SUPPORTING MATERIAL

Identification of allosteric disulfides from prestress analysis

Beifei Zhou, Ilona B. Baldus, Wenjin Li, Scott A. Edwards, and Frauke Gräter

Tables

Designation	χ_1	χ_2	χ_3	χ_2'	χ_1'	Designation	χ_1	χ_2	χ_3	χ_2'	χ_1'
-LHSpiral	-	-	-	-	-	-RHHook	-	+	+	-	-
+/-LHSpiral	+	-	-	-	-	-LHHook	-	-	-	+	-
-/+RHHook	+	+	+	-	-	-RHStaple	-	-	+	-	-
+/-RHSpiral	+	+	+	+	-	-RHSpiral	-	+	+	+	-
-/+LHHook	+	+	-	-	-	+/-LHHook	+	-	-	+	-
+RHSpiral	+	+	+	+	+	+/-RHHook	+	-	+	+	-
+/-LHStaple	+	+	-	+	-	-LHStaple	-	+	-	+	-
+LHHook	+	-	-	+	+	+LHSpiral	+	-	-	-	+
+/-RHStaple	+	-	+	-	-	+RHHook	+	+	+	-	+
+RHStaple	+	-	+	-	+	+LHStaple	+	+	-	+	+

Table S1: Classification of disulfide bonds based on the signs of five dihedral angles χ_1, χ_2, χ_3, χ_2', χ_1'. There are twenty classes in total.

	Initial set	without metal ions and ligands	resolution cutoff (2.5 Å) and residue cutoff (300)	MD simulation
number of structures	18,723	5193	1221	667
number of disulfide bonds	178,920			2360

Table S2: Parameters and sample sizes of filtering structures for MD simulations.

Proteins	Disulfide Bonds	χ_1	χ_2	χ_3	χ_2'	χ_1'	Cα -Cα' (nm)	DSE (kJ $\cdot$ mol^{-1})	Designation
CD4	Cys16-84	34.91 ± 28.71	-71.59 ± 6.04	-78.55 ± 0.74	-153.3 ± 0.63	-35.4 ± 15.43	0.66 ± 0.00	18.84 ± 13.82	+/-LHSpiral
	Cys130-159	-67.22 ± 1.48	-90.04 ± 2.65	106.05 ± 0.44	-86.32 ± 2.41	-69.15 ± 1.27	0.40 ± 0.00	16.35 ± 0.93	-RHStaple
vWFC1	Cys3-29	62.26 ± 1.09	120.97 ± 2.74	80.55 ± 0.63	61.7 ± 1.18	-95.1 ± 13.25	0.59 ± 0.00	22.37 ± 5.60	+/-RHSpiral
	Cys27-37	-66.48 ± 3.4	-84.28 ± 5.07	105.32 ± 1.39	-92.67 ± 4.05	-61.92 ± 2.47	0.41 ± 0.00	15.03 ± 1.67	-RHStaple
	Cys24-65	53.66 ± 2.93	-72.83 ± 39.77	82.07 ± 3.03	-125.41 ± 32.68	55.85 ± 2.86	0.65 ± 0.01	11.95 ± 6.56	+RHStaple
	Cys42-66	53.02 ± 2.44	-118.62 ± 44.22	-86.45 ± 2.37	97.92 ± 44.56	48.92 ± 2.37	0.60 ± 0.02	19.55 ± 8.99	+LHHook
	Cys50-71	56.51 ± 15.67	-93.57 ± 17.87	-87.07 ± 5.6	-48.8 ± 69.63	51.84 ± 3.26	0.63 ± 0.02	9.5 ± 9.63	+LHSpiral

Table S3: Classification of disulfides based on dihedrals and dihedral strain energy (DSE) in CD4 and vWFC1.

