Optimal Interstrand Bridges for Collagen-like Biomaterials

I. Caglar Tanrikulu and Ronald T. Raines*

Computational Methodology and Results

Computational Design and Evaluation of Disulfide Bridges Between $(PPG)_{10}$ Strands

All calculations were performed on Intel Xeon 2.33-GHz processors at the Materials and Process Simulation Center in California Institute of Technology (Pasadena, CA). Computational models were built on the crystal structure of the $(PPG)_{10}$ trimer (PDB entry 1kf6).[1] Hydrogens were added using Reduce[2] (ver.3.03), and the model was fully minimized. All minimizations were carried out to a 0.2 kcal/mol/Å RMS-force convergence criterion using conjugate gradient minimization on MPSim[3] without solvation. The forces on the model peptides were described by the DREIDING force-field[4] without atomic charges. Following initial optimization, the backbone coordinates of the trimeric structure were set to be immutable, and were kept that way throughout.

Proline residues in the Xaa and Yaa positions selected for linkage were replaced with Cys, Hcy, or Tnv. The new side-chains were minimized on trimers that contain both Cys-analogs on neighboring positions, and the pre-disulfide formation energies, $E_{x\cdot y}$ were determined. The thiol hydrogens were omitted from these models to maintain the same number of atoms before and after disulfide formation. The configuration of the disulfide-bonded linkers were optimized, first by minimization and then by multiple rounds of simulated annealing. Post-disulfide formation energies, $E_{x–y}$ were determined on the resulting models. The energy change due to disulfide formation was calculated as $E_{\text{strain}} = E_{x–y} – E_{x\cdot y}$. The intent of this metric is to quantify changes in bonding and steric interactions, and allow their comparison across linkers. The omission of solvation and Coulombic contributions from the model help reduce noise and focus results on local interactions.

Linker Strain in Relation to Torsion Angles

The *h–c* linker is predicted to be least strained by the computational analysis, as shown in Table S1. Strained bond lengths and bond angles are responsible for the poor performance of *c–c* and *c–h* linkers, whereas high torsion energies and, for *t–t* and *t–c*, van der Waals energies are observed for linkers other than *h–c*. Overall, linkers that have a longer side chain at the Xaa position than Yaa ($n_{\text{Xaa}} > n_{\text{Yaa}}$) perform better than those with shorter ($n_{\text{Xaa}} < n_{\text{Yaa}}$) or analogous side chains ($n_{\text{Xaa}} = n_{\text{Yaa}}$). The torsion angles across each C–C, C–S, or S–S bond on the linker models are presented in Table S2. The C–C and C–S torsion angles on the linkers can be categorized into two groups: near-eclipsed ($|\chi|$ = 0°–30° or 90°–150°) or near-staggered ($|\chi|$ =

30°–90° or 150°–180°). The high torsion energies correlate with the occurrence of near-eclipsed torsion angles across C–C bonds, most often between C^{β} and C^{γ}. The *h*–*c* linker is the only one that does not harbor strained torsion angles according to this simple criterion.

Experimental Methodology and Results

Peptide Synthesis

All $(PPG)_{10}$ and $(PPG)_{10}$-variants were synthesized on Wang or polyethylene glycol-based resins using a Prelude (Protein Technologies) peptide synthesizer at room temperature using standard Fmoc chemistry at the University of Wisconsin–Madison Biotechnology Center (www.biotech.wisc.edu/services/peptidesynth). Condensation of Fmoc-ProProGly-OH tripeptide segments was employed with all peptides wherever applicable, except for the C-terminal PPG section where proline residues were added individually on glycine-preloaded resins. Synthesis of Fmoc-ProProGly-OH was described previously,[5] and Fmoc-*S*-trityl-L-homocysteine was from Chem-Impex (Wood Dale, IL). Fmoc removal was achieved in piperidine (20% v/v in DMF), and peptide building blocks (4 equiv), activated through treatment with HCTU and NMM, were coupled to the free amine of the growing chain for 60 min.

