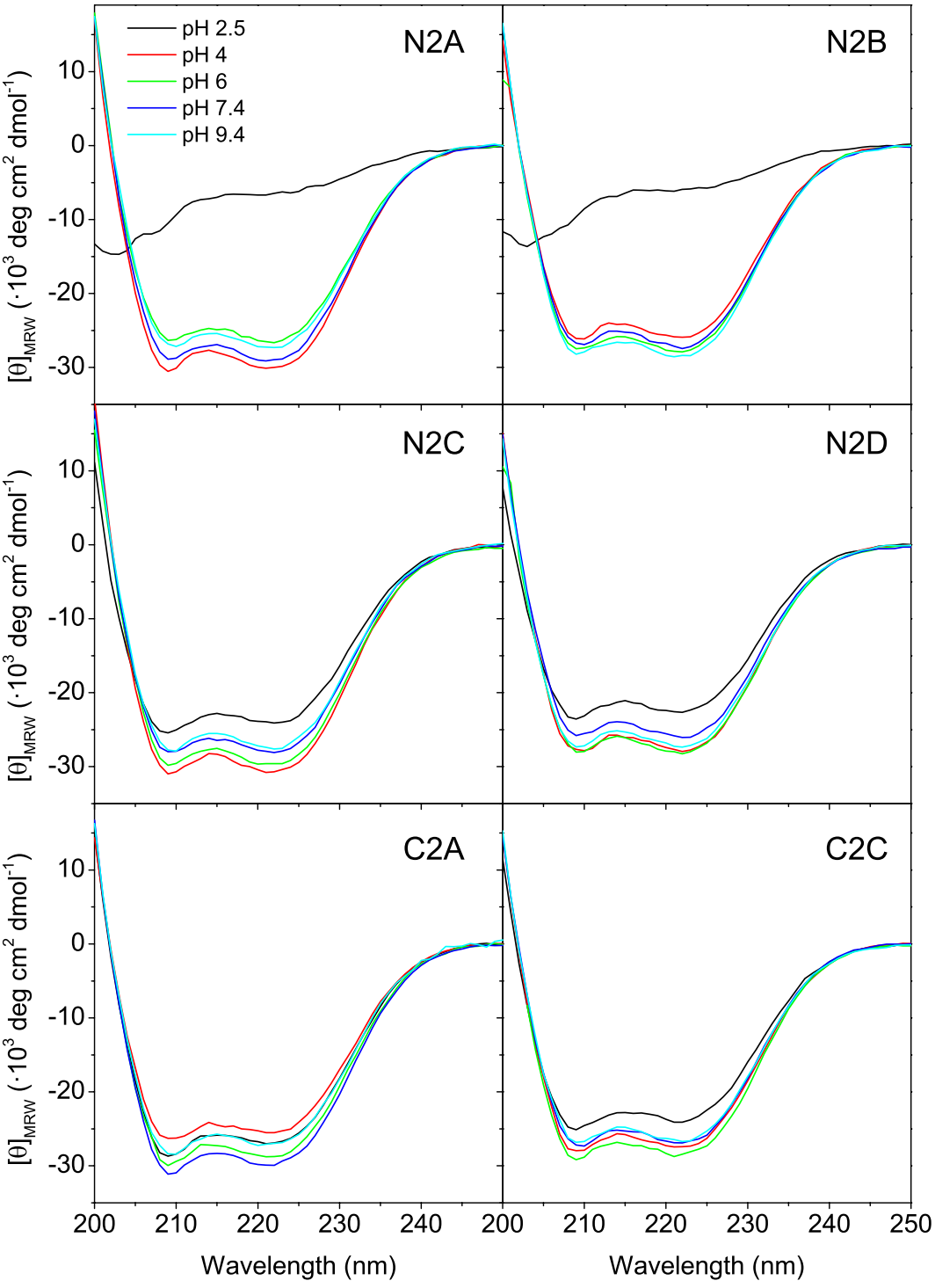


Figure S2: Far-UV CD spectra of the CoVS-HR1 miniproteins at 25°C and different pH. Spectra were recorded at a protein concentration of 15 µM in different buffers (20 mM glycine/HCl pH 2.5; 50 mM sodium acetate pH 4; 50 mM sodium cacodilate pH 6; 50 mM sodium phosphate pH 7.4; 50 mM sodium carbonate pH 9.4). The data are normalized as mean-residue molar ellipticity.

Table S3. Hydrodynamic radii (R_h) of the CoVS-HR1 proteins measured by dynamic light scattering (DLS) at 25°C

	Experimental R_h (nm)					Calculated Rh (monomer)[a]
Protein ↓ pH →	**2.5**	**4.0**	**6.0**	**7.4**	**9.4**	
N2A	3.3	2.3	2.2	2.5	2.2	2.3
N2B	3.0	2.6	2.5	2.6	2.6	2.4
N2C	2.5	2.3	2.3	2.3	2.2	2.3
N2D	2.8	2.6	2.5	2.4	2.6	2.4
C2A	3.1	3.0	3.4	3.3	3.3	2.3
C2C	2.7	2.5	2.7	3.3	3.1	2.3
L3B [b]	3.0	3.6	2.8	3.1	3.4	3.3
L3C [b]	2.9	3.0	3.2	3.0	3.0	3.3

[a] Calculated with HYDROPRO software [2] using the models of the miniproteins.
[b] Data taken from [3].

Table S4. Thermodynamic parameters of unfolding of the CoVS-HR1 miniproteins measured by DSC.

Protein	pH	T_m (°C) [(a)]	$\Delta H_{U,m}$ (kJ mol^{-1}) [(a)]	ΔC_{pU} (kJ mol^{-1}) [(b)]
N2A	2.5	4.5	0.06	5.1
	4	44.6	204	
	6	53.1	247	
	7.4	53.4	252	
	9.4	48.5	261	
N2B	2.5	4.9	-1.4	6.2
	4	42.3	234	
	6	53.2	298	
	7.4	51.9	300	
	9.4	47.7	295	
N2C	2.5	35.5	166	4.3
	4	65.4	301	
	6	76.2	340	
	7.4	81.1	359	
	9.4	72.5	400	
N2D	2.5	34.2	185	4.2
	4	63.1	314	
	6	74.2	351	
	7.4	77.7	368	
	9.4	70.7	391	
C2A	2.5	34.2	185	n.d.
	4	63.1	314	
	6	74.2	351	
	7.4	77.7	368	
	9.4	70.7	391	
C2C	2.5	34.2	185	n.d.
	4	63.1	314	
	6	74.2	351	
	7.4	77.7	368	
	9.4	70.7	391	

(a) Determined from non-linear least-squares fittings using the two-state unfolding model for the N miniproteins and from direct peak integration for the C miniproteins.

(b) Determined from linear fittings of $\Delta H_{U,m}$ vs T_m values for each miniprotein. Data at pH 9.4 were excluded from the linear fittings. n.d.: Not determined

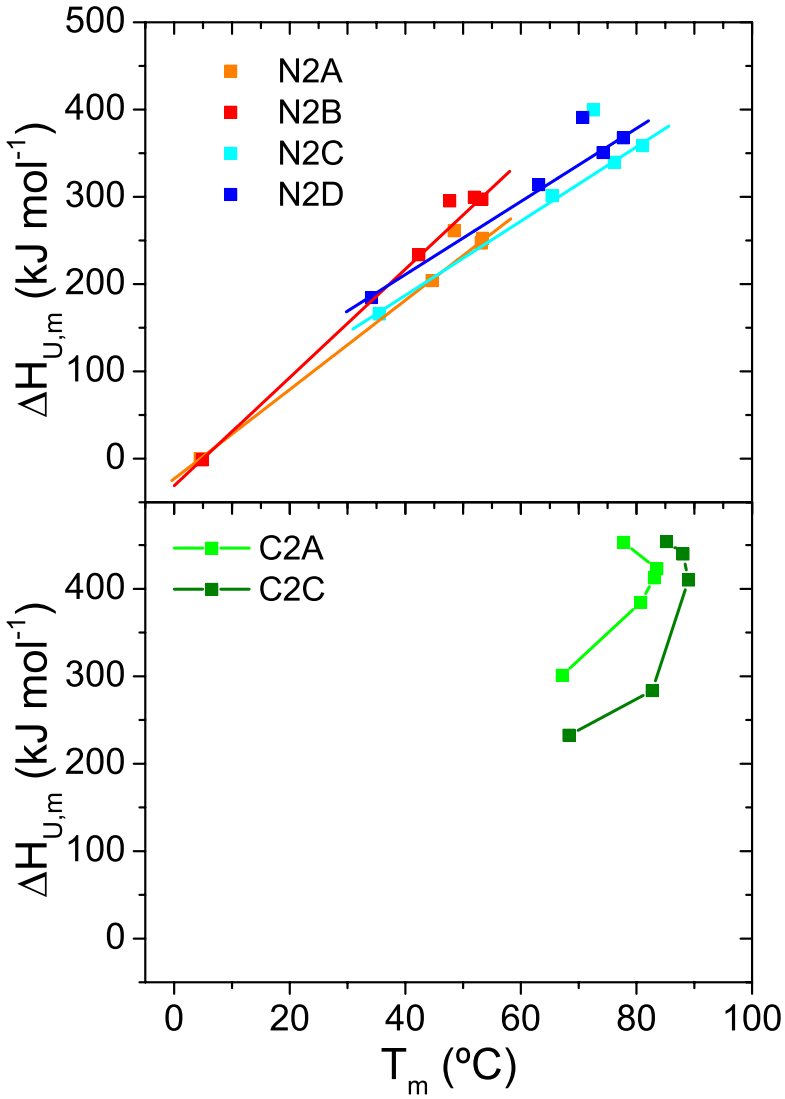


Figure S3. Plots of unfolding enthalpy ($\Delta H_{U,m}$) versus melting temperature (T_m) for the CoVS-HR1 miniproteins. The unfolding of the N-miniproteins follows the two-state model. The lines represent linear fits of the data between pH 2.5 and pH 7.4. The values for the C-miniproteins have been determined by direct integration of the unfolding peaks.

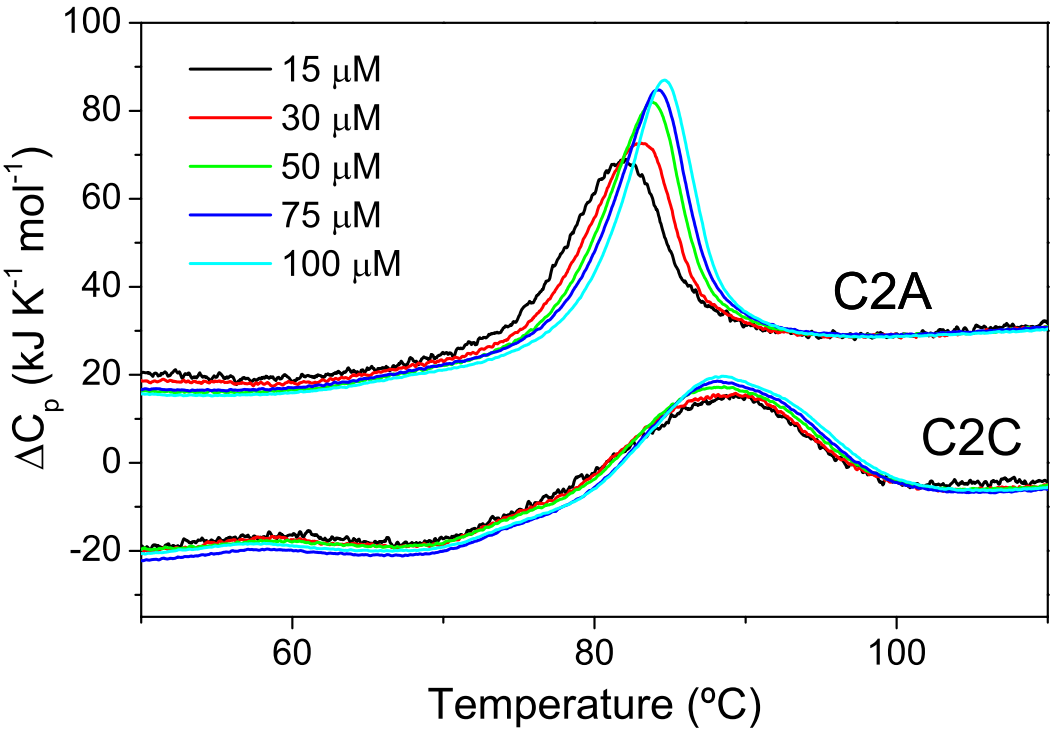
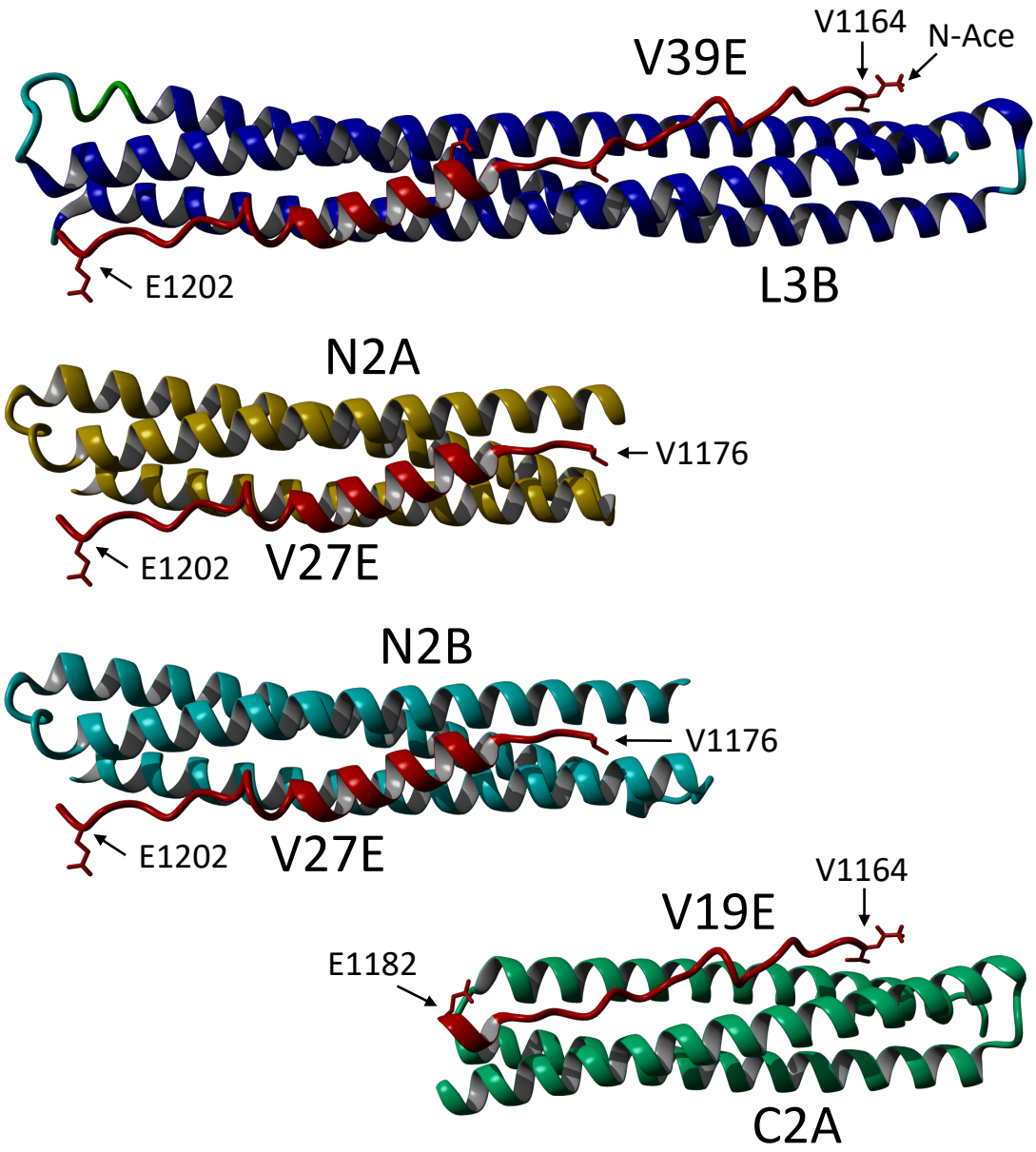


Figure S4. Effect of protein concentration upon the DSC thermograms of the C miniproteins. The experiments were carried out ay pH 7.4, in 50 mM sodium phosphate buffer, at a scan rate of 90 °C/h at the indicated protein concentrations. The curves of C2A have been displaced vertically by 40 kJ K^{-1} mol^{-1} for the sake of clarity.



V39E: **VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE**
V27E: **VVNIQKEIDRLNEVAKNLNESLIDLQE**
V19E: **VDLGDISGINASVVNIQKE**

Figure S5: Model structures of the hypothetical complexes between the CoVS-HR1 miniproteins with HR2-derived peptides. The crystallographic structure of the complex between L3B and V39E (HR2 residues 1164-1202) is also shown for reference (PDB id: 7ZR2 [3]). The model of the complexes between the N2A, N2B and C2A miniproteins and the complementary HR2 peptides (V27E: HR2 residues 1176-1202; V19E: HR2 residues 1164-1182) have been created by structural alignment of each miniprotein with the L3B-V39E complex. The sequences of the HR2 peptides are shown at the bottom. The three peptides contain a C-terminal SGGY tag and are N-acetylated and C-amidated.

Table S5: Thermodynamic parameters of binding of HR2 peptides to the CovS-HR1 proteins measured by ITC at 25 °C.

Protein	HR2 peptide	K_b ($\cdot 10^3$ M^{-1})	K_d (μM)	ΔG_b (kJ mol^{-1})	ΔH_b (kJ mol^{-1})	$T\cdot\Delta S_b$ (kJ mol^{-1})	n
N2A	V27E	270 ± 100	3.7 ± 1.3	–31.0 ± 0.9	–9.2 ± 0.9	+21.8 ± 1.8	1.10 ± 0.05
N2B[(b)]		107 ± 23	9.3 ± 2.0	–29.9 ± 0.6	–68 ± 5	–33 ± 5	0.91 ± 0.04
N2C		340 ± 130	3.0 ± 1.1	–31.6 ± 0.9	+6.1 ± 0.5	+39.3 ± 1.6	1.21 ± 0.05
N2D		360 ± 80	2.8 ± 0.6	–31.7 ± 0.6	+4.6 ± 0.2	+35.7 ± 0.8	1.09 ± 0.03
C2C	V39E	121 ± 11	8.3 ± 0.8	–29.0 ± 0.2	–6.2 ± 0.2	+22.8 ± 0.4	1.02 ± 0.02

[(b)] Parameters measured at 37 °C

Table S6: Thermodynamic parameters of binding of HR2 peptides to the CoVS-HR1 proteins measured by DSC using a model of two-state unfolding coupled to ligand binding (NL ⇔ N + L⇔ U + L).

Protein	HR2 peptide	K_b [(a)] ($\cdot 10^3$ M^{-1})	ΔH_b [(a)] (kJ mol^{-1})	$\Delta C_{p,b}$ [(a)] (kJ K^{-1} mol^{-1})	K_d (μM)
N2A	V27E	34.1 ± 0.6	–47.7 ± 0.5	–3.38 ± 0.04	29.3
N2B		27.6 ± 0.7	–23.4 ± 1.1	–3.45 ± 0.05	36.2
N2C		100 ± 10	–54 ± 5	–2.65 ± 0.25	10
N2D		110.5 ± 4.9	–44.4 ± 1.3	–3.66 ± 0.04	9.1
N2A	V39E	13.6 ± 0.4	–35.9 ± 0.9	–3.64 ± 0.04	73.3
N2B		19.6 ± 0.4	–41.1 ± 0.8	–3.14 ± 0.04	51.0
N2C		27.7 ± 1.1	–35.6 ± 0.8	–4.06 ± 0.06	36.0
N2D		52.2 ± 1.8	–36.2 ± 0.8	–4.90 ± 0.05	19.2

[(a)] Parameters at 37 °C determined from the global fitting of the DSC thermograms in presence of peptides at different peptide:protein ratio. Uncertainties correspond to 95% confidence intervals of the fittings.

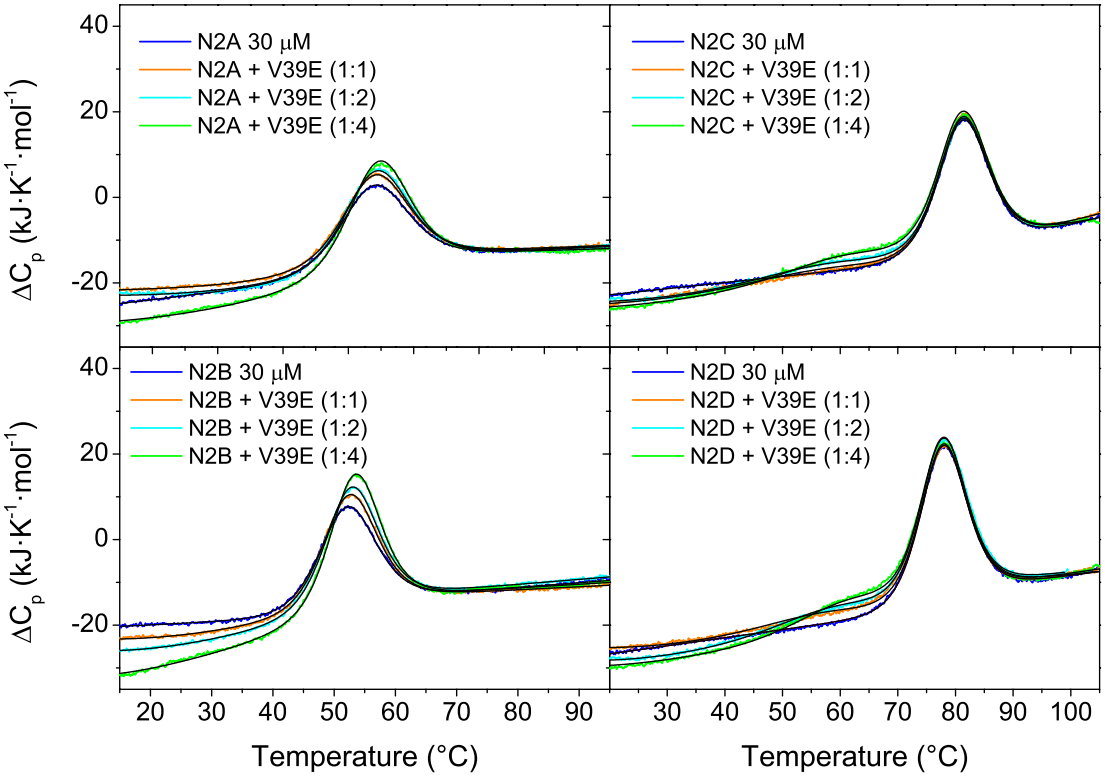


Figure S6: DSC thermograms of the N miniproteins in absence and in presence of three different concentrations of the V39E peptide. Experiments were carried out at pH 7.4 in 50 mM sodium phosphate buffer with a protein concentration of 30 μM. The black curves represent the global fits using a model of binding coupled to unfolding (NL ⇔ N + L⇔ U + L).

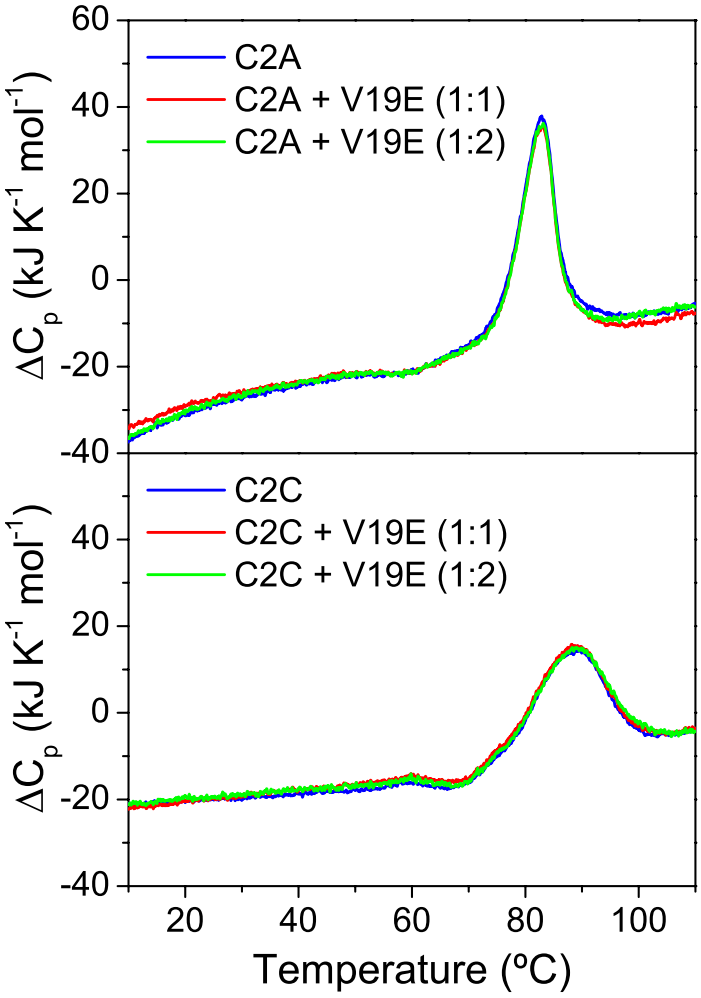


Figure S7: DSC thermograms of C2A and C2C in absence and in presence of two different concentrations of the V19E peptide. Experiments were carried out at pH 7.4 in 50 mM sodium phosphate buffer with a protein concentration of 30 μM.

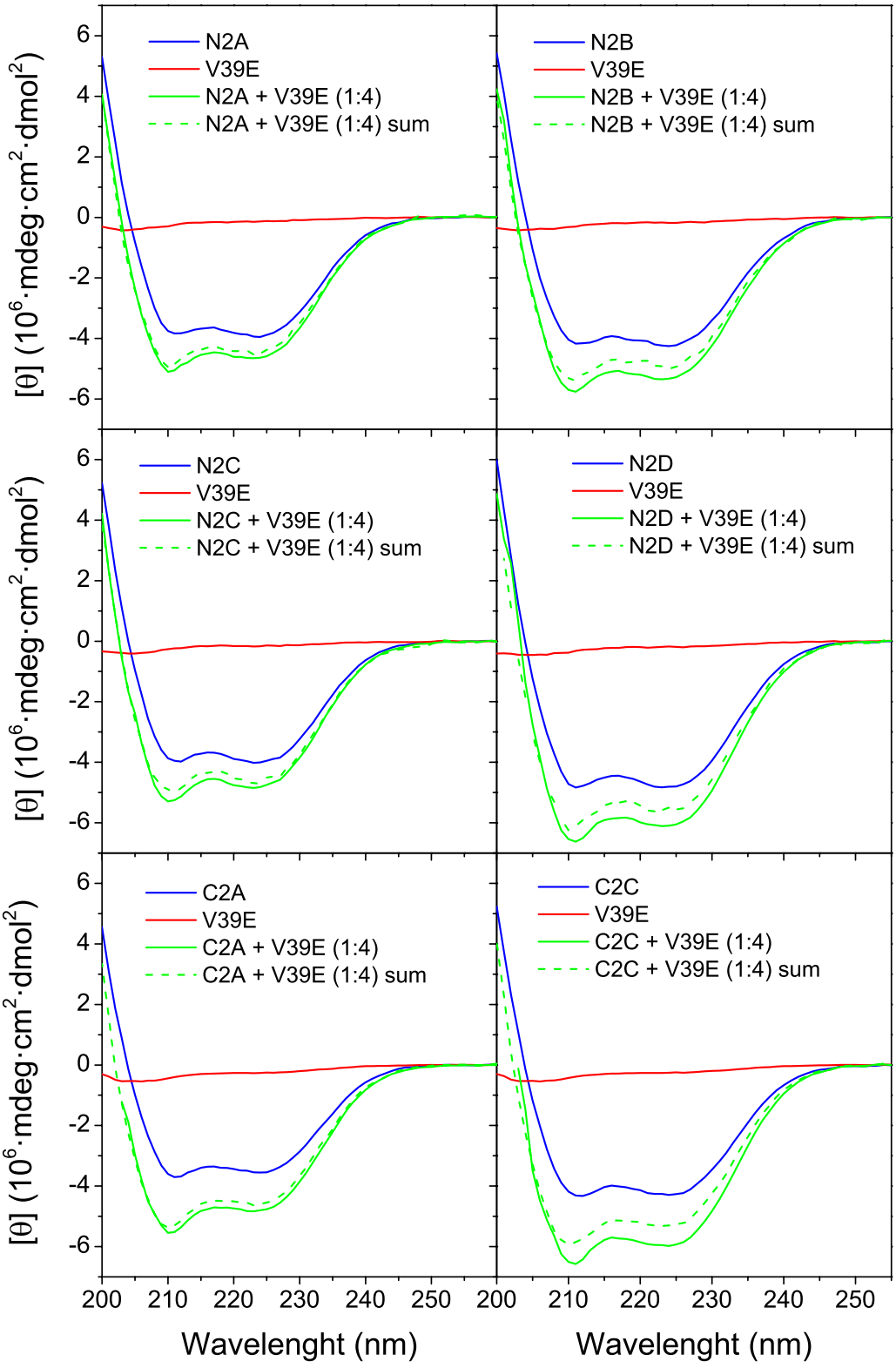


Figure S8: Far-UV CD spectra of the CoVS-HR1 miniproteins in presence of a four-fold molar excess of V39E peptide. The spectra have been recorded at 25 °C in 50 mM sodium phosphate buffer at pH 7.4. The spectra have been normalized per mole of molecules for a proper comparison. The dashed lines represent the spectra that would be obtained for mixtures of non-interacting molecules.

References

[1] F. Madeira, M. Pearce, A.R.N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, Search and sequence analysis tools services from EMBL-EBI in 2022, Nucleic Acids Res. 50 (2022) W276–W279. https://doi.org/10.1093/nar/gkac240.

[2] A. Ortega, D. Amorós, J. García de la Torre, Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models, Biophys. J. 101 (2011) 892–898. https://doi.org/10.1016/j.bpj.2011.06.046.

[3] M. Cano-Muñoz, D. Polo-Megías, A. Cámara-Artigas, J.A. Gavira, M.J. López-Rodríguez, G. Laumond, S. Schmidt, J. Demiselle, S. Bahram, C. Moog, F. Conejero-Lara, Novel chimeric proteins mimicking SARS-CoV-2 spike epitopes with broad inhibitory activity, Int. J. Biol. Macromol. 222 (2022) 2467–2478. https://doi.org/10.1016/j.ijbiomac.2022.10.031.