Supporting Information

Müller et al. 10.1073/pnas.1300547110

SI Materials and Methods

Cloning, Mutagenesis, Expression, and Purification of the Constant Fragment Domains. The Cμ2 (Ile225–Ala339), Cμ3 (Asp344–Lys443), Cμ4 (Glu446–Lys558), and Cμ4tp (Glu446–Tyr576) genes were obtained by PCR amplification using the mouse IgM heavy-chain cDNA as template. Amino acid numbering is according to the UniProt entry P01872 (IGHM_MOUSE) with the V_H (122 amino acids) added. The genes were cloned into the pET28a expression vector (Merck/Novagene) via the NcoI and HindIII restriction sites. The C337S, C414S, and C575S mutations were introduced using the QuikChange site-directed mutagenesis strategy (Stratagene/Agilent). Expression was performed in *Escherichia coli* BL21 cells at 37 °C. At an OD_{600} of 0.6–0.8, expression was induced with 1 mM isopropylthio-β-galactoside. Cells were harvested after overnight growth and inclusion bodies were prepared as previously described (1). Inclusion bodies were solubilized in 50 mM Tris·HCl (pH 7.5), 10 mM EDTA, 8 M urea, and 2 mM β-mercaptoethanol at 4 °C for 2 h. Insoluble components were removed by centrifugation (46,000 × *g*, 25 min, 4 °C). The supernatant was applied to a Q-Sepharose column equilibrated with 50 mM Tris (pH 7.5), 10 mM EDTA, and 5 M urea. The protein was present in the flow-through, and refolding was carried out via dialysis into 250 mM Tris·HCl (pH 8.0), 100 mM L-arginine, 10 mM EDTA, 1 mM oxidized glutathione, and 0.5 mM reduced glutathione at 4 °C overnight. To remove misfolded protein and remaining impurities, the protein was applied to a HiLoad Superdex75 26/60-pg gel-filtration column (GE Healthcare) equilibrated in PBS buffer. For NMR (chemical shift perturbation analysis) and SAXS measurements, a Cμ3C414S–Cμ4 construct was cloned, which included the cysteine at position 337 (Cys337–Lys558) and the protein was expressed, refolded, and purified as disulfide linked dimer. Isotope-labeled proteins were expressed in M9 minimal medium containing ^{15}N ammonium chloride as the only nitrogen source and in addition ^{13}C glucose as the only carbon source. Partially deuterated protein was expressed in 70% D_2O-containing medium. All constructs were sequenced and the mass of the purified proteins was verified using matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

Analytical Gel Filtration. The oligomeric state of the domains were determined by analytical size-exclusion chromatography (SEC) measurements using a Shimadzu HPLC System and an analytical Superdex75 10/300 GL column (GE Healthcare) equilibrated with PBS (8.09 mM Na_2HPO_4, 1.76 mM KH_2PO_4, 137 mM NaCl, and 2.65 mM KCl, pH 7.4) at 20 °C. All experiments were performed with a flow rate of 0.5 mL/min. Fifty-microliter samples were injected at concentrations of 1, 10, and 100 μM, and the elution profiles were detected by absorbance at 280 nm or by fluorescence emission at 350 nm. For the Cμ4tp domain, two Superdex200 5/150 GL columns (GE Healthcare) in series and a Dawn Heleos multiangle light scattering detector (Wyatt) were used in addition.

Analytical Ultracentrifugation. Analytical ultracentrifugation was carried out with a ProteomLab XL-I (Beckman-Coulter) equipped with absorbance optics. All experiments were performed using PBS at 20 °C. For the determination of the masses of the individual domains and mutants, sedimentation velocity experiments were performed at a rotor speed of 50,000 rpm. Sedimentation was detected at 230, 280, or 300 nm, depending on protein concentration. For the Cμ4tp oligomer, sedimentation velocity runs were performed with protein concentrations of 1–20 μM at 35,000 rpm. Data analysis was carried out by the C(s) method with TI (time invariant) and RI (radial invariant) noise fitting using SEDFIT version 14.1 software (2).

For K_d measurements, sedimentation equilibrium runs with protein samples of different concentrations were performed. Sedimentation was detected at 230, 280, or 300 nm, depending on protein concentration. Data from various concentrations and rotor speeds were fitted globally to a self-association model (3) in Origin (OriginLab) according to Eq. **S1**:

$$a(r) = c_1(r_0)\varepsilon_1 d \exp\left[M_b \frac{\omega^2}{2RT}\left(r^2 - r_0^2\right)\right] + K_{12} c_1(r_0)^2 2\varepsilon_1 d \exp\left[2M_b \frac{\omega^2}{2RT}\left(r^2 - r_0^2\right)\right]. \quad \textbf{[S1]}$$

The $c_1(r_0)$ value denotes the molar concentration at a reference position r_0; M_b, ε_1, and d designate the buoyant mass, molar extinction coefficient, and the optical path length, respectively. K_{12} is the association constant for the monomer/dimer equilibrium.

Fluorescence and CD Measurements. CD measurements were carried out at 25 °C in a Jasco J-720 spectropolarimeter. Far UV (FUV)-CD spectra were recorded from 195–260 nm at a protein concentration of 10 μM in 1 mm quartz cuvettes. Near UV (NUV)-CD spectra were collected form 260–340 nm at a protein concentration of 40 μM in 2-mm quartz cuvettes. All spectra were accumulated 16 times and buffer-corrected. For temperature-induced equilibrium unfolding, the change in ellipticity at 205 nm was recorded and the rates for heating and cooling were set to 10 °C h^{-1}.

Fluorescence measurements were performed in a Spex FluoroMax-4 spectrofluorimeter (Instruments SA). To determine the maximum difference between native and guanidium chloride (GdmCl)-induced unfolded protein, intrinsic tryptophan fluorescence spectra were obtained from 310–450 nm, with exitation at 280 nm and excitation and emission slit width set to 1 nm and 6 nm, respectively. Measurements were carried out using a protein concentration of 1 μM in 1-cm quartz cuvettes.