Peptides were cleaved from the resin and deprotected in reagent R (90:5:3:2 TFA:thioanisole:ethanedithiol:anisole; 1.5-2.0 ml), precipitated from methyl *t*-butyl ether below 0 °C, and isolated by centrifugation. Dried crude peptides were dissolved in 0.1% v/v TFA and filtered and purified by preparative HPLC at 45 °C using gradients of CH_3CN/water containing 0.1% v/v TFA on a Shimadzu Prominence unit equipped with a Macherey–Nagel VarioPrep 250/21 C18 column. All peptides were >90% pure according to analytical HPLC and MALDI–TOF mass spectrometry (MS). MALDI–TOF analysis was carried out on an Applied Biosystems Voyager DE-Pro mass spectrometer at University of Wisconsin–Madison Biophysics Instrumentation Facility (BIF; www.biochem.wisc.edu/bif). Both single peptides and linked-dimers were analyzed using a 10:1 matrix mixture of 2-(4-hydroxyphenylazo)benzoic acid: α-cyano-4-hydroxycinnamic acid to suppress the reduction of disulfide bonds due to in-source decay.[6]

Synthesis of Linked-Dimers

Disulfide-linked strand dimers were produced by coupling 2,2′-dithiobis(5-nitropyridine)-(DTNP)-activated thiols on an *s3y* strand with free thiols on an *s1x* strand (Figure 3), as described previously.[7] Briefly, dried *s3y* peptides (1.2–1.4 mM) and DTNP (5 equiv) were dissolved in reaction solvent (3:1 v/v HOAc:water; degassed and argon-saturated), and the resulting solution was stirred vigorously for ≥6 h. The reaction was stopped by addition of 1.5 reaction-volumes of HOAc. The reaction mixture was freeze-dried, sonicated in aqueous 0.1% v/v TFA, and filtered through a 0.2-μm membrane. DTNP and 5-nitro-2-pyridinethiol remaining in the freeze-dried filtrate were removed by washing (3×) the lyophilized powder with diethyl ether:methylene chloride (7:3, v/v). Isolated *s3y*-Npys was freeze dried and weighed.

Coupling of *s1x* and *s3y*-Npys was initiated by mixing equimolar amounts of the components (2.2 mM final concentration) dissolved previously in degassed and argon-saturated 50 mM NH_4OAc buffer, pH 5.3. Reactions were stirred under argon for ≥6 h, after which the solution

was acidified and the solvent was removed by lyophilization. Disulfide-linked *s1x–s3y* dimers were isolated by HPLC and analyzed by MALDI–TOF mass spectrometry as described above. For *x–y* linked-dimers, (*m*/*z*) $[M + H]^+$ calcd 5072.8, found 5073.2 for *c–c*; calcd 5086.8, found 5086.1 for *c–h*; calcd 5100.8, found 5100.6 for *h–h*; calcd 5086.8, found 5086.3 for *h–c*; calcd 2530.3, found 2530.0 for *s2*. Analytical HPLC results for purified linked-dimers in 50 mM HOAc are presented in Figure S1.

Sample Preparation

Samples were prepared and experiments were conducted as reported previously.[8] All peptides and linked dimers were dissolved in 50 mM HOAc to 0.7 mM concentration based on weight. Many of the critical experiments in this study involve mixtures of peptides and linked dimers, and preparation of equimolar mixtures of the components is of vital importance. To ensure equimolar mixtures, relative concentrations were determined by monitoring peptides at 214 nm on the HPLC, and concentrations of all components in interrogated samples were matched to that of 60 μM *s2* peptide. The extinction coefficients of all strands were assumed to be identical at 214 nm. The common presence of *s2* in all mixtures allows for consistent concentrations across all samples prepared. To facilitate the formation of the thermodynamic product for peptide association, mixtures were heated to ≥55 °C, and annealed to 4 °C over 4 h. Samples were left at 4 °C for at least 48 h before data acquisition.