Figures

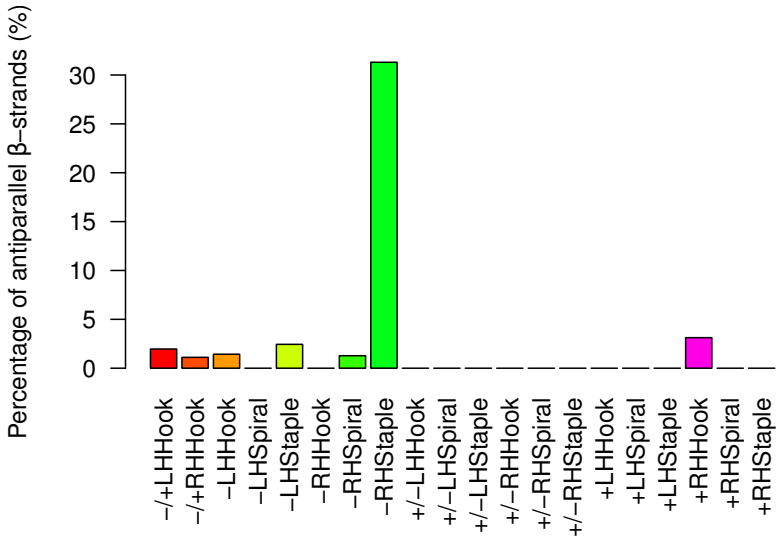

Figure S1: Ratio of the antiparallel β-strands for each type of disulfide bonds. 667 proteins are analyzed, and for each protein, the last 5 ns including 500 frames were taken into account. The secondary structure was defined by the DSSP algorithmm.

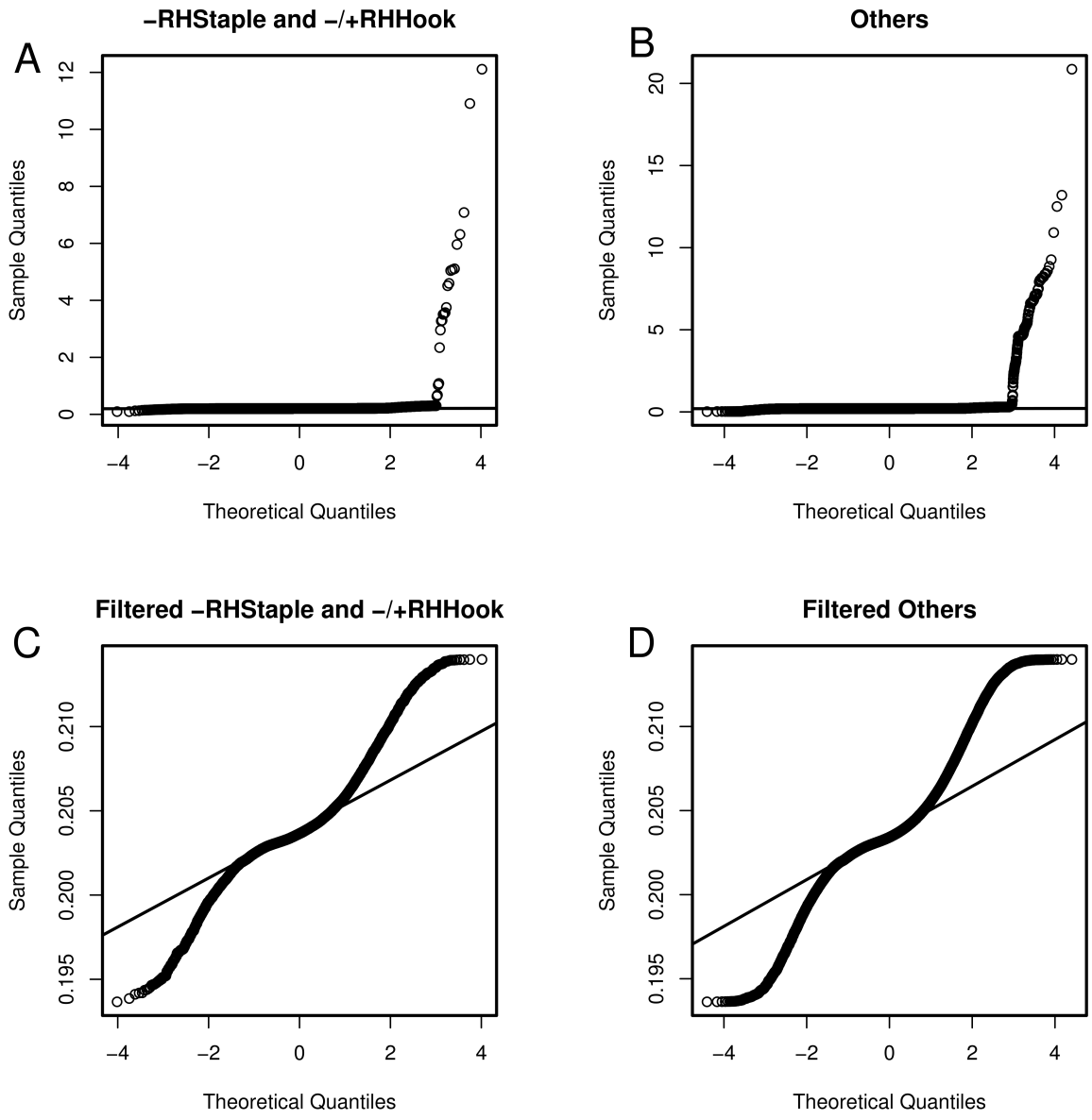

Figure S2: Quantile-quantile plot. (A) qq plot for bond lengths of the -RHStaple and the -/+RHHook disulfides. (B) qq plot for bond lengths of other common disulfides. (C) qq plot with the filtered data of bond lengths for the -RHStaple and the -/+RHHook disulfides. (D) qq plot with the filtered data of bond lengths for other common disulfides. Filtering is described in the Methods section.

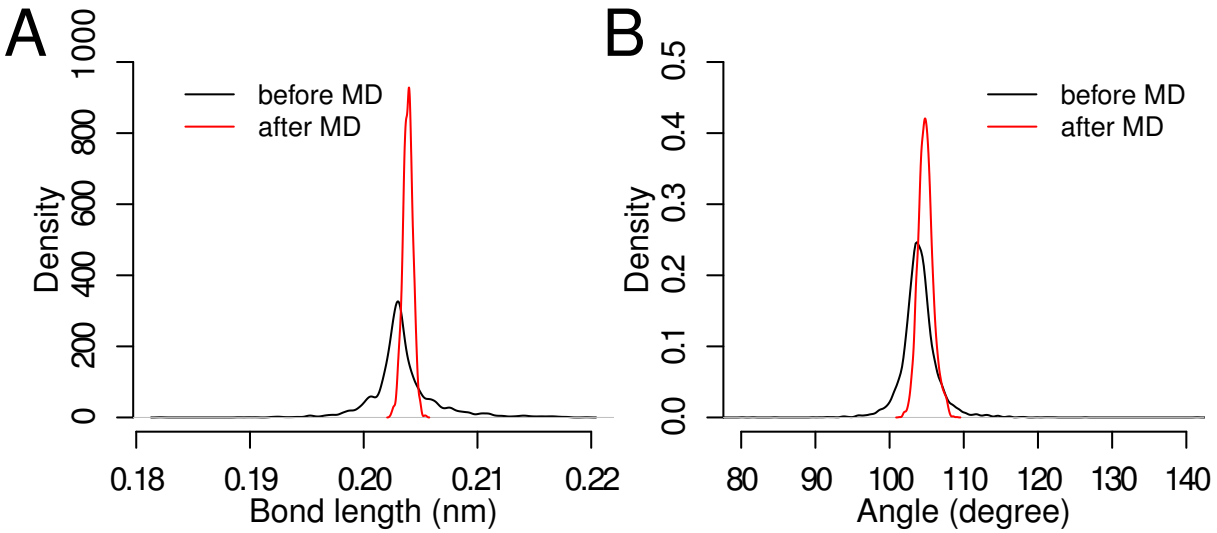

Figure S3: Comparison of (*A*) the distributions of S-S bond lengths and (*B*) the distributions of C-S-S angles before (*black*) after (*red*) MD simulations.

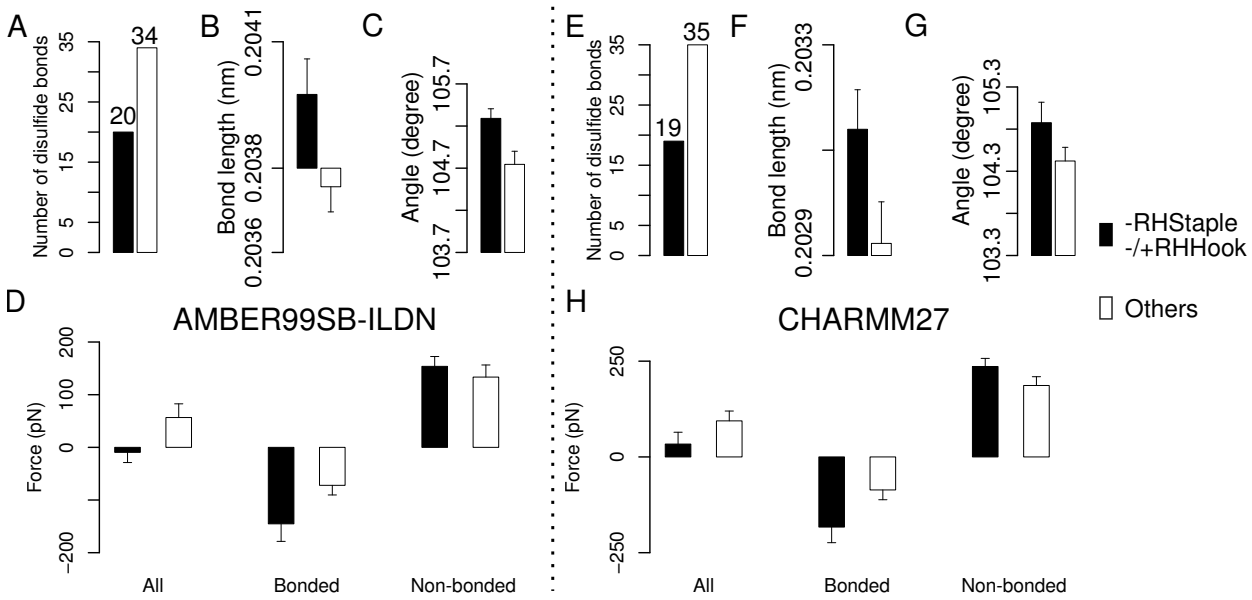

Figure S4: Comparison of the prestress calculated from the AMBENER99SB-ILDN and CHARMM27 force fields. Error bars represent SEM. There are 13 proteins (pdb codes: 1KPT, 1MWP, 2BLG, 2EA3, 1CDH, 1BOQ, 1H34, 1JIS, 2MEA, 1CKH, 1F0W, 152L, 3TGL). (*A*)-(*D*) AMBER99SB-ILDN. (*E*)-(*H*) CHARMM27. (*A*) Number of disulfides in the -RHStaple and the -/+RHHook configuration (black) and the other three types of disulfides in the test set (white). (*B*) S-S bond length. (*C*) C-S-S angles. (*D*) Mean prestress between two adjacent cysteines including all interactions (*left block*), only including bonded interactions (*middle block*), and only including non-bonded interactions (*right block*).

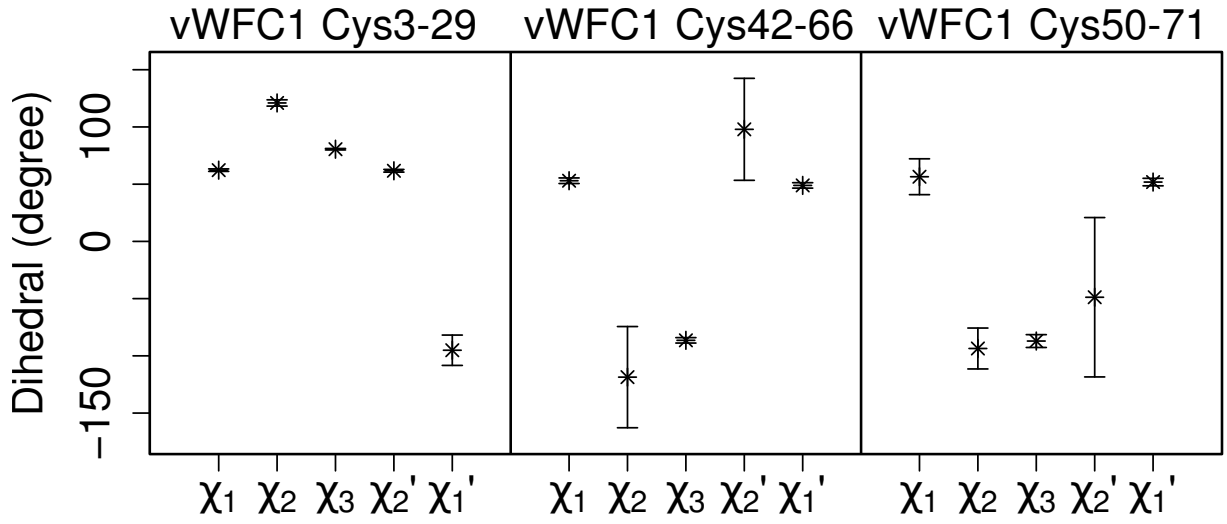

Figure S5: Dihedral angles for disulfide bonds in the vWFC1 domain. Dihedral angles for the other two disulfide bonds of vWFC1 are given in the main text (Fig. 3 *B*).

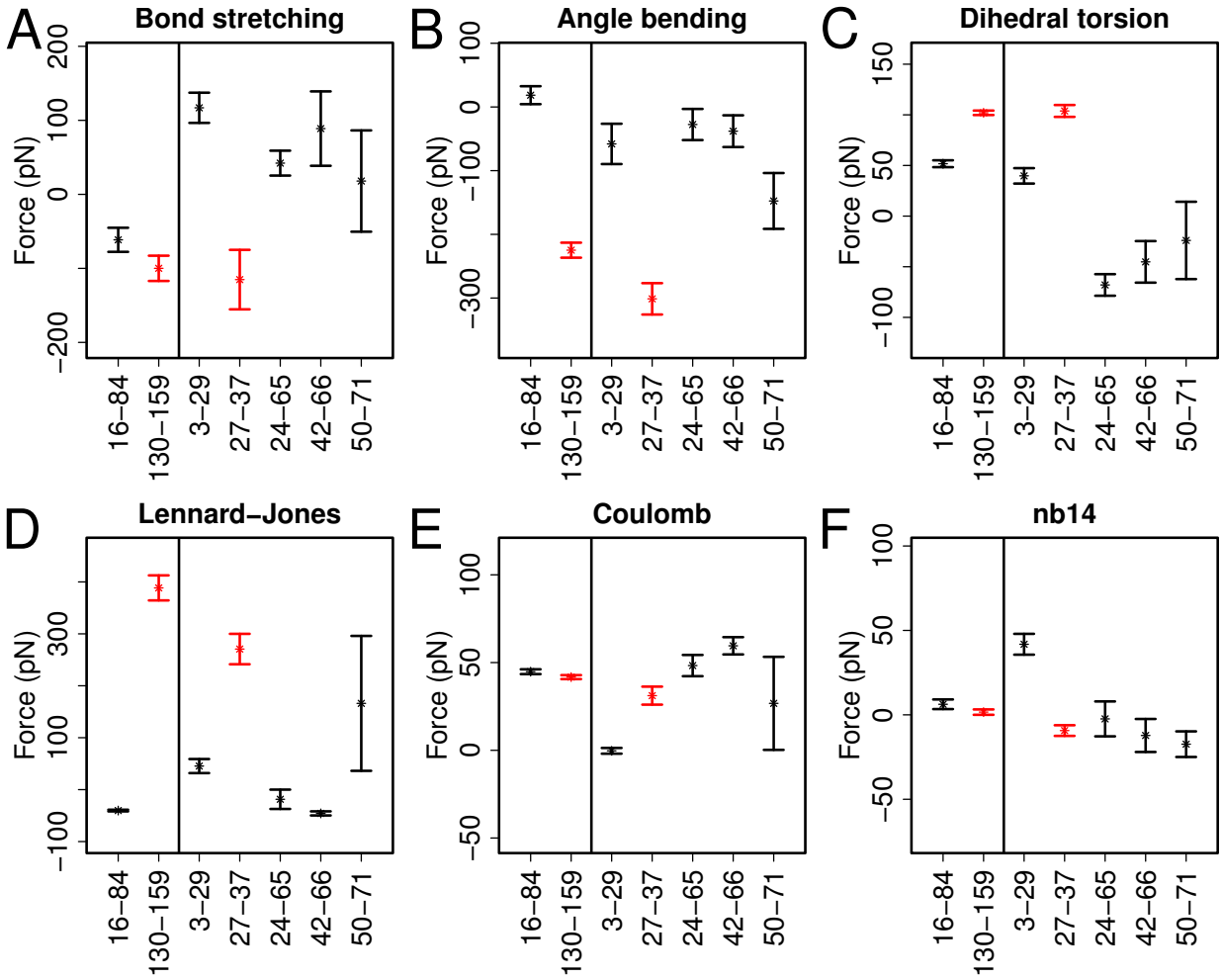

Figure S6: Prestress of disulfides (from Fig. 4) separated into contributions from different interactions. (*A*) Bond stretching interactions. (*B*) Angle bending interactions. (*C*) Dihedral torsion interactions. (*D*) Lennard-Jones interactions. (*E*) Coulomb interactions. (*F*) nb14 interactions. In each subfigure, the x-axis shows disulfide bond pairs, and with two bars on the left for CD4, and the five on the right for vWFC1. Averages and standard errors were obtained from ten independent trajectories, each 50 ns in length. Red: -RHStaple, black: other disulfides.

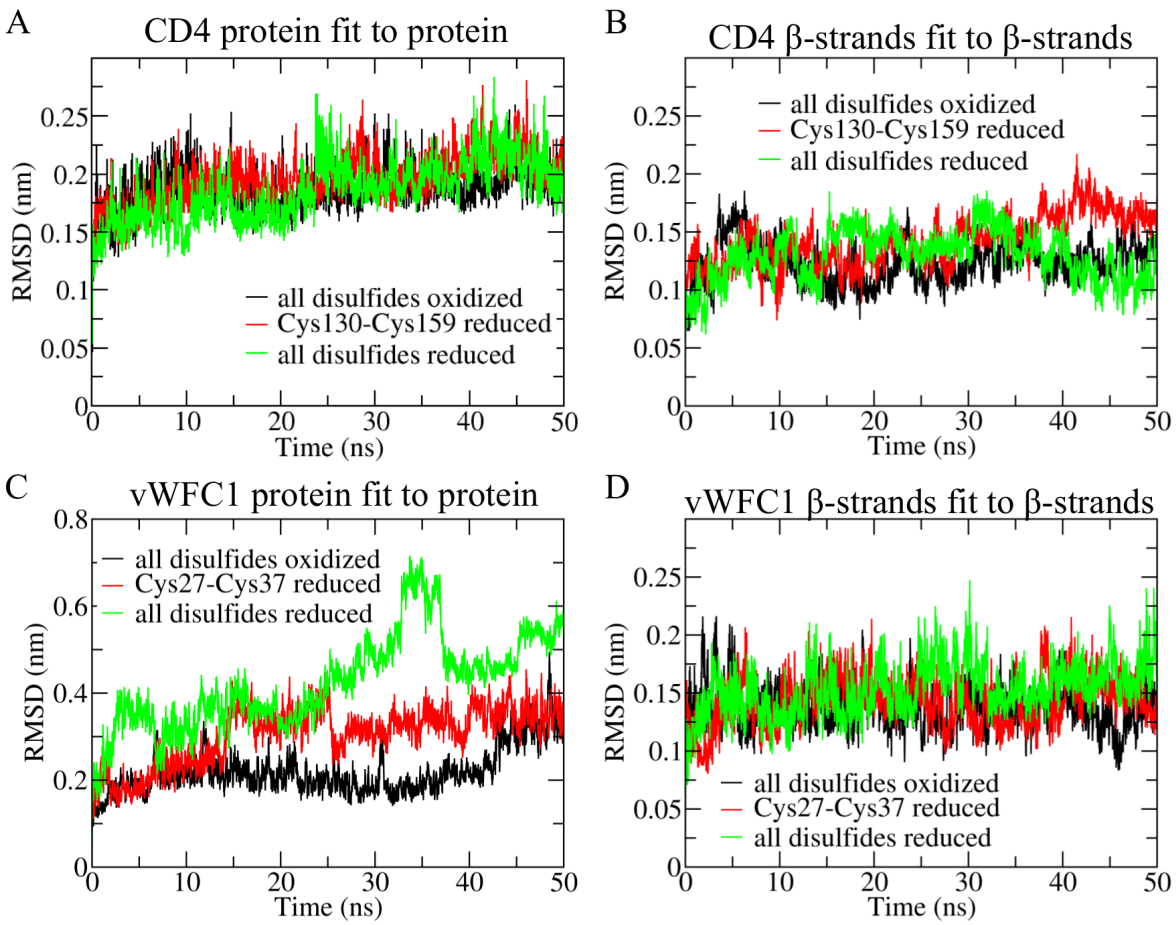

Figure S7: Reduction of disulfide bonds. Structural fits were done to all atoms of the specified group. (*A*) RMSD (root-mean-square deviation) of CD4 in different oxidized and reduced states as indicated. (*B*) RMSD of two adjacent β-strands in CD4. (*C*) RMSD of vWFC1 in different oxidized and reduced states as indicated. (*D*) RMSD of two adjacent β-strands in vWFC1.

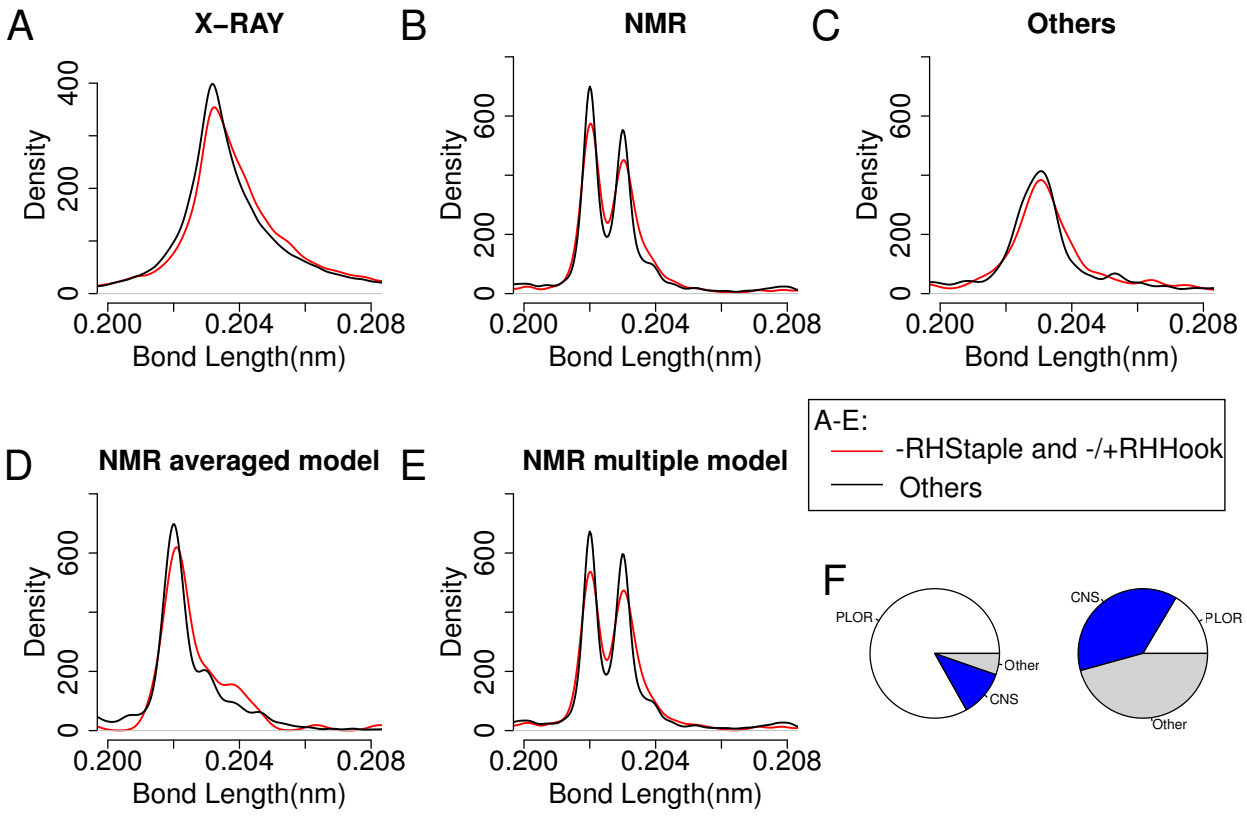

Figure S8: Statistical analysis of disulfide bonds in static protein structures. The red curve represents the -RHStaple and the -/+RHHook disulfides, the black curve represents other disulfides. S-S bond lengths for structures obtained with different experimental techniques, (*A*) X-ray, (*B*) NMR, (*C*) other methods, (*D*) NMR averaged models, and (*E*) NMR multiple models. (*F*) Proportion of softwares used for predicting structures. The left one shows the proportions of disulfides in the first peak of (*E*), the right one shows the proportion in the second peak.

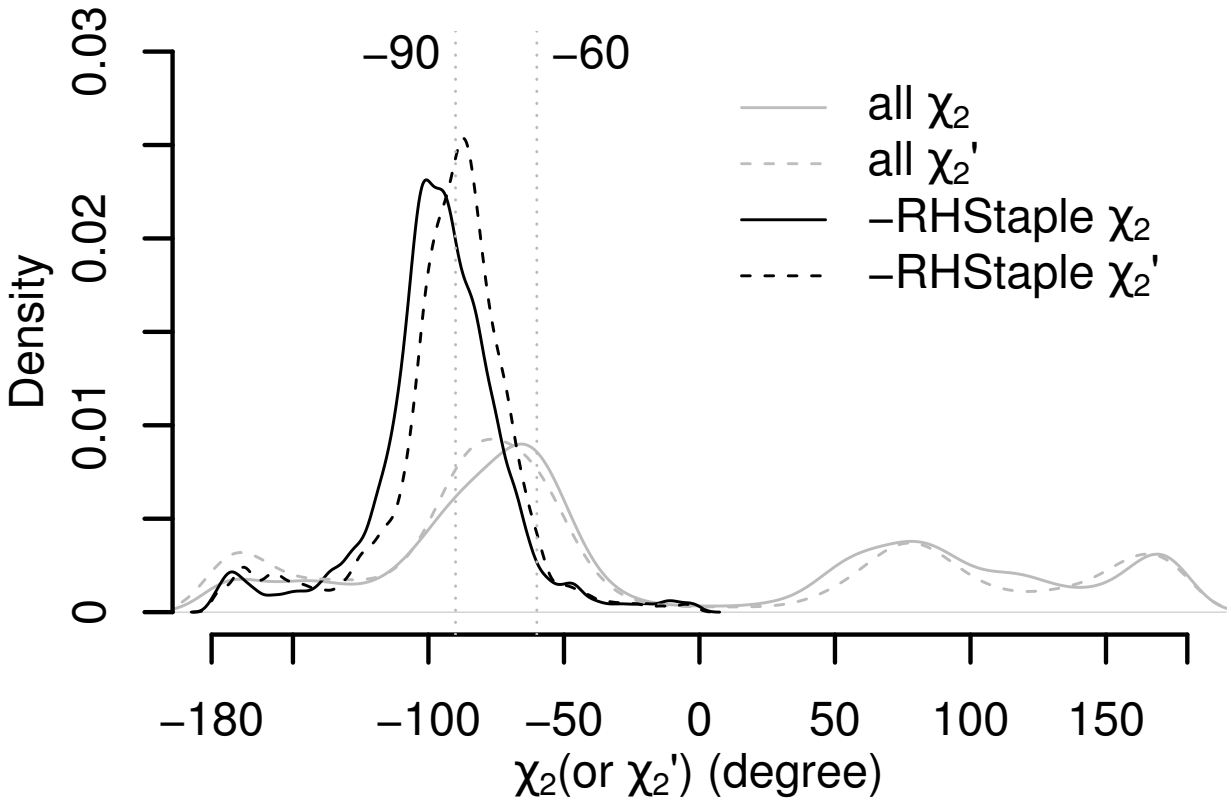

Figure S9: The distribution of χ_2 and χ_2' angles for all disulfide bonds and for the -RHStaple disulfide bonds .