GdmCl-induced equilibrium unfolding was monitored by fluorescence and CD spectroscopy. For fluorescence measurements, the excitation wavelength was set to 280 nm and the emission was determined at 360 nm with slit width of 1.5 nm and 6 nm, respectively. For CD measurements, the change in ellipticity at 212 nm (Cμ2), 213 nm (Cμ3), 220 nm (Cμ4), or 221 nm (Cμ4tp) was monitored. All measurements were performed at a protein concentration of 10 μM. For GdmCl-induced unfolding transitions, the samples were incubated over night at 25 °C at different GdmHCl concentrations before measurements.

Data were evaluated using Origin (OriginLab); for GdmCl-induced unfolding transitions, a two-state model was applied (4).

Crystallization and Structure Determination. Crystallization of the Cμ2 and Cμ4 domains of IgM was performed by the sitting drop vapor diffusion method at 20 °C. Ten milligrams per milliliter (800 μM) of protein solution in 10 mM Tris·HCl (pH 7.5) were mixed with an equal volume of buffer C2 [1.0 M lithium chloride, 0.1 M MES (pH 5.4), and 23% PEG 6000] for the Cμ2 domain and buffer C4 [15% PEG 8000 and 0.1M Hepes (pH 7.2)] for the Cμ4 domain. In both cases, crystals grew to a final size of ~50 × 50 × 25 μm^3 within 48 h. Crystals were either soaked in 2 M lithium sulfate at pH 5.4 (Cμ2 domain) or buffer C4 with additional 25% PEG 200 (Cμ4 domain) before being cooled in a

stream of liquid nitrogen at 100 K. Native diffraction datasets were collected using synchrotron radiation at the X06SA-beamline (Swiss Lightsource) (Cμ2 domain at 1.3 Å resolution, Table S1) or on a Bruker Microstar/X8 Proteum (Bruker AXS Inc.) with a Cu rotating anode (λ = 1.54 Å; Cμ4 domain at 1.9 Å resolution, Table S1).

For structure determination of the Cμ2 domain, data were processed using the program package XDS (5). The self-rotation function, calculated with MOLREP (6), revealed noncrystallographic twofold symmetries indicating two Cμ2 domain molecules in the asymmetric unit cell. Determination of the crystal structure was performed by molecular replacement using the program PHASER (7). Coordinates of the corresponding domain of the human IgE (1O0V, residues 228–328) were applied as starting model for the ligand structure. After rigid body and preliminary positional refinement with REFMAC (8) as well as noncrystallographic symmetry averaging using the software package MAIN (9), both Cμ2 domains were traced. The refined model resulted in improved density and allowed finalization of the last missing secondary structures and loop connections. The structures were completed in successive rounds using the interactive 3D graphic programs COOT (10) and anisotropically refined with REFMAC5 (8). Water molecules were located either by using ARP/wARP (11) within the CCP4i GUI (12) or by using COOT (10). All solvent molecules were verified manually in COOT (10). Temperature factors were corrected with restraints between bonded atoms and between noncrystallographic symmetry related atoms using translation/libration/screw (TLS) parameters with current crystallographic values of R_{work} = 0.138 and R_{free} = 0.168 (13). Data collection and refinement statistics are summarized in Table S1. Coordinates were confirmed to have good stereochemistry from the Ramachandran plot using PROCHECK (14), which displays 99.0% of residues in the most favored region and 1.0% of residues in the additionally allowed regions. The asymmetric unit cell contains two Cμ2 domains (with the N-terminal Met and the last two Ala residues at the C terminus being structurally disordered) and 291 water molecules (Table S1). The data of the Cμ2 domains have been deposited in the Protein Data bank (PDB ID code 4JVU).

For structure determination of the Cμ4 domain, data were processed with the Proteum 2 software suite (version 2006; Bruker AXS Inc.). MOLREP revealed twofold symmetries suggesting four Cμ4 domains in the asymmetric unit cell. Initial phases were received by Patterson search calculations applying PHASER and using the coordinates of the previously determined model structure of the Cμ2 domain as starting model. Similarly as described for the structure elucidation of the Cμ2 domain, cycles of model building, fourfold noncrystallographic symmetry averaging for electron density map improvement and TLS refinement yielded the final Cμ4 model. Data collection and refinement statistics are summarized in Table S1. The stereochemical quality of the Cμ4 domains were assessed with PROCHECK, displaying in total 97.9% of residues in the most favored region and 2.1% of residues in the additionally allowed regions. The asymmetric unit cell contains four Cμ4 structures (with the N-terminal five amino acids being structurally disordered) and 428 water molecules (Table S1). The data of the Cμ4 domains have been deposited in the Protein Data bank (PDB ID code 4JVW). Molecular illustrations were prepared using the PyMOL Molecular Graphics System (DeLano Scientific).

NMR Spectroscopy and Structure Calculation. NMR data were acquired at 298 K on a Bruker Avance III 600-MHz spectrometer equipped with a TCI CryoProbe head. Processing was done with NMRPipe (15) and analysis was carried out using Sparky (16). The backbone resonances were assigned manually based on HNCACB, HN(CO)CACB, HNCO, and HN(CA)CO experiments (17). Side chains were assigned using an (H)CCH-TOCSY experiment. Distance restraints were extracted from ^{15}N and ^{13}C edited NOESY spectra with mixing times of 70 ms (17, 18). Finally, ^{15}N R_1 and R_2 relaxation rates, as well as $\{^1H\}$-^{15}N heteronuclear NOE values, were measured at 600 MHz proton Larmor frequency at 298 K (19). All data were acquired with a 300-μM uniformly ^{15}N,^{13}C-labeled Cμ3C414S sample, except for the stereospecific assignment of valine and leucine side chains, which was based on a ^1H-^{13}C heteronuclear multiple quantum correlation acquired on a 10% ^{13}C-labeled sample as previously described (20, 21).

NOE cross-peaks were assigned automatically by the software CYANA 3.0 (22) and then inspected manually for assignment and completeness. The distance restraints extracted from the CYANA calculation as well as torsion angles derived by TALOS+ (23) were used in a water refinement calculation in Aria 2.2 (24). Finally, the quality of the structure ensemble was validated using WHATIF (25) and PROCHECK (26). The data of the Cμ3 domain have been deposited in the Protein Data bank (ID code 4BA8).

To map the Cμ3/Cμ4 interface, chemical shift perturbation analysis was performed, following chemical shifts in the HNCA spectra of the ^{15}N-labeled dimeric Cμ3C414S–Cμ4 construct compared with the labeled Cμ3C414S and plotted for each residue in Cμ3C414S. The chemical shift perturbations have been calculated according to Eq. **S2**:

$$CSP(\text{ppm}) = \sqrt{\left(\frac{\delta 15N}{10}\right)^2 + \delta 1H^2}. \qquad \textbf{[S2]}$$

The chemical shift perturbations of the different nuclei have been normalized using their gyromagnetic ratio.

Small Angle X-Ray Scattering Measurements. Small angle X-ray scattering (SAXS) data for solutions of Cμ4tp were recorded on an in-house SAXS instrument (SAXSess mc^2; Anton Paar GmbH) equipped with a Kratky camera, a sealed X-ray tube source, and a one-dimensional reverse-biased silicon diode array detector (Mythen 1K; Dectris). The scattering patterns were measured with a 30-min exposure time (three frames, each 10 min) for several solute concentrations in the range from 188 to 751 μM. Radiation damage was excluded based on a comparison of individual frames of the 30 min exposures, where no changes were detected. A range of momentum transfer of $0.01 < s < 0.7$ $Å^{-1}$ was covered ($s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 1.5$ Å is the X-ray wavelength).

SAXS data for solutions of Cμ4tp and Cμ3C414S–Cμ4 in the presence of 5 mM DTT were recorded at the X33 beamline of the European Molecular Biology Laboratory at Deutsches Elektronen Synchrotron using a MAR345 image plate detector. The scattering patterns were measured with a 2-min exposure time (eight frames, each 15 s) for several solute concentrations in the range from 110 to 578 μM. Radiation damage was excluded based on a comparison of individual frames of the 2-min exposures, where no changes were detected. Using the sample–detector distance of 2.7 m, a range of momentum transfer of $0.01 < s < 0.6$ Å-1 was covered ($s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 1.5$ Å is the X-ray wavelength).

All SAXS data were analyzed with the package ATSAS. The data were processed with the SAXSQuant software (version 3.9) and desmeared using the program GNOM (27). The forward scattering, I(0), the radius of gyration, R_g, the maximum dimension, D_{max}, and the interatomic distance distribution functions, (P(R)), were computed with the program GNOM (27). The masses of the solutes were evaluated by comparison of the forward scattering intensity with that of a human serum albumin reference solution (molecular mass 69 kDa). The structure of the Cμ4tp hexamer was modeled using the program CORAL (28). The crystal structure of Cμ4 was taken as input, a C6 symmetry

was defined, and the flexible N- and C-terminal regions (residues E446–H448 and Gly557–Tyr576) were randomized. No connections between the Cμ4 dimers have been defined owing to lack of experimental data defining the interface. A total of 50 structures were calculated, and the best scoring structure (based on the solvation free energy gain upon formation of the interface, $\Delta^{i}G$) was selected to model the flexible C terminus. To test whether the additional Cμ4 interface observed in the crystal structures can be used to construct the structure of the hexamer we calculated an ensemble of structures using the crystal structure of the Cμ4 dimer of dimers as input. The calculations were carried out using the program CORAL (28) with the same settings as described above, but defining C3 symmetry and distance restraints between the dimers of dimers according to the additional interface observed in the crystal structure. Because the Cμ4 dimers of dimers might adopt orientations different from those in the hexamer, we used all possible combinations of distance restraints (i.e., contacts of chain A with chain B/C of the next Cμ4tp dimer of dimers, chain A–chain C/B, chain D–chain B/C, and chain D–chain C/B). The resulting structures are in worse agreement with the experimental data compared with the calculation in which none of the Cμ4 interfaces in the hexamer was restrained (Fig. S6). C-alpha coordinates of the Cμ4tp hexamer of dimers SAXS model have been deposited in the Protein Data Bank (ID code 4BLE).

To model the structure of the entire IgM, the Cμ4tp hexamer modeled based on SAXS data was taken as input for its extension with Cμ2, Cμ3, and the flexible linkers/termini (Cμ2–Cμ3, Cμ3–Cμ4, and the tail piece extending Cμ4). The following input and data were used in the rigid body docking program CORAL: the atomic resolution structures of (*i*) Cμ2, (*ii*) Cμ3, (*iii*) the SAXS model of the Cμ4 hexameric ring, (*iv*) the disulfide linkages (C414 and C575), and (*v*) the Cμ3–Cμ4 interface derived from NMR data as input. The connections of Cμ4 to the Cμ3 domains has been made based on the NMR chemical shift perturbations observed for the free Cμ3C414S domain compared with the Cμ3C414S–Cμ4 dimer (Fig. S4). Based on that, ambiguous distance restraints defining that any proton in Cμ3 residues F354 and F358 has to be within a distance closer than 6 Å to any proton of Cμ4 have been defined. A total of 50 structures were calculated, and the structure with the best fit to the experimental data was selected to prepare Fig. 4*C*.

1. Mayer M, Buchner J (2004) Refolding of inclusion body proteins. *Methods Mol Med* 94:239–254.
2. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78(3):1606–1619.
3. Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Sci* 11(9):2067–2079.
4. Bolen DW, Santoro MM (1988) Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of delta G degrees N-U values in a thermodynamic cycle. *Biochemistry* 27(21):8069–8074.
5. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* 26:795–800.
6. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):22–25.
7. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4): 658–674.
8. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240–255.
9. Turk D (1992) Improvement of a programme for molecular graphics and manipulation of electron densities and its application for protein structure determination. PhD thesis (Technische Universität München).
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. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6(5):458–463.
12. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
13. Brünger AT (1992) Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355(6359):472–475.
14. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291.
15. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293.
16. Goddard TD, Kneller D (2007) SPARKY3 (Univ of California, San Francisco).
17. Sattler M, Schleucher J, Griesinger C (1999) *Prog Nucl Magn Reson Spectrosc* 34: 93–158.
18. Otting G, Wüthrich K (1990) Heteronuclear filters in two-dimensional [1H,1H]-NMR spectroscopy: Combined use with isotope labelling for studies of macromolecular conformation and intermolecular interactions. *Q Rev Biophys* 23(1):39–96.
19. Farrow NA, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33(19):5984–6003.
20. Szyperski T, Neri D, Leiting B, Otting G, Wüthrich K (1992) Support of 1H NMR assignments in proteins by biosynthetically directed fractional 13C-labeling. *J Biomol NMR* 2(4):323–334.
21. Coggins BE, et al. (2003) Structure of the LpxC deacetylase with a bound substrate-analog inhibitor. *Nat Struct Biol* 10(8):645–651.
22. Güntert P (2009) Automated structure determination from NMR spectra. *Eur Biophys J* 38(2):129–143.
23. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
24. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M (2003) Refinement of protein structures in explicit solvent. *Proteins* 50(3):496–506.
25. Vriend G (1990) WHAT IF: A molecular modeling and drug design program. *J Mol Graph.* 8(1):52–56, 29.
26. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8(4):477–486.
27. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* 25:495–503.
28. Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* 89(2): 1237–1250.

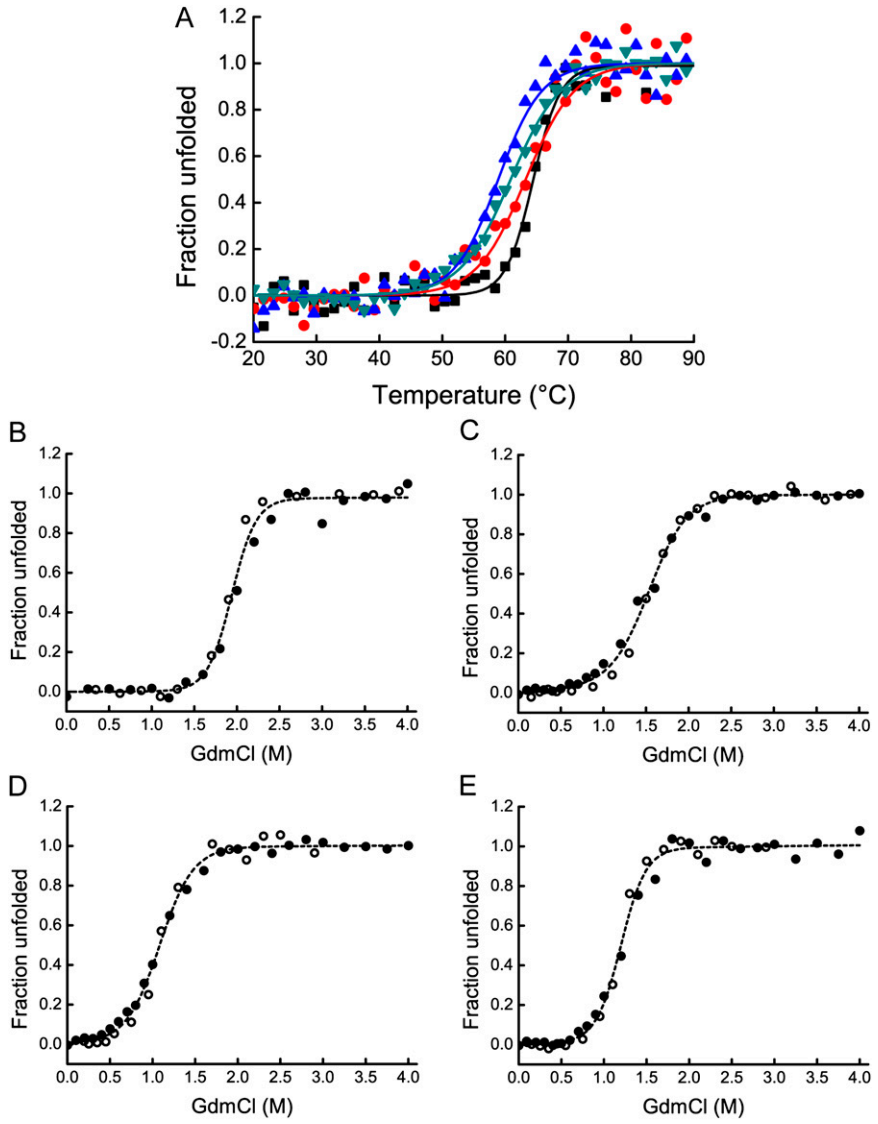

Fig. S1. Stability against thermal and chemical unfolding. (*A*) Temperature denaturation of the Cμ2 (black), Cμ3C414S (red), Cμ4 (blue), and Cμ4tpC575S (green) domains, followed by the far-UV CD signal at 205 nm at a heating rate of 20 °C/h. Data were fitted by a Boltzmann function. GdmCl-induced equilibrium unfolding (filled circles) and refolding (open circles) of the (*B*) Cμ2, (*C*) Cμ3C414S, (*D*) Cμ4, and (*E*) Cμ4tpC575S domains. GdmCl-induced unfolding/refolding transitions were monitored by intrinsic tryptophan fluorescence. Data were fitted by a two-state unfolding model.

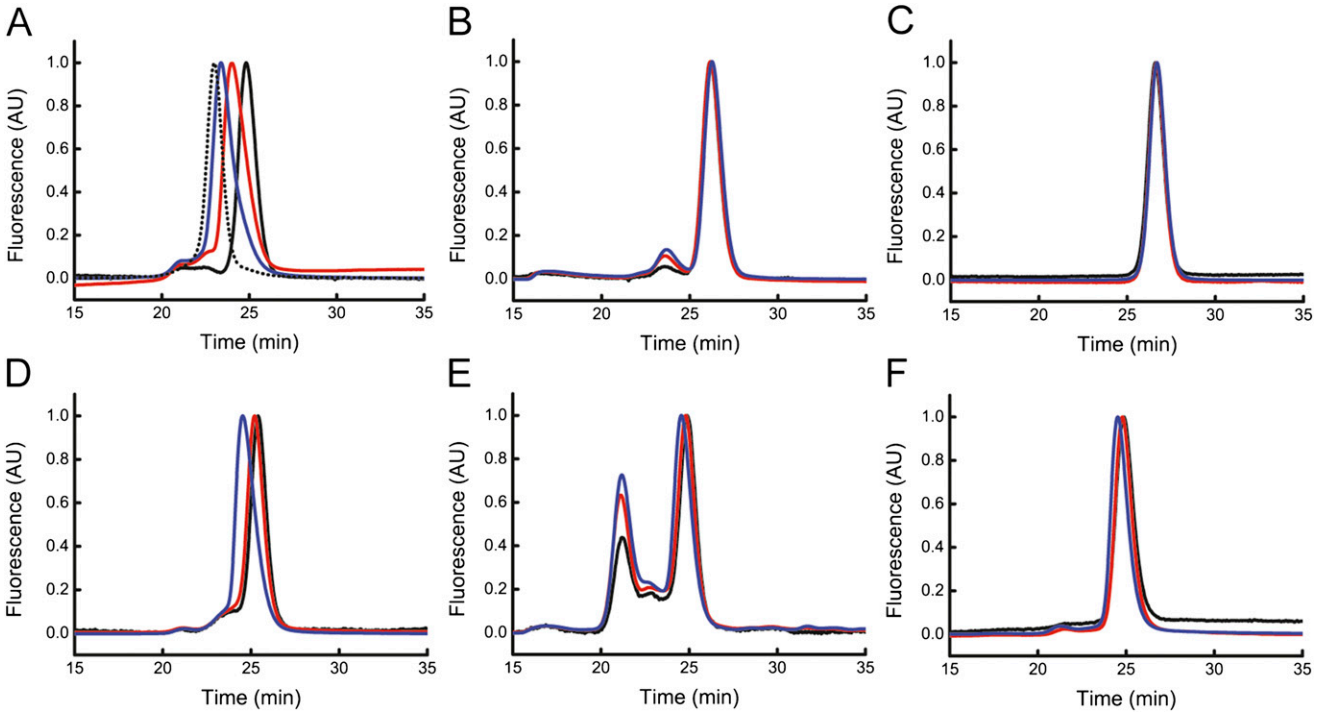

Fig. S2. Characterization of domain oligomerization. SEC profiles of (*A*) Cμ2 (dotted line) and Cμ2C337S domain, (*B*) Cμ3 domain, (*C*) Cμ3C414S domain, (*D*) Cμ4, (*E*) Cμ4tp domain, and (*F*) Cμ4tpC575S domain. The proteins were loaded with 1 μM (black line), 10 μM (red line), and 100 μM (blue line). The fluorescence signal was normalized to the highest peak intensity.

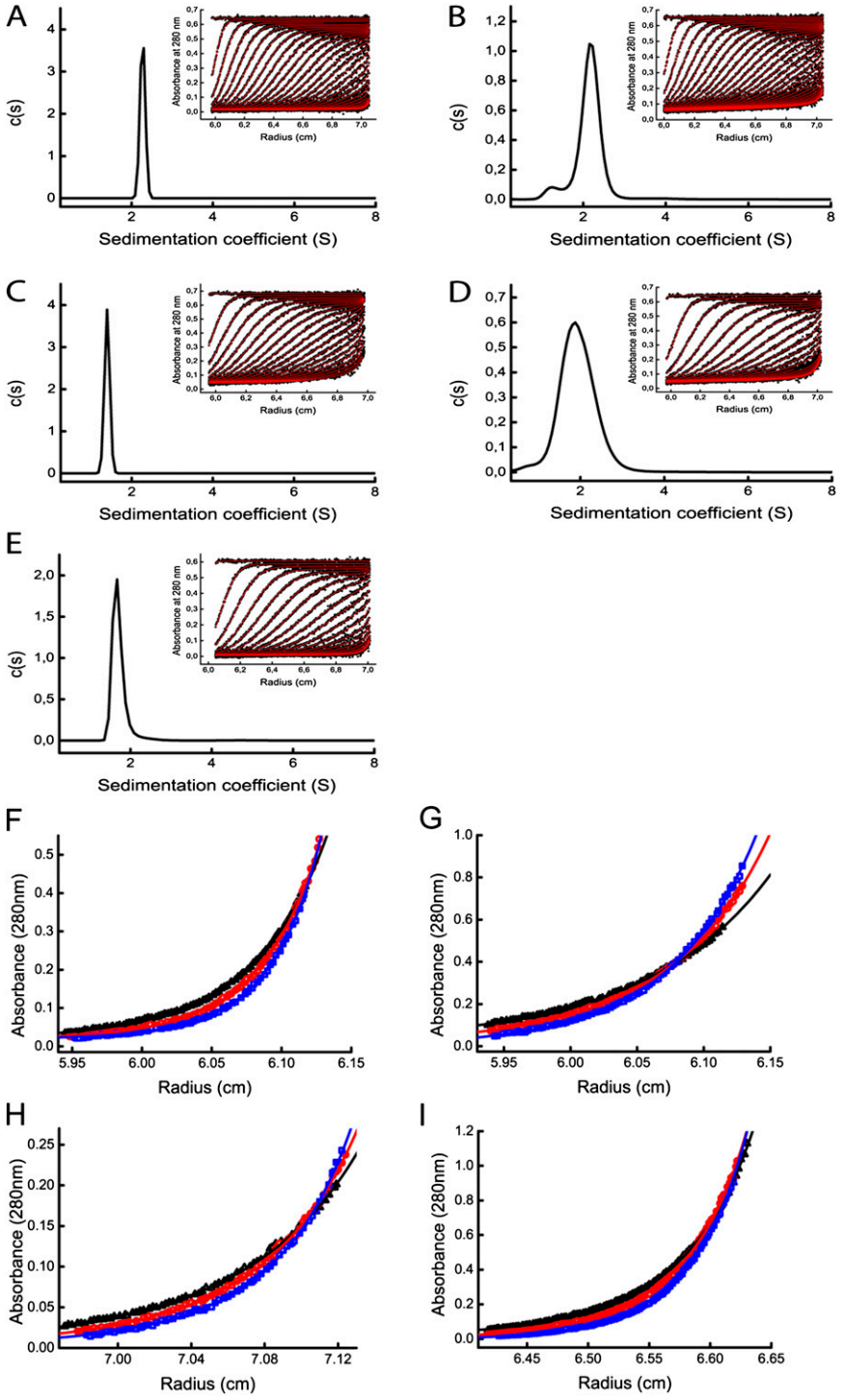

Fig. S3. Sedimentation velocity analytical ultracentrifugation runs of (*A*) Cμ2, (*B*) Cμ2C337S, (*C*) Cμ3C337S, (*D*) Cμ4, and (*E*) Cμ4tpC575S. Runs were performed at 50,000 rpm in PBS buffer. Representative analytical ultracentrifugation sedimentation equilibrium runs of (*F*) Cμ2C337S (10 μM), (*G*) Cμ3C414S (25 μM), (*H*) Cμ4 (5 μM), and (*I*) Cμ4tpC575S (20 μM). Runs were performed at 34,000 rpm (black triangles), 38,000 rpm (red circles), and 42,000 rpm (blue squares) in PBS buffer. Data were fitted globally to a self-association monomer–dimer equilibrium model (solid lines).

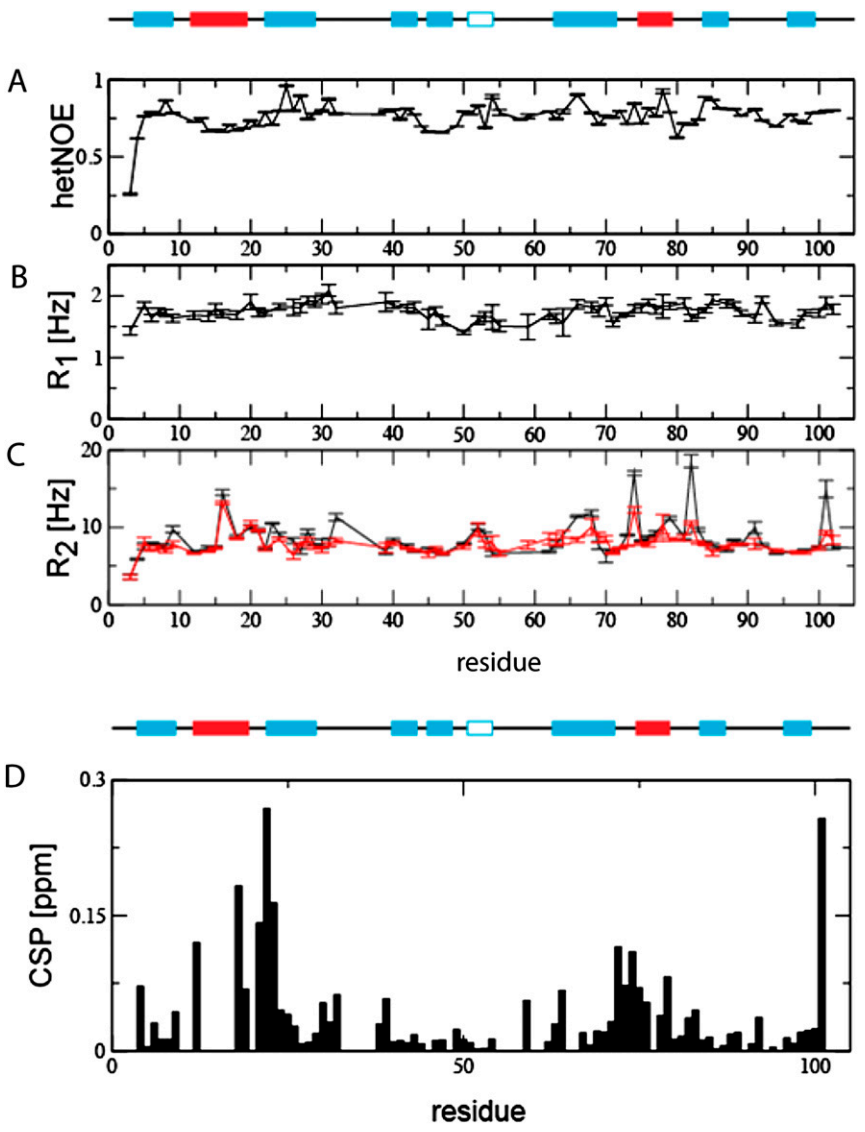

Fig. S4. Experimental ^{15}N backbone amide relaxation parameters and chemical shift perturbation (CSP) of the free Cμ3C414S domain compared with the Cμ3C414S–Cμ4 dimer. (*A*) hetNOE values, (*B*) longitudinal (R_1) rates, and (*C*) comparison of transverse relaxation rates R_2, measured directly (black) and derived from $R_{1\rho}$ (red). All experiments have been recorded at a Larmor frequency of 600 MHz at 298 K. Correlation times values were determined from ^{15}N relaxation data as described in *SI Materials and Methods*. (*D*) Chemical shifts in free form (Cμ3C414S) and the tandem construct (Cμ3C414S–CH4) have been assigned using 3D HNCA spectra. The secondary structure is pictured above (α-helixes in red, β-sheets in blue).

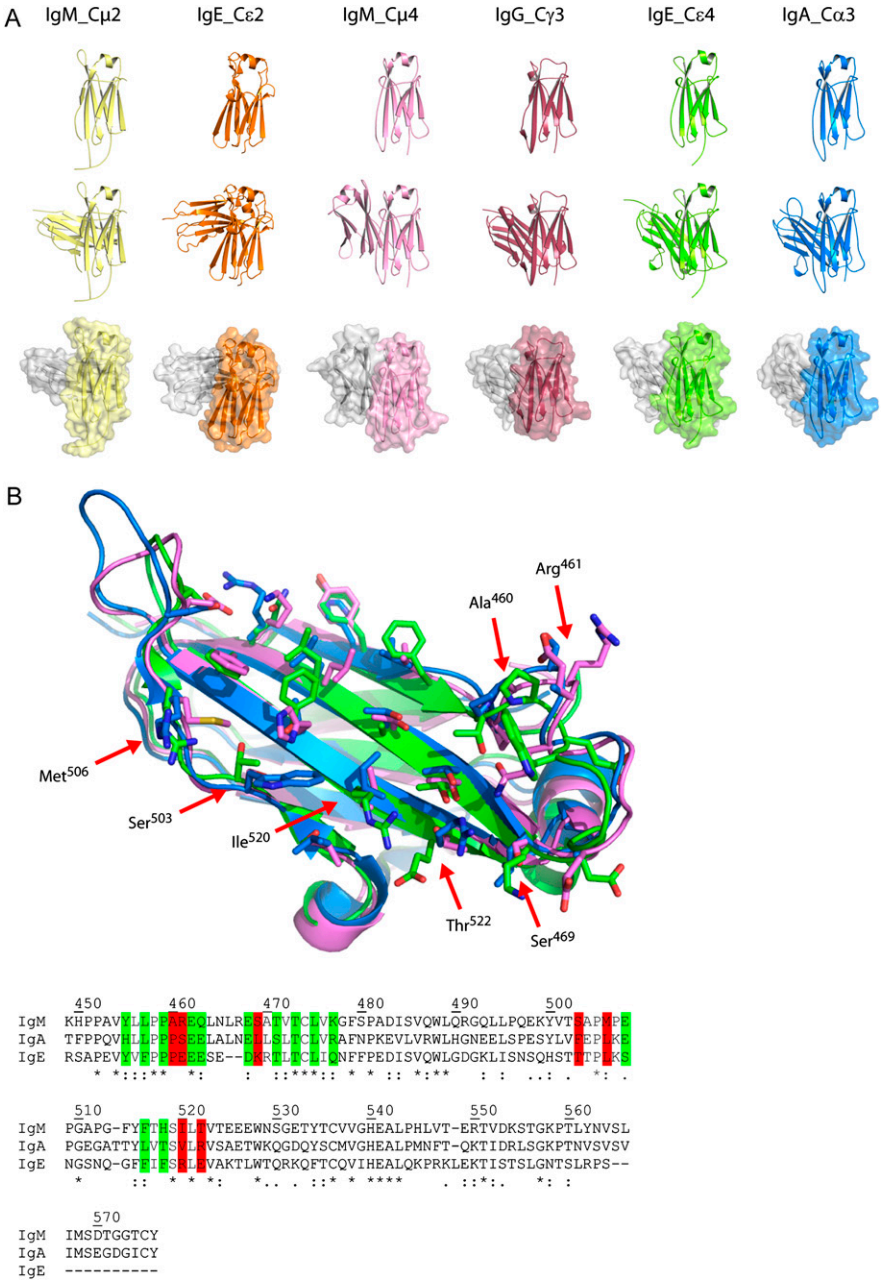

Fig. S5. Comparison and alignments of Ig constant domains. (*A*) Comparison of the various Ig dimerization modes (cartoon representation). (*Top*) Display of the respective chain A. (*Middle*) Display of the dimer form. (*Bottom*) Surface representation of the dimers. (*B*) Structure and sequence alignments of IgM Cμ4 (pink), IgA Cα3 (blue), and IgE Cε4 (green). The amino acids involved in the dimer interface are colored in green. Interface residues that are not conserved are shown in the alignment in red and highlighted in the structure by red arrows. The following coordinates were used (IgA, 1OW0; IgE,1O0V; IgG, 3HKF).

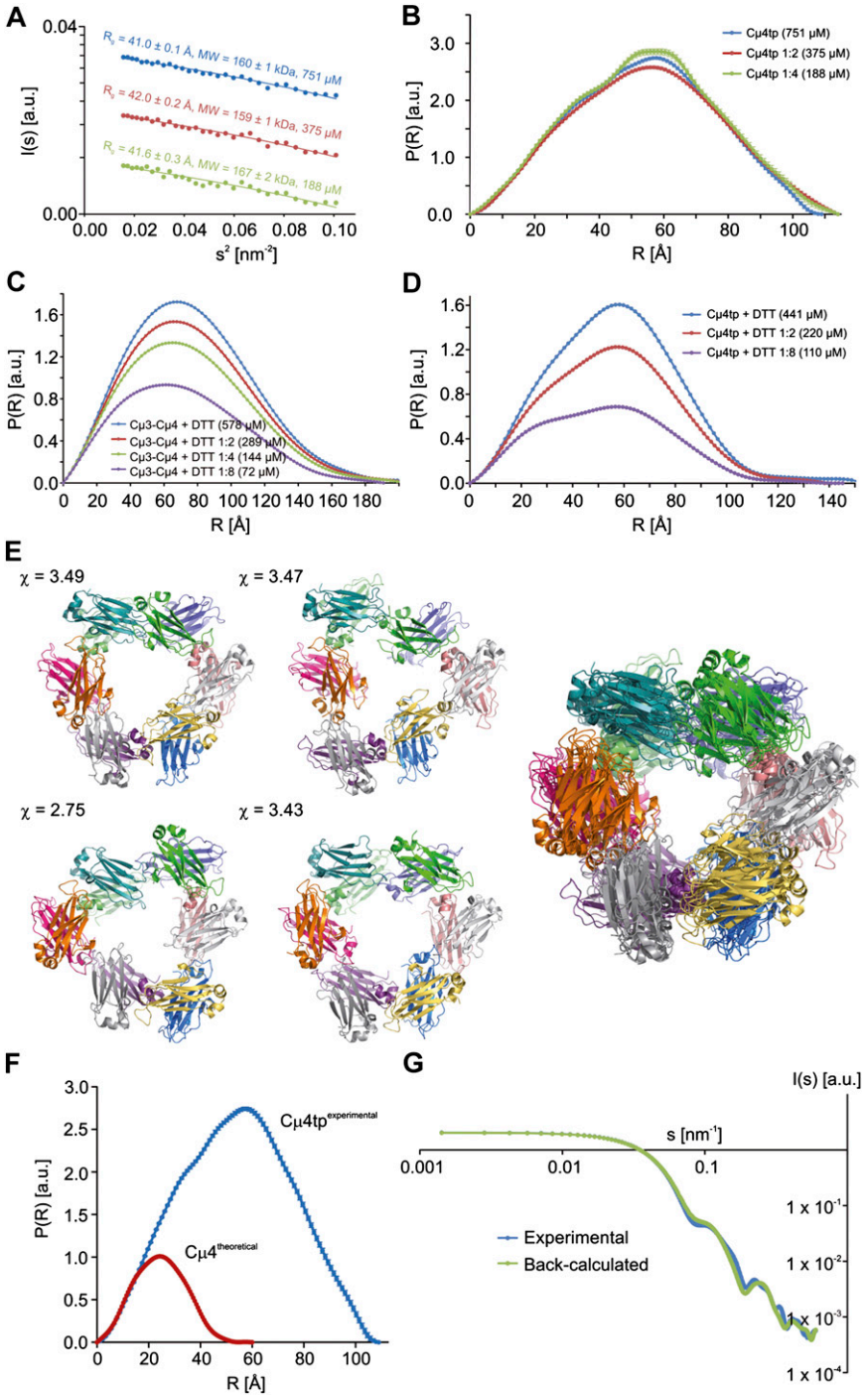

Fig. S6. Summary of the SAXS data and modeling. (*A*) Guinier plots of unmerged Cμ4tp scattering data collected at different concentrations. The R_gs and calculated molecular weights (based on BSA standard) are shown. (*B–D*) SAXS data showing a comparison of the experimental radial density distributions of Cμ4tp (*B*) and Cμ4tp (*C*) and Cμ3C414S–Cμ4 (*D*) in the presence of 5 mM DTT at different concentrations. (*E*) Modeling of the hexamer structure using the program CORAL and defining C3 symmetry. Different combinations of the additional Cμ4 interface observed in the crystal structures were used as restraints (*SI Materials and Methods* gives details). The fit (χ) between the experimental and back-calculated data are shown. (*F*) SAXS data showing a comparison of the experimental radial density distribution of Cμ4tp (blue) and the Cμ4 dimer (red) [back-calculated from the crystal structure determined here, using the program Crysol (1)]. (*G*) Comparison of experimental Cμ4tp SAXS data with SAXS data back-calculated from the CORAL model of the Cμ4tp hexamer. Both the s and I(s) axes are shown in a logarithmic representation. The angular ranges from 0.0014–0.6 nm^{-1} are compared.

1. Svergun DI, Barberato C, Koch MHJ (1995) CRYSOL – A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst* 28:768–773.

Table S1. Crystallographic statistics

Parameters	Cμ2 domain	Cμ4 domain
Crystal parameters		
Space group	C2	C2
Cell constants	a = 92.44 Å, b = 44.78 Å, c = 54.57 Å α = γ = 90°; β =118°	a = 169.16 Å, b = 41.21 Å, c = 67.1 Å α = γ = 90°; β = 92.3°
Data collection		
Beamline	X06SA, SLS	CuKα, rotating anode
Wavelength, Å	1.0	1.5418
Resolution range, Å*	40–1.30 (1.4–1.30)	40–2.0 (2.1–2.0)
No. of observations	15,5963	15,4481
No. of unique reflections†	48,207	31,721
Completeness, %*	99.3 (99.6)	100 (100)
R_{merge}, %*,‡	3.8 (30.9)	4.6 (35.4)
I/σ (I)*	16.1 (5.2)	19.0 (3.1)
Refinement (REFMAC5)		
Resolution range, Å	15–1.3	15–2.0
No. of reflections working set	45,796	28,566
No. of reflections test set	2,414	1,537
No. of nonhydrogen	1,615	3,356
No. of solvent water	291	428
R_{work}/R_{free}, %§	13.8/16.8	19.7/23.4
Rmsd bond, Å¶/angles, °	0.019/1.857	0.007/1.256
Average B-factor, Å^2	20.0	26.1
Ramachandran plot, %‖	99.0/1.0/0.0	97.9/2.1/0.0
PDB ID code	4JVU	4JVW

*The values in parentheses of resolution range, completeness, R_{merge}, and I/σ (I) correspond to the last resolution shell.
†Friedel pairs were treated as different reflections.
‡$R_{merge}(I) = \sum_{hkl}\sum_j |[I(hkl)_j - I(hkl)]| / [\sum_{hkl} I_{hkl}$, where $I(hkl)_j$ is the j^{th} measurement of the intensity of reflection hkl and <I(hkl)> is the average intensity.
§$R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where R_{free} is calculated without a sigma cutoff for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections.
¶Deviations from ideal bond lengths/angles.
‖Number of residues in favored region/allowed region/outlier region.

Table S2. NMR and refinement statistics for Cμ3C414S domain

Parameters	For all residues	For residues 3-29,38-51,66-100
NOE-based distance restraints		
Intraresidual, sequential	772	
Medium range ($2 \le |i-j| \le 4$)	158	
Medium range ($|i-j| \ge 5$)	563	
Total	1,493	
Violated	0	
Other restraints		
ϕ + ψ dihedral restraints	100	
Coordinate precision		
Backbone, A	0.49 ± 0.10	
Heavy atom, A	0.94 ± 0.11	
Common Interface for NMR structure Generation		
Red, %	16	11
Orange, %	18	14
Green, %	67	67
Whatcheck		
First-generation packing quality	2.918 ± 1.251	
Second-generation packing quality	4.617 ± 1.785	
Ramachandran plot appearance	0.886 ± 0.227	
χ_1/χ_2 rotamer normality	−3.471 ± 0.486	
Backbone conformation	1.647 ± 0.186	
Ramachandran plot, %		
Most favored regions	91.9	
Allowed regions	7.0	
Generously allowed regions	0.9	
Disallowed regions	0.3	
PDB ID code	4BA8	

Other Supporting Information Files

Dataset S1 (TXT)