Circular Dichroism (CD) Spectroscopy

CD spectra for all samples were acquired at 4 °C with a 1-nm band-pass in 0.1-cm pathlength quartz cuvettes , using an averaging time of 3 s. For thermal denaturation experiments the CD signal was monitored at 226 nm while the sample was heated from 4 to 64 °C in 3-°C steps over 4.5 h. All samples were prepared at a strand concentration of 180 μM except for *s1x* + *s2* + *s3y* mixtures, which were prepared at a total peptide concentration of 270 μM in anticipation of complex melting transitions. Due to weak signal, the *s1h* peptide was tested at 360 μM as well as 180 μM to allow for a more accurate determination of the T_m value. CD data from denaturation experiments were converted to fraction folded, and data near 50% folded were used to obtain the T_m values for each sample, which are reported in Table S3.

CD data were acquired on CD spectrometers from Aviv Biomedical (Lakewood, NJ). These spectrometers were an Aviv 202SF and 420 at the BIF, and an Aviv 420 in the Gellman Laboratory of the Department of Chemistry, all of them equipped with a 5-cell thermoelectric sample changer. Data on 180 μM *s2* were collected in every experiment to allow comparison of instrument performance and to account for differences in wavelength calibration. Our results were well reproducible between instruments (T_m = 36.5 ± 0.6 °C for 180 μM *s2*; n = 5). All CD spectra and melting curves discussed in this study are shown in Figure S2.

Analytical Ultracentrifugation (AUC)

Sedimentation equilibrium experiments were performed at the BIF with a Beckman XL-A analytical ultracentrifuge equipped with an An-60 Ti rotor. Samples were prepared at a strand concentration of 180 μM, but were diluted to a concentration of 90 μM before the experiment. Sample (100 μl) and buffer (110 μL) were placed in a cell with a 12-mm double-sector charcoal-filled centerpiece (Epon). Experiments were run at 4 °C at speeds of 12, 22, 32, and 42 k rpm,

and gradients were monitored at 231 nm until they were superimposable when recorded 4 h apart. A buffer density of 1.00037 g/ml and a partial specific volume of 0.7275 ml/g calculated based on amino acid content for $(PPG)_{10}$ was used. Equilibrium gradients at 4 °C were modeled as single and multiple non-interacting species through nonlinear least-squares fits to gradient data. Analysis was performed with programs written for IGOR PRO (WaveMetrics, Lake Oswego, OR) by D. R. McCaslin (University of Wisconsin). Non-sedimenting baselines between 0.03–0.05 OD were applied for all samples, whereas 0.07 OD was used for the *x–y·s2* construct featuring a *c–c* linker. Plots of gradients and fits are shown in Figure S3.

Overall, the data collected on *x–y·s2* variants fitted best to a triple-helical model (single species with MW = 7.6 kDa). A similar model does not explain the *s2* data. In addition to triple helices, consideration of free $(PPG)_{10}$ strands in the model was necessary for satisfactory description of *s2* gradients. Among *x–y·s2* variants, the samples that feature *h–c*, *c–h* and *h–h* linkers behave as triple-helical units. Even though the *c–c* case forms gradients similar to others, models that consider triple-helices work best either with large baseline corrections, or when strand-dimers are considered in the model in addition to triple helices. It is likely that the actual composition for this sample includes triple helices, together with low levels of strand-monomers and dimers. Such issues are not observed with linkers predicted to be less strained.

References

(1) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2002**, *11*, 262–270.

(2) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735–1747.

(3) Lim, K. T.; Brunett, S.; Iotov, M.; McClurg, R. B.; Vaidehi, N.; Dasgupta, S.; Taylor, S.; Goddard, W. A. *J. Comput. Chem.* **1997**, *18*, 501–521.

(4) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(5) Jenkins, C. L.; Vasbinder, M. M.; Miller, S. J.; Raines, R. T. *Org. Lett.* **2005**, *7*, 2619–2622.

(6) Huwiler, K. G.; Mosher, D. F.; Vestling, M. M. *J. Biomol. Tech.* **2003**, *14*, 289–297.

(7) (a) Rabanal, F.; DeGrado, W. F.; Dutton, P. L. *Tetrahedron Lett.* