Supporting Information to:

Single-chain protein mimetics of the N-terminal heptad-repeat region of gp41 with potential as anti HIV-1 drugs

Sara Crespillo[a], Ana Cámara-Artigas[b], Salvador Casares[a], Bertrand Morel[a], Eva S. Cobos[a], Pedro L. Mateo[a], Nicolas Mouz[c], Christophe E. Martin[c], Marie G. Roger[c], Raphaelle El Habib[d], Bin Su[e], Christiane Moog[e] and Francisco Conejero-Lara[a]

[a] Departamento de Química Física e Instituto de Biotecnología, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

[b] Department of Chemistry and Physics, University of Almería, Agrifood Campus of International Excellence (ceiA3), Carretera de Sacramento s/n, 04120 Almería, Spain

[c] PX'Therapeutics, 7 parvis Louis Néel – BP 50 MINATEC/BHT-52B, 38040 Grenoble Cedex 9, France

[d] Sanofi Pasteur S.A., Campus Mérieux, 1541, avenue Marcel Mérieux, 69280 Marcy l'Etoile, France

[e] INSERM U1109, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, 3 rue Koeberlé, 67000 Strasbourg, France.

SI Methods:

Protein expression and purification

The DNA encoding the protein sequences was synthesized by Geneart (Regensburg, Germany) and inserted in a pM1800 vector provided by Sanofi Pasteur. To facilitate purification by Ni-Sepharose affinity chromatography, the protein sequences were histidine tagged at the C-terminus with the sequence GGGGSHHHHHH. For protein expression, BLR(DE3) or BL21(DE3) *E. coli* cells (Novagen, reference: 69053) were transformed with the plasmids and cultured in auto-inducible medium in the presence of

30 μg·mL^{-1} of Kanamycin for 4 hours at 37°C and for 20 h more at 25°C, 30°C or 37°C for optimal expression, depending of the specific protein. Cells were collected by centrifugation, resuspended in lysis buffer (50 mM Tris pH 7.5 containing 500 mM NaCl, 1 μg·mL^{-1} pepstatine, 1 μg·mL^{-1} leupeptine, 1 mM β-mercaptoethanol and 1 mM EDTA) and lysed by disintegration in a high-pressure homogenizer. The cell lysates were incubated for 30 min with 250 U/L of benzonase and the soluble and insoluble fractions were separated by ultracentrifugation at 15900 g for 1 h at 4°C. The insoluble fraction was discarded for all constructs except for covNHR2, which was purified from inclusion bodies to optimize production yield. In this case the protein was solubilized in 50 mM Tris pH 7.5 buffer containing 500 mM NaCl, 1 mM β-mercaptoethanol and 6 M guanidine under shaking at 8°C overnight.

The constructs were purified with a Ni-Sepharose Fast Flow column (Amersham GE Healthcare), previously equilibrated in 50 mM Tris pH 7.5 buffer, 1 mM β-mercaptoethanol, 500 mM NaCl. The equilibration buffer contained additionally 8 M urea and 1 mM Glycine in the case of covNHR3 or 500 mM urea and 0.1% Chaps in the cases of covNHR3A and covNHR3B constructs. The proteins were eluted using an imidazole gradient (0 - 1 M) in the same equilibration buffer. The pooled fractions were dialyzed against 50 mM Tris pH 7.5 buffer containing 25 mM NaCl. The proteins were then further purified by an ionic exchange chromatography step on a SP Sepharose column (Amersham GE Healthcare), with a 0 – 1 M NaCl gradient in the same buffer. Sample concentrations were determined by UV absorption measurement at 280 nm. Protein identity and purity was assessed using SDS-PAGE, Western blotting, mass spectrometry and N-terminal sequencing.

Molecular size characterization
The molecular size of the proteins was measured by light scattering using a DynaPro MS-X instrument (Wyatt, Santa Barbara, CA). Prior to the measurements, the sample solutions were centrifuged for 30 min at 14000 rpm and 4°C in a bench microfuge. The measurements were made in a 30 μL quartz sample cuvette. Sets of dynamic light scattering (DLS) data were acquired by averaging 50 acquisitions of an acquisition time of 10 s each. The intensity autocorrelation curves were analyzed with the Dynamics V6 software to obtain the hydrodynamic radius distributions. To determine the weight-

averaged molecular mass, scattering intensities were measured at different protein concentrations and analyzed using the Debye plot as described in the Fig S1.

Circular dichroism

CD measurements were performed with a Jasco J-715 spectropolarimeter (Tokyo, Japan) equipped with a thermostatized cell holder. Measurements of the far-UV CD spectra (260-200 nm) were made with a 1 mm path length quartz cuvette. Spectra were recorded at a scan rate of 100 nm/min, 1 nm step resolution, 1 s response and 1 nm bandwidth. The resulting spectrum was usually the average of 5 scans. Near-UV CD spectra (350-250 nm) were measured using a 5 mm cuvette and were typically the average of 20 scans. Each spectrum was corrected by baseline subtraction using the blank spectrum obtained with the buffer and finally the CD signal was normalized to molar ellipticity ($[\theta]$, in deg·dmol^{-1}·cm^{2}). The percentage of α-helical structure was estimated from the far-UV CD spectra as described elsewhere (1).

Differential scanning calorimetry

The thermal stability of the protein constructs was characterized by differential scanning calorimetry using a VP-DSC microcalorimeter (Microcal Inc., Northampton, MA). Scans were run from 5 to 125°C at a scan rate of 90°C·h^{-1}. Protein concentration was typically 20 µM. Baselines recorded with buffer in both cells were obtained before each experiment and subtracted from the experimental thermograms obtained with the samples. Reheating runs were carried out to determine the thermal reversibility of the protein denaturation process, with the exact same parameters used for the main scan. After correction of the dynamic response of the instrument, the partial molar heat capacity (C_p) was calculated from the experimental DSC thermograms using the instrument's software.

Limited proteolysis

CovNHR3-ABC (50 µM) and T20 (100 µM) were submitted to proteolysis by thermolysin (Sigma-Aldrich) at 37ºC in 50 mM Hepes pH 7.2, 5 mM CaCl$_2$, 0.1 M NaCl. Proteolysis was stopped by addition of 20 mM EDTA and cooling on ice. Samples were then analyzed by SDS-PAGE and fragments were separated by reversed-phase HPLC and identified by electrospray-ionization mass spectrometry in a Bruker Autoflex TOF analyzer.

Peptide titrations

The interaction between the covNHR constructs and the gp41 CHR region was investigated by mixing each covNHR protein with exogenous synthetic peptides encompassing different segments the CHR region. Four peptides were assayed: T20 (gp41 sequence, 127-162), Q32Q (gp41 sequence 110-141), W34L (also known as C34, sequence 117-150), and Q20S (sequence 110-129). T20 (Fuzeon®) was purchased from Roche Pharma and purified additionally by reversed-phase HPLC in a C-18 Vydac column, using a water-acetonitrile gradient. The other synthetic peptides were purchased from EZ Biolab Inc. (Carmel, IN, USA). Peptides were normally used in N-acetylated and C-amydated form but unprotected peptides were also used in ITC experiments (see below). Titration experiments monitored by near-UV CD were performed to estimate the binding affinity of the covNHR constructs for each CHR peptide. Binding was assessed by monitoring the recovery of a negative near-UV CD band at 293 nm, which is characteristic of the interaction between the CHR region and the conserved hydrophobic pocket of NHR (2). Two stock solutions of the covNHR constructs were prepared at 50 µM in 50 mM phosphate buffer, pH 7.4: one without CHR peptide and another containing CHR peptide concentration of about 500 µM. The two stock solutions were mixed at the appropriate proportions to reach the desired peptide:protein ratios and to maintain constant the protein concentration during the titration. The mixtures were incubated for at least 1 h prior to their analysis. The near-UV CD spectra of each mixture were recorded and normalized ellipticity data at 293 nm were plotted versus the peptide concentration and fitted to the equations according to a 1:1 binding model.

Isothermal titration calorimetry

A thermodynamic characterization of the binding between the covNHR constructs and the CHR peptides was made by isothermal titration calorimetry (ITC). ITC experiments were carried out in a VP-ITC microcalorimeter from MicroCal Inc. (Northampton, MA). The protein solutions were titrated with 28 injections of the peptide solution using a profile of variable injection volumes, at 480 s intervals. Concentrations of the protein used for the titrations were in the range of 12-20 µM, while the ligands in the syringe were typically at 0.3-0.5 mM. The experiments were carried out in 50 mM phosphate

buffer, pH 7.4. All experiments were performed at 25°C. As a blank, an independent experiment with only buffer in the calorimeter's cell was performed with the same peptide solution to determine the corresponding heats of dilution.

The experimental thermograms were baseline corrected and the peaks were integrated to determine the binding heats produced by each ligand injection. Finally, each heat was normalized per mole increase in the total ligand concentration. The resulting binding isotherm was fitted using an appropriate binding model as described in the results, allowing the determination of the binding constant, K_b, the enthalpy change, ΔH_b, and the binding stoichiometry, n, for each interaction. From these values, the Gibbs energy and entropy of binding could be derived as $\Delta G_b = -RT \ln K_b$ and $T \cdot \Delta S_b = \Delta H_b - \Delta G_b$.

X-ray crystallography.

Initial screening of crystallization conditions was performed by the sitting-drop vapour-diffusion method using the Crystal Screen HT kit (Hampton Research). 2 μL of protein solution (15 mg·mL^{-1} in 10 mM Tris buffer pH 8.0) were mixed with an equal volume of the reservoir solution and vapor equilibrated against 0.2 ml of the precipitant solution at 25 °C. Best crystals were obtained at pH 4.6 with 2 M sodium chloride and 0.1 M sodium acetate trihydrate as precipitant.

Crystals were transferred into a cryoprotectant solution containing the reservoir solution and 10% PEG300 before flash-cooling in liquid nitrogen. Data sets were collected at 100 K on the ID-29 beamline at the ESRF (Grenoble, France) and the Xaloc beamline at ALBA (Barcelona, Spain). Data were indexed and integrated with the program XDS (3, 4). Data scaling was performed using the program Aimless (5) from de CCP4 suite (6). The solution and refinement of the structures were performed using the suite PHENIX (7). Crystals of covNHR-ABC belonged to the R32 space group with one molecule in the asymmetric unit (AU). Molecular-replacement phasing using PHASER (8) was performed with the coordinates of the model generated from the crystallographic structure of the gp41 mimetic 5-Helix in complex with an antibody (PDB code 2XRA) (9), modified with the inverted α-helix. Manual model-building was performed using COOT (10). Refinement was performed using phenix.refine in PHENIX (11). The final cycles of refinement were performed using TLS refinement (12). Quality of the structure was checked using MOLPROBITY (13) and POLYGON (14). The crystallographic parameters and statistics of the data collection and the refinement statistics are collected in Supplementary Table S5.

HIV inhibition assays

The HIV inhibitory activity of the covNHR constructs was studied using two neutralizing assays: i) on TZM-bl infected with HIV-1 SF-162 pseudoviruses (classified as tier 1) and ii) on PBMC using HIV-1 primary isolates from various subtypes and sensitivity to neutralization (15). Inhibition of virus replication was analyzed as a reduction in Tat-induced luciferase reporter gene expression after a single round of infection of TZM-bl. The detection of inhibitory activity was assessed at different concentrations of inhibitor. Neutralization of primary isolates was performed as previously described (15, 16). Various primary isolates were used: HIV-1 SF162 (Tier 1, subtype B, R5), QH0692 (Tier 2, subtype B, R5), TV01 (Tier 1, subtype C, R5), KON (Tier 2/3, subtype CRF02 AG, X4), 92UG024 (subtype D, X4), 89.6 (subtype B, X4R5), Du174 (subtype C, R5), 92Br025 (subtype C, R5). These viruses have been obtained through the Program EVA Centre for AIDS Reagents (CFAR, NIBSC, UK) and their inhibitory characteristics have been extensively analyzed in a comparative study (15). Inhibition of the percentage of infected CD4 T lymphocytes was measured by intracellular p24 staining. Samples were analyzed at different concentrations for the detection of the percentage of infected cells after a single cycle of infection. The concentration of inhibitor inducing an 80% reduction of the percentage of infected cells (IC80) was recorded.

Table S1: Amino acid sequences of the covNHR constructs designed in this work. Residues in bold correspond with mutations engineered in the 1st generation constructs. Residues in bold and underlined correspond to mutations engineered for the 2nd generation constructs. All constructs were expressed in E. coli with a N-terminal methionine and a C-terminal His-tag of sequence GGGGSHHHHHH.

WT gp41	NHR sequence 30-81
	ARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
	Helix 1
CovNHR1	ARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR2	ARQ**E**LSGI**E**Q**K**QNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3	ARQ**E**LSGI**E**Q**K**QNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-A	ARQELSGIEQKQNNLLRAIEAQQHLLQLTVW**K**IKQLQARILAVERYLKDQQL
CovNHR3-AC	ARQELSGIEQKQNNLLRAIEAQQHLLQLTV**SK**IKQLQARILAVERYLKDQQL
CovNHR3-B	ARQELSGIEQKQNNLLR**Q**IEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-AB	ARQELSGIEQKQNNLLR**Q**IEAQQHLLQLTVW**K**IKQLQARILAVERYLKDQQL
CovNHR3-ABC	ARQELSGIEQKQNNLLR**Q**IEAQQHLLQLTV**SK**IKQLQARILAVERYLKDQQL
	Loop 1
CovNHR1	GKAD
CovNHR2	GKAD
CovNHR3	GKGNQ
CovNHR3-A	GKGNQ
CovNHR3-AC	GKGNQ
CovNHR3-B	GKGNQ
CovNHR3-AB	GKGNQ
CovNHR3-ABC	GKGNQ
	Helix 2
CovNHR1	PQQDKLYREVALIRAQLQKIGWVTLQLLHQQAEIARLLNNQQQVIGSLLQR
CovNHR2	PQQDKLYREVALIRAQLQKIGWVTLQLLHQQAEIARLLNNQ**EQE**IGSL**K**QR
CovNHR3	PQQDKLYREVALIRAQLQKIGWVTLQLLHQQAEIARLLNNQ**EQE**IGSL**K**QR
CovNHR3-A	PQQDKLYREVALIRAQLQKI**EWE**TLQLLHQQAEIARLLNNQEQEIGSLKQR
CovNHR3-AC	PQQDKLYREVALIRAQLQKI**ESE**TLQLLHQQAEIARLLNNQEQEIGSLKQR
CovNHR3-B	PQQDKLYREVALIRAQLQKIGWVTLQLLHQQAEI**ERE**LNNQEQEIGSLKQR
CovNHR3-AB	PQQDKLYREVALIRAQLQKI**EWE**TLQLLHQQAEI**ERE**LNNQEQEIGSLKQR
CovNHR3-ABC	PQQDKLYREVALIRAQLQKI**ESE**TLQLLHQQAEI**ERE**LNNQEQEIGSLKQR
	Loop 2
CovNHR1	GILD
CovNHR2	GILD
CovNHR3	GLIDG
CovNHR3-A	GLIDG
CovNHR3-AC	GLIDG
CovNHR3-B	GLIDG
CovNHR3-AB	GLIDG
CovNHR3-ABC	GLIDG
	Helix 3
CovNHR1	PLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR2	PLLSGI**D**Q**R**QNNL**K**RAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3	PLLSGI**D**Q**R**QNNL**K**RAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-A	PLLSGIDQRQNNLKRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-AC	PLLSGIDQRQNNLKRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-B	PLLSGIDQRQNNLKRAIEAQ**K**HLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-AB	PLLSGIDQRQNNLKRAIEAQ**K**HLLQLTVWGIKQLQARILAVERYLKDQQL
CovNHR3-ABC	PLLSGIDQRQNNLKRAIEAQ**K**HLLQLTVWGIKQLQARILAVERYLKDQQL

Table S2. Apparent hydrodynamic radius (R_h) of covNHR proteins determined by dynamic light scattering (DLS) in 50 mM sodium phosphate buffer pH 7.4 and 50 mM glycine buffer pH 2.5, at different protein concentrations. The apparent R_h values of all constructs at acid pH are highly consistent with the value of 2.65 nm calculated for a rigid model of covNHR2 using the software HYDROPRO (17) indicating a monomeric state. In contrast, at pH 7.4 the R_h values of some of the constructs are considerably higher. For covNHR3 the R_h values increase with the protein concentration, suggesting self-association into oligomeric species at pH 7.4. The approximate uncertainty of the R_h values is 0.5 nm.

Construct	Apparent R_h of the main component (nm)	
	pH 2.5	pH 7.4
covNHR2	2.6	3.1
covNHR3	2.6	3.7
covNHR3-A	2.5	2.8
covNHR3-B	2.6	3.5
covNHR3-AB	2.6	2.9
covNHR3-AC	2.5	3.0
covNHR3-ABC	2.6	2.7
[covNHR3] (µM)		
9	2.8	3.3
18.5	2.8	3.5
37.5	2.4	3.6
75	2.5	4.2
135	2.4	4.5

Table S3: Thermodynamic parameters for the thermal unfolding of the covNHR constructs derived from the fitting of the DSC thermograms using a model of two sequential transitions (N $\rightleftarrows$ I $\rightleftarrows$ U). Uncertainties of the parameters correspond to the standard error of the fits.

Construct	pH	$T_{m,1}$ (°C)	$\Delta H_{m,1}$ (kJ mol^{-1})	$T_{m,2}$ (°C)	$\Delta H_{m,2}$ (kJ mol^{-1})	$\Delta h^{(a)}$ (kJ res^{-1})	Maximum fraction of intermediate I[b]
covNHR2	2.5	81.3±0.2	172±1	87.6±0.02	344±2	3.17	0.51
	3.0	85.3±0.2	162±2	86.8±0.04	369±2	3.26	0.39
	4.0	94.9±0.1	188±1	102.3±0.04	330±2	3.17	0.53
covNHR3	2.5	83.5±0.2	220±2	86.6±0.03	351±3	3.54	0.43
	3.0	87.6±0.2	190±3	84.7±0.06	356±4	3.35	0.27
	4.0	95.2±0.2	195±2	99.3±0.04	348±2	3.33	0.45
	5.0	92.7±2	207±2	110.1±0.06	370±2	3.53	0.79
covNHR3-A	2.5	89.6±0.13	195±1	95.0±0.02	392±1	3.60	0.50
covNHR3-B		86.8±0.06	374±3	89.6±0.06	353±3	4.46	0.44
covNHR3-AB		93.7±0.09	272±4	91.1±0.06	329±5	3.69	0.26
covNHR3-AC		91.5±0.2	215±4	88.9±0.07	345±5	3.44	0.27
covNHR3-ABC	2.5	91.7±0.2	211±6	87.1±0.1	394±7	3.71	0.23
	3.0	97.4±0.6	263±11	83.1±0.15	321±11	3.58	0.07
	4.0	114.8±1.2	221±22	100.1±0.2	367±21	3.61	0.09
	5.0	116.9±0.6	158±12	108.8±0.3	412±14	3.50	0.21

[a]Total specific enthalpy calculated as the sum of the enthalpies of the two transitions and normalized by the number of residues, excluding the histidine tag.

[b]Estimated from the parameters of the fits.

Table S4. Apparent dissociation constants of the interaction between covNHR proteins and CHR peptides. Titration experiments were monitored by near-UV CD. Binding was assessed by monitoring the recovery of the negative near-UV CD band at 293nm. Normalized ellipticity data at 293nm were plotted versus the peptide concentration and fitted to the following equations according to a 1:1 binding model. For each titration point the ellipticity is given by:

$$[\theta] = [\theta]_0 + \Delta[\theta] \cdot f \qquad (2)$$

where $[\theta]_0$ is the ellipticity signal of the free protein, f the fraction of protein the complex with the peptide and $\Delta[\theta]$ the máximum ellipticity change. The value of f is determined from the equilibrium dissociation constant, K_d and the total concentrations of protein, M_T, and ligand, L_T, for each titration point:

$$f = \frac{(K_d + M_T + L_T) - \sqrt{(K_d + M_T + L_T)^2 - 4 \cdot M_T \cdot L_T}}{2 \cdot M_T} \qquad (3)$$

Protein	CHR peptide	pH	Apparent K_D (µM)
covNHR2	W34L	2.5	7.4 ± 1.4
	Q32Q	2.5	9 ± 3
covNHR3	Q32Q	2.5	11.9 ± 2.3
	W34L	7.4	2.4 ± 0.4
	Q32Q	7.4	2.6 ± 0.5
	Q20S	7.4	17 ± 3
covNHR3-ABC	W34L	7.4	1.5 ± 0.3
	Q32Q	7.4	2.5 ± 1.3
	Q20S	7.4	15.9 ± 2.0

Table S5. Thermodynamic magnitudes for the binding of the W34L peptide (with protected and unprotected termini) to covNHR3 and covNHR3-ABC at pH 7.4 and 25°C. The magnitudes have been derived from the analysis of ITC binding isotherms using a model of several sets of different and independent sites. The errors correspond to 95% confidence intervals of the fits.

Protein	K_b (×10^{-6}) [a] (M^{-1})	K_d (μM)	ΔH_b (kJ mol^{-1})	n [b]	ΔG_b (kJ mol^{-1})	$T \cdot \Delta S_b$ (kJ mol^{-1})
			W34L			
covNHR3	**4.4 ± 0.5**	**0.23 ± 0.03**	**-51.7 ± 0.2**	**2.37 ± 0.02**	**-38.0 ± 0.3**	**-13.7 ± 0.5**
	0.028 ± 0.005	35 ± 6	-43 ± 3	1 (f)		
covNHR3-ABC	**25 ± 5**	**0.040 ± 0.008**	**-47.6 ± 0.1**	**1.22 ± 0.01**	**-42.2 ± 0.6**	**-5.4 ± 0.6**
	0.017 ± 0.008	59 ± 28	-6.3 ± 1.3	2 (f)		
			W34L$_{un}$			
covNHR3	**14 ± 9**	**0.07 ± 0.05**	**-19.5 ± 0.2**	**1 (f)**	**-40.8 ± 1.6**	**21.3 ± 1.8**
	0.10 ± 0.06	10 ± 6	-25 ± 8	1 (f)		
	0.001 ± 0.004	1000 ± 4000	-297 ± 1461	1 (f)		
covNHR3-ABC	**5.5 ± 1.1**	**0.18 ± 0.04**	**-20.47 ± 0.16**	**1.22 ± 0.01**	**-38.5 ± 0.5**	**18.0 ± 0.7**
	0.007 ± 0.003	140 ± 60	-21 ± 4	2 (f)		

[a] Values of the high affinity sites have been highlighted in bold.

[b] The number of sites marked with (f) had to be fixed in the fittings.

Table S6. X-ray data collection and refinement statistics for covNHR3-ABC structure determination

Wavelength (Å)	0.96869
Resolution range (Å)	17.8 - 3.1 (3.2 - 3.1)
Space group	R32
Unit cell	57.956 57.956 257.671 90 90 120
Total reflections	28217 (4637)
Unique reflections	3278 (582)
Multiplicity	8.6 (8.0)
Completeness (%)	99.6 (100.0)
Mean I/sigma(I)	130.8 (41.9)
Wilson B-factor	43.97
R-merge %	5.3 (18.1)
R-work %	24.69 (34.72)
R-free %	28.93 (43.63)
RMS (bonds) Å	0.005
RMS (angles) degrees	0.59
Ramachandran favored (%)	99
Ramachandran outliers (%)	0
Average B-factor	72.50

Statistics for the highest-resolution shell are shown in parentheses.
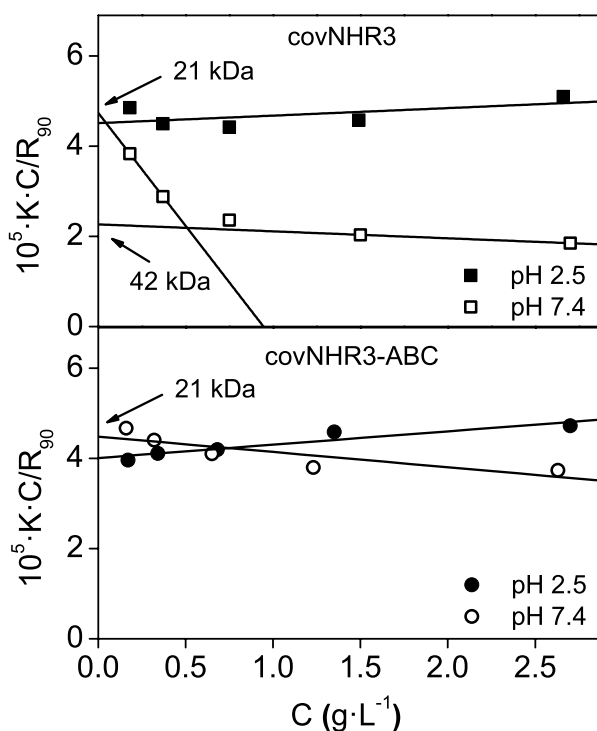
Figure S1: Debye plots for the static light scattering (SLS) intensity of covNHR3 and covNHR3-ABC solutions measured at different protein concentrations at pH 2.5 and 7.4. Lines are drawn only for visual purposes. Static scattering intensities were analyzed using equation 1, valid for particles significantly smaller that the wavelength of the incident radiation:

$$\frac{K \cdot C}{R_\theta} = \frac{1}{M_w} + 2 \cdot A_2 \cdot C \qquad (1)$$

where the K is an optical constant of the instrument, C is the particle mass concentration, R_Θ is the Rayleigh ratio of scattered to incident light intensity, M_w is the weighted-average molecular weight, A_2 is the 2nd virial coefficient that is representative of inter-particle interaction strength. M_w can be determined from the intercept. At pH 2.5 the plots are quite linear for both covNHR3 and covNHR3-ABC and the intercepts correspond approximately to 21 kDa, very consistent with the molar mass of the protein monomers (20.1 kDa). At pH 7.4 the plot of covNHR3 has a larger dependence with the concentration and considerable curvature, suggesting a self-association process. Extrapolation of the data in the low concentration range is consistent with the molecular weight of a monomer, whereas a similar extrapolation in the high concentration range indicates a dimer. In the case of covNHR3-ABC at pH 7.4 the Debye plot is much more linear indicating an essentially monomeric state.
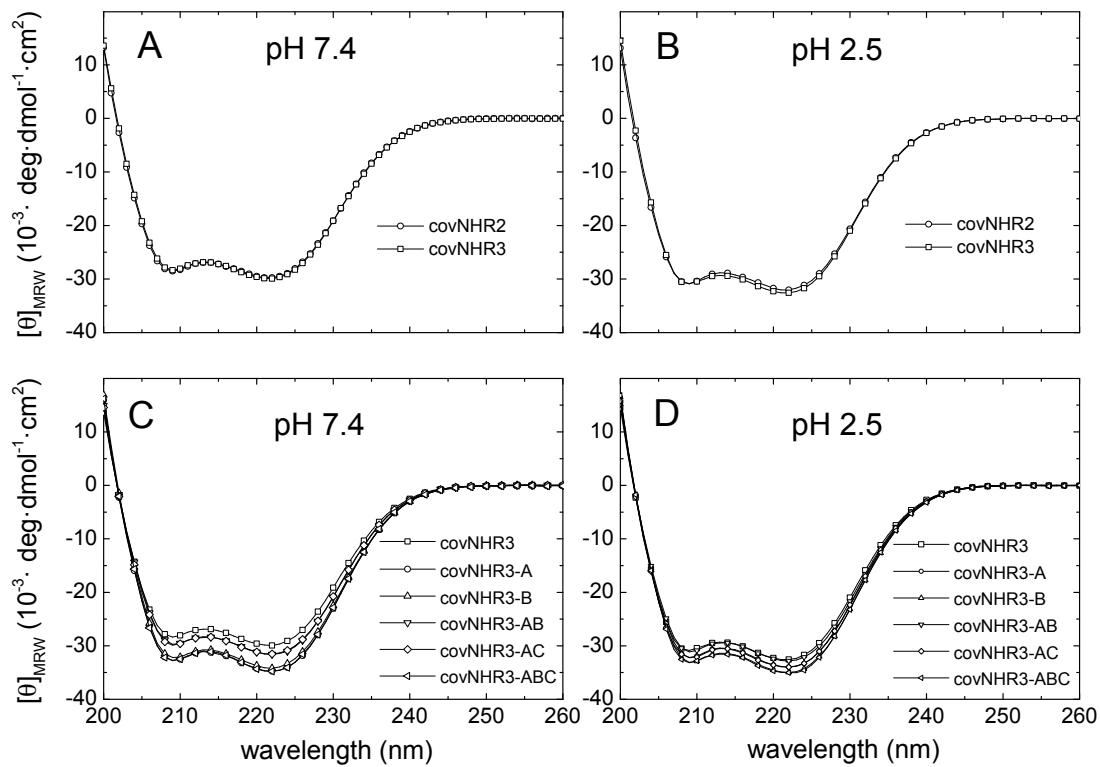
Figure S2: Secondary structure of covNHR proteins. **A)** and **B)** Far-UV CD spectra of the covNHR2 and covNHR3 constructs at pH 7.4 and pH 2.5. **B)** and **C)** Far-UV CD spectra of 2nd generation covNHR constructs at pH 7.4 and pH 2.5 in comparison with covNHR3. All spectra have been obtained at 25°C at a concentration of about 15 μM and normalized to mean-residue ellipticity units. Estimated alpha-helical percentages range between 81% and 90%, depending on the pH and the construct.
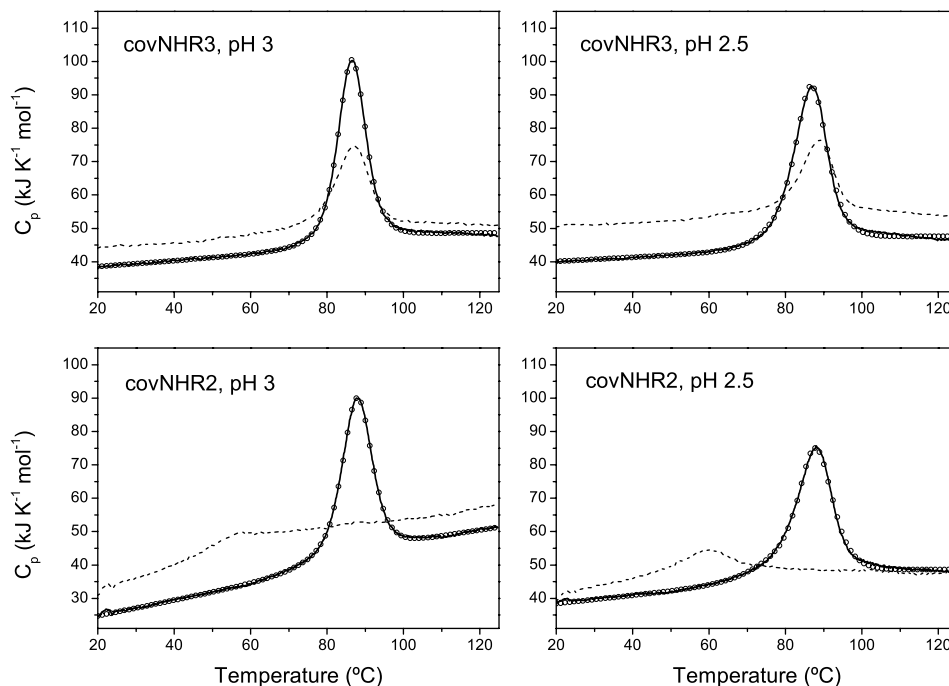
Figure S3: DSC thermograms measured by DSC for the thermal unfolding of covNHR2 and covNHR3 at acid pH. Solid lines represent the experimental first scans. Dashed lines represent the second scans of the same sample after cooling. While the unfolding transitions of covNHR3 are partially reversible by about 50% in a second scan of the same sample, those of covNHR2 are considerably shifted to lower temperature, indicating that this construct cannot refold well after heating up to 125°C. The unfolding transitions are slightly asymmetric and could not be fitted perfectly to a two-state unfolding model (N $\rightleftarrows$ U) but followed very well a model of two sequential transitions (N $\rightleftarrows$ I $\rightleftarrows$ U). Open symbols correspond to the best fits according to the latter model. The thermodynamic parameters derived from these fits are collected in Table S3. The heat capacity changes of unfolding could not be reliably determined from each individual scan due to some baseline distortions.
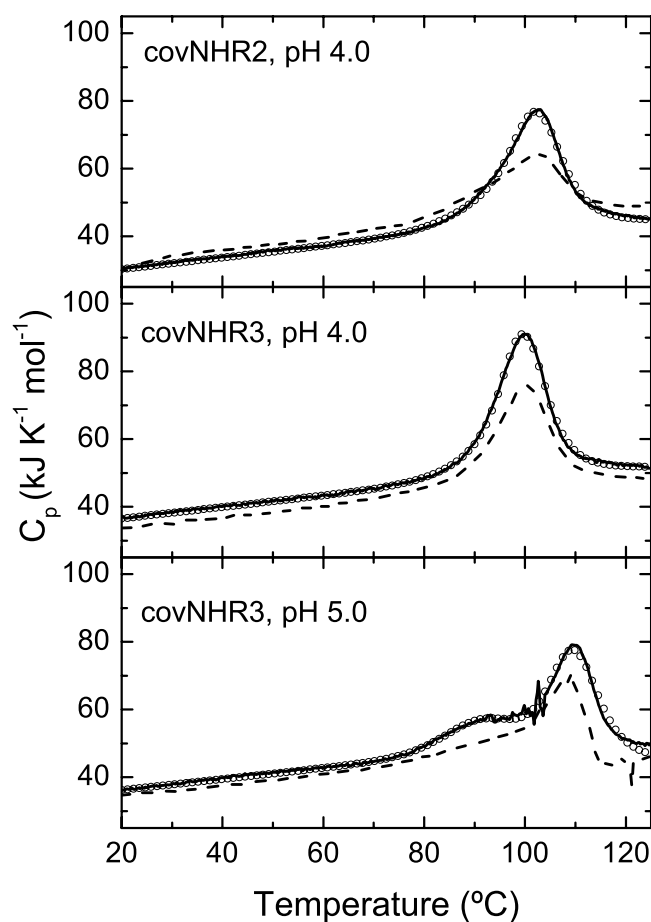
Figure S4: DSC thermograms for the thermal unfolding of covNHR2 and covNHR3 at pH 4.0 and 5.0. The DSC unfolding curves at pH 4.0 are partially reversible but considerably asymmetric. At pH 5.0, the thermal stability increases but two transitions become segregated indicating a decreased structural cooperativity. Solid lines represent the experimental first scans. Dashed lines represent the second scans. Open symbols correspond to the best fits according to a model of two sequential transitions. The thermodynamic parameters are shown in Table S3.
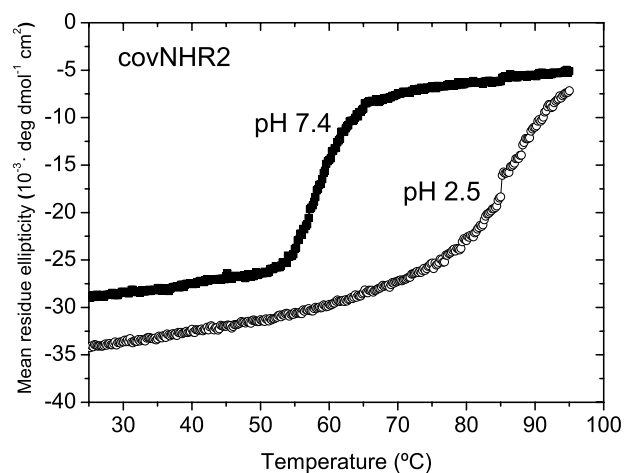
Figure S5: Thermal unfolding of covNHR2 followed by monitoring the CD signal at 222 nm, which is characteristic of the α-helix structure, at pH 2.5 and pH 7.4. The unfolding transitions can be observed as strong decreases in negative ellipticity due to loss of helical structure and can be associated to the unfolding peaks observed by DSC at pH 2.5 (87°C) and pH 7.4 (60°C), respectively. Both experiments were carried out at a protein concentration of 15 μM.
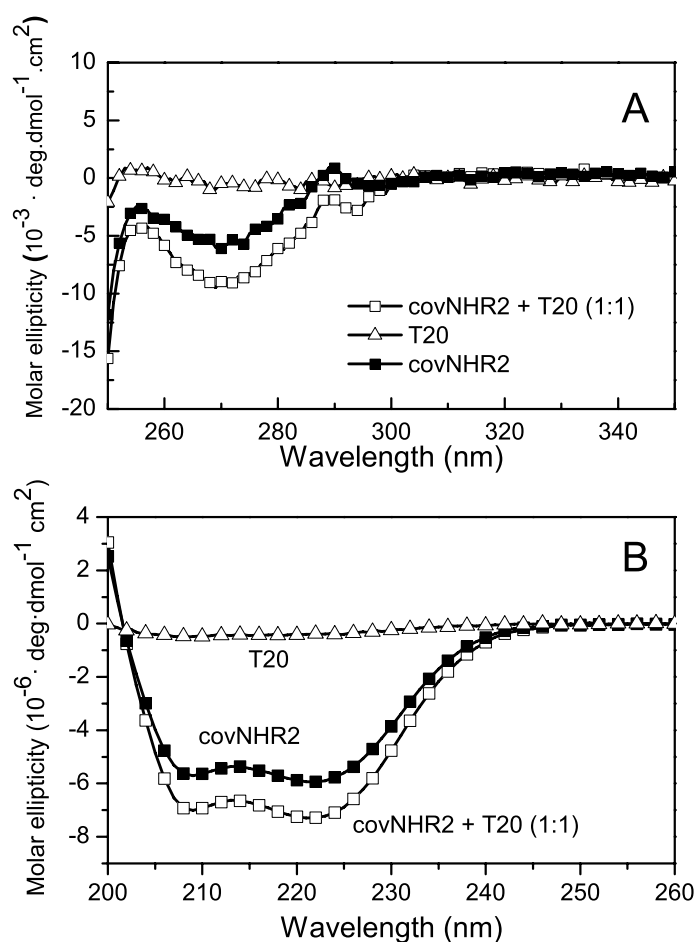
Figure S6: Interaction of covNHR2 with T20. Near-UV CD (A) and far-UV CD (B) spectra at pH 2.5 and 25°C of equimolar mixtures between covNHR2 and T20 compared with the similar spectrum of the isolated moieties. The far-UV CD spectra were recorded at a concentration of 15 μM, whereas the near-UV CD spectra were obtained at 50 μM. The near-UV CD spectra indicate a significant gain in an ellipticity band centered at ca. 270 nm, suggesting the implication of a tyrosine side chain (likely Y127 from T20) in the interaction. The far-UV CD spectrum of the mixture shows a considerable increase in negative ellipticity in comparison with the isolated molecules, indicating acquisition of α-helical structure due to the binding. However, the mixtures were prone to form insoluble precipitates, especially at pH 7.4, precluding a more detailed characterization of the binding.
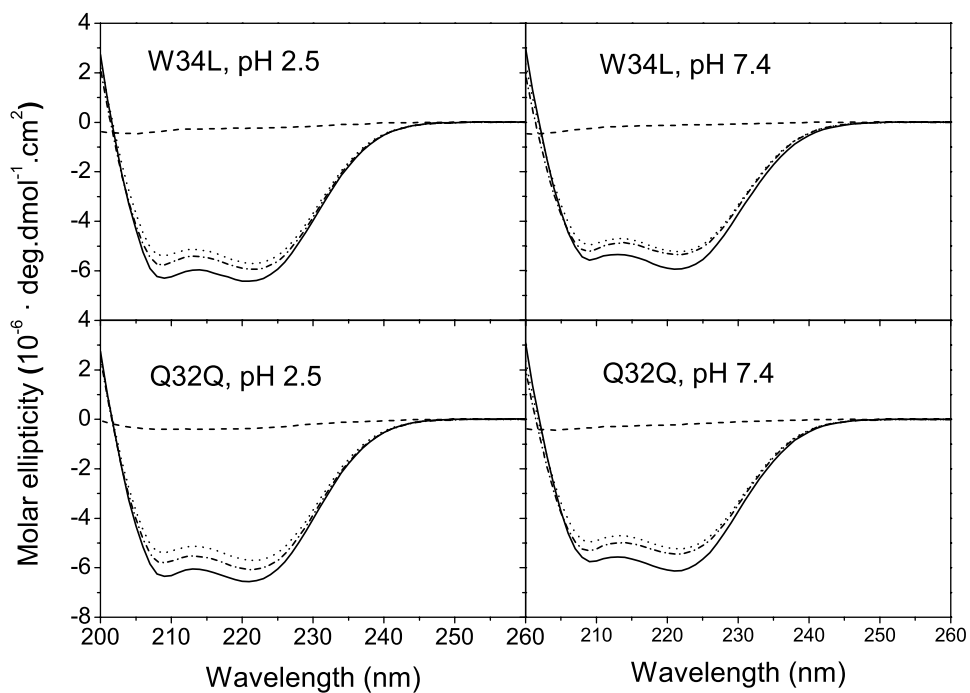
Figure S7: Interaction between covNHR3 with each of the peptides Q32Q and W34L monitored by far-UV CD at pH 2.5 and 7.4 and 25°C. Equimolar mixtures between each covNHR construct and each peptide were prepared at a concentration of about 15 μM. CD spectra were compared with those of the pure samples. Spectra are normalized per mole and represented as follow: 1:1 mixtures (solid lines); free covNHR construct (dotted lines); free peptide (dashed lines); sum of the spectra of the free molecules (dashed-dotted lines). The mixtures showed significant increases in α-helical structure relative to the isolated moieties, according to the negative ellipticity values in the far-UV CD spectra.
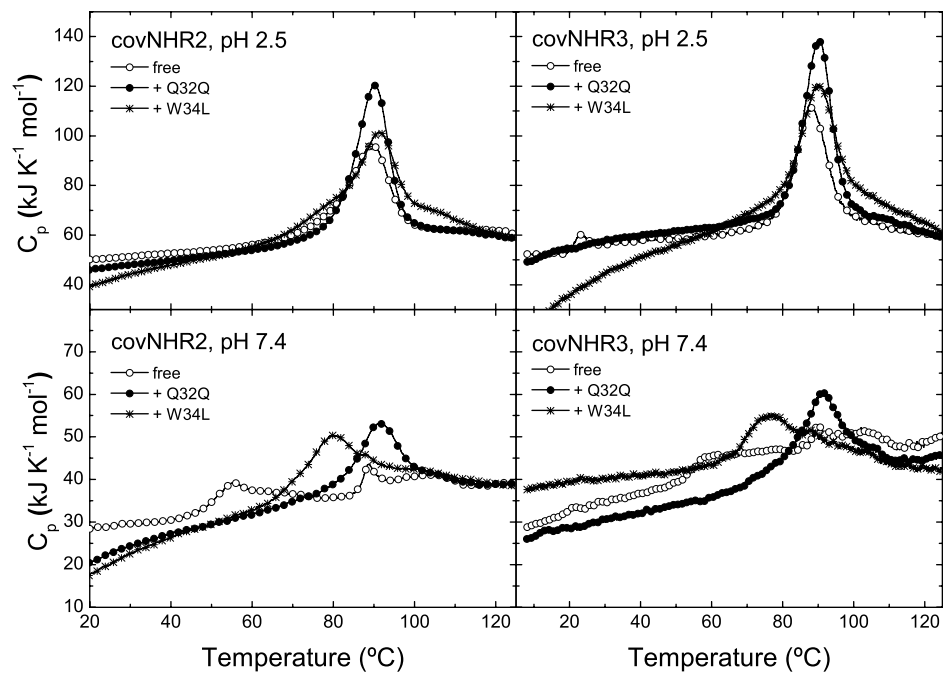
Figure S8: DSC analysis of equimolar mixtures of covNHR constructs and CHR peptides at pH 2.5 and 7.4. Experiments were carried out at a concentration of 20 μM for each mixture. The thermograms of the free covNHR molecules are also shown for comparison.
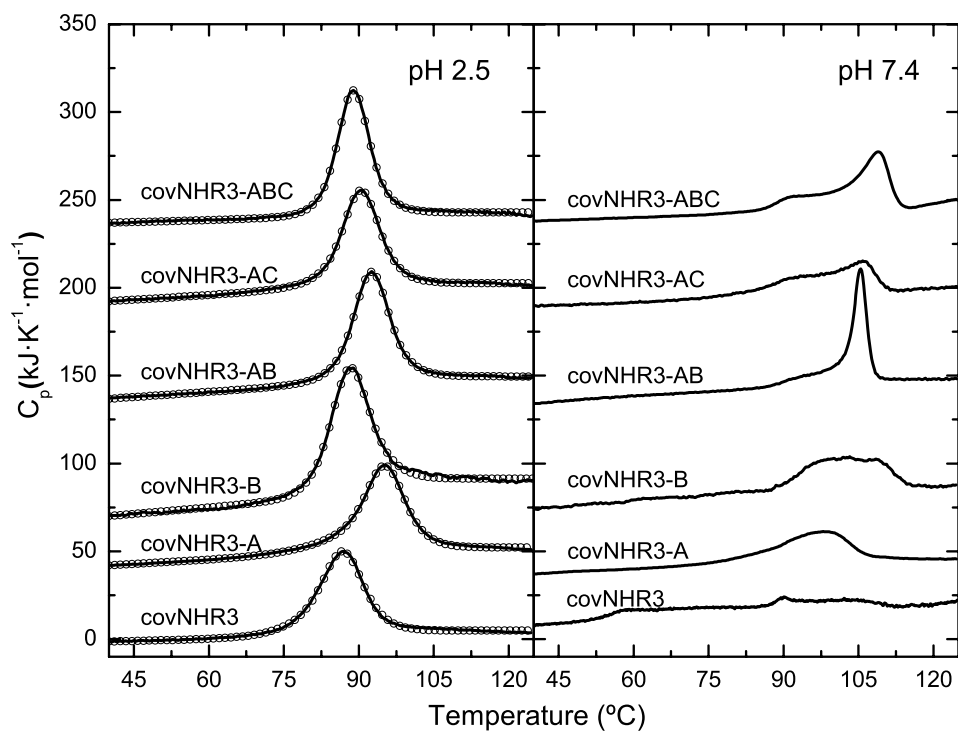
Figure S9: Thermal unfolding of the 2nd generation covNHR proteins analyzed by DSC at different pH values in comparison with covNHR3. Experimental DSC thermograms are represented in solid lines for each construct. The open symbols shown at pH 2.5 represent the best fits of the curves using an unfolding model of two-sequential transitions. The C_p curves have been artificially displaced along the vertical axis for clarity.
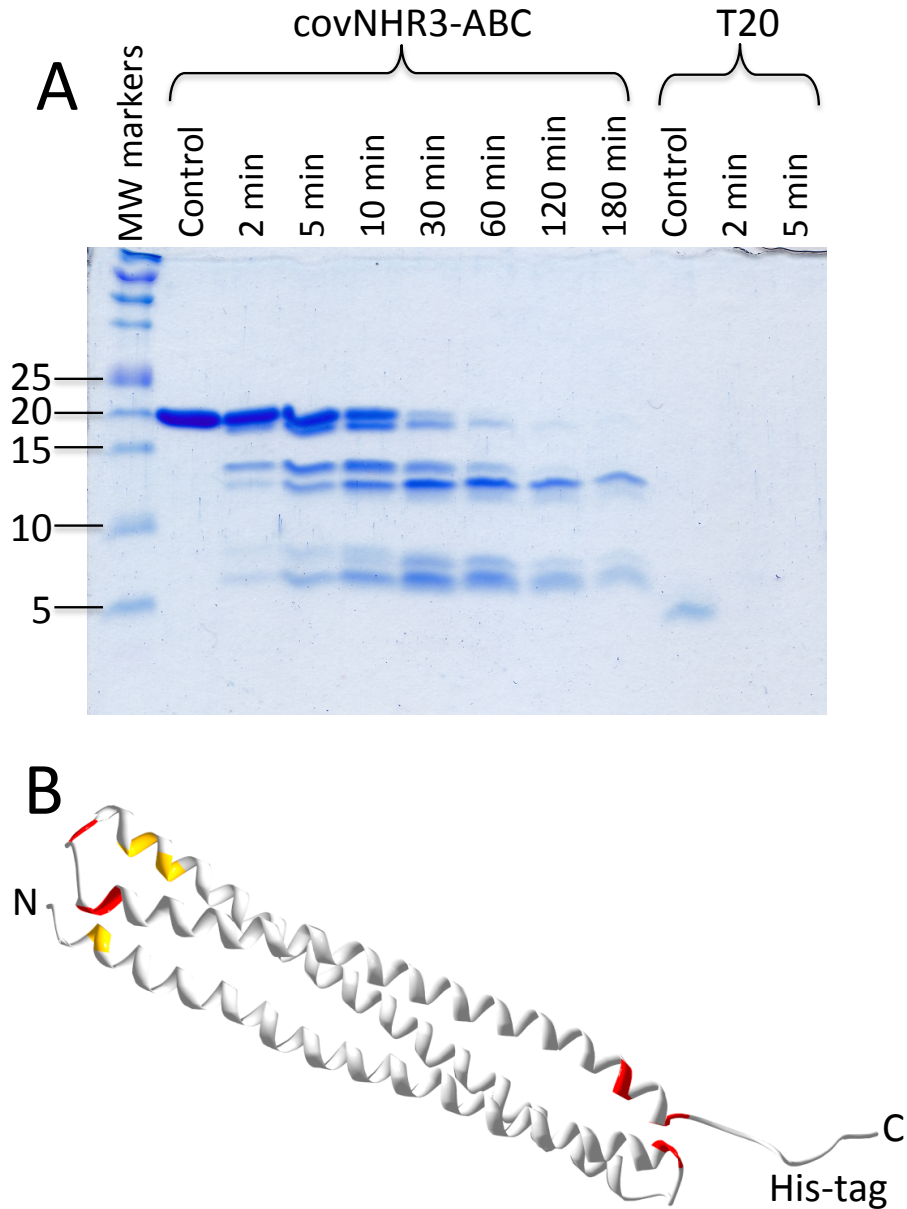
Figure S10: Resistance against proteolysis of covNHR3-ABC and T20. (A) SDS-PAGE analysis of 50 μM covNHR3-ABC or 100 μM T20 samples submitted to proteolytic digestion with 0.1 μM thermolysin at 37ºC, in 50 mM Hepes pH 7.2, 0.1 M NaCl, 5 mM CaCl$_2$ during the times indicated. While T20 is rapidly degraded during the first 5 min, covNHR3-ABC is much more resistant, showing few sites accessible to proteolysis (B) Main cleavage sites for proteolytic attack by thermolysin indicated on a model of covNHR3-ABC. Fragments were identified by mass-spectrometry analysis, taking into account sequence specificity of the protease. Main cleavage sites are indicated in red and secondary cleavage sites in orange. The C-terminal His-tag becomes rapidly cleaved.

Cleavage also occurs at the chain termini and the flexible loops. Fragments encompassing one (A2-Q52 and L117-Q163) and two helices (L6-G110) remain for more than 3 h, suggesting that mutual association protects them from proteolytic attack.
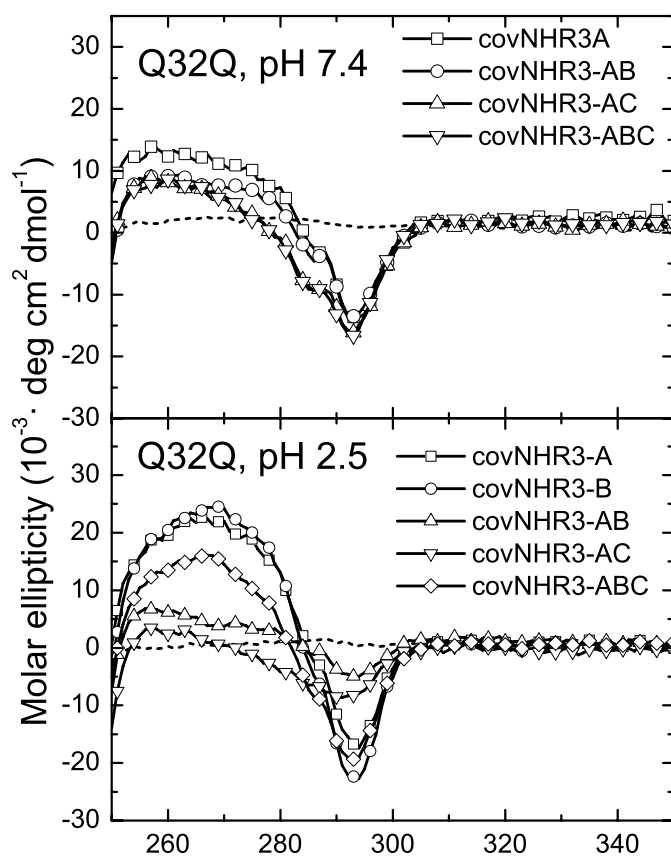
Figure S11: Near-UV CD spectra of equimolar mixtures of the second-generation covNHR constructs and the Q32Q peptide at pH 2.5 and 7.4. Experiments were carried out at a concentration of about 20 μM for each mixture. The covNHR3-B could not be tested at pH 7.4 due to precipitation of the mixture. The spectra of the free Q32Q peptide are also shown in dashed lines for comparison.
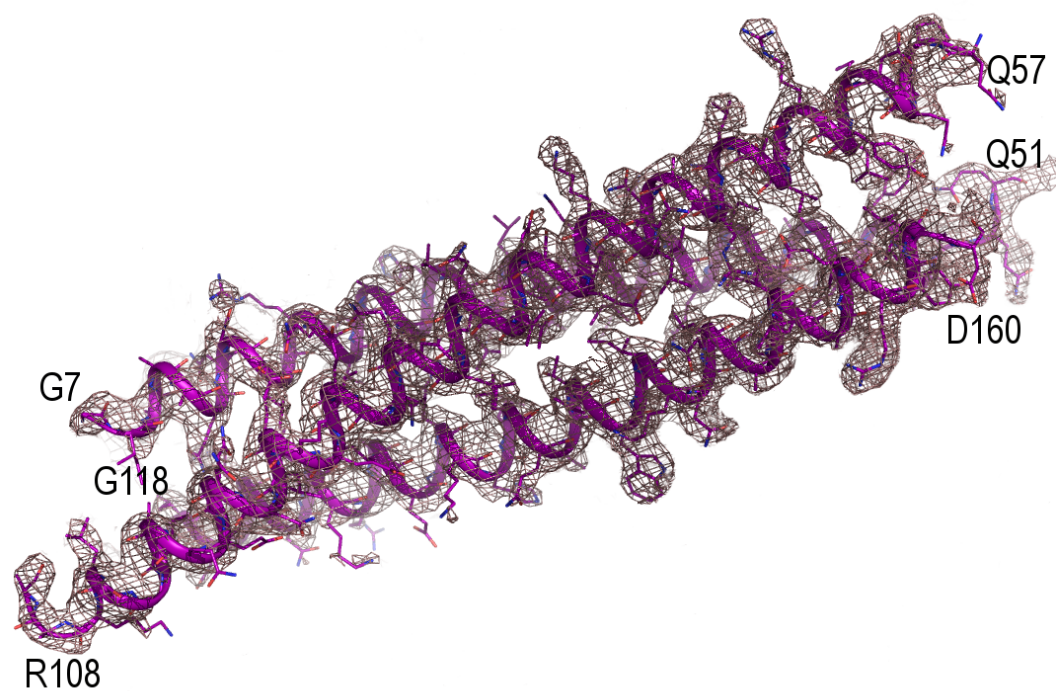
Figure S12: Overall fold of the trimeric coiled-coil superimposed on the corresponding 2Fo−Fc map contoured to 1 σ. The parallel helices (residues Gly7-Gln51 and Arg108-Asp160) and the designed antiparallel helix (residues Glu57-Gly118) were modeled in the electron density maps. The loops connecting these helices are not visible in the difference electron density maps.

Supplementary references

1. Luo P & Baldwin RL (1997) Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 36(27):8413-8421.

2. Peisajovich SG, Blank L, Epand RF, Epand RM, & Shai Y (2003) On the interaction between gp41 and membranes: the immunodominant loop stabilizes gp41 helical hairpin conformation. *J Mol Biol* 326(5):1489-1501.

3. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.

4. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133-144.

5. Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):282-292.

6. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760-763.

7. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.

8. Bunkoczi G*, et al.* (2013) Phaser.MRage: automated molecular replacement. *Acta Crystallogr D Biol Crystallogr* 69(Pt 11):2276-2286.

9. Sabin C*, et al.* (2010) Crystal structure and size-dependent neutralization properties of HK20, a human monoclonal antibody binding to the highly conserved heptad repeat 1 of gp41. *PLoS Pathog* 6(11):e1001195.

10. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.

11. Afonine PV*, et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352-367.

12. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122-133.

13. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12-21.

14. Urzhumtseva L, Afonine PV, Adams PD, & Urzhumtsev A (2009) Crystallographic model quality at a glance. *Acta Crystallogr D Biol Crystallogr* 65(Pt 3):297-300.

15. Fenyo EM*, et al.* (2009) International network for comparison of HIV neutralization assays: the NeutNet report. *PLoS One* 4(2):e4505.

16. Holl V*, et al.* (2004) Involvement of Fc gamma RI (CD64) in the mechanism of HIV-1 inhibition by polyclonal IgG purified from infected patients in cultured monocyte-derived macrophages. *J Immunol* 173(10):6274-6283.

17. Garcia De La Torre J, Huertas ML, & Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* 78(2):719-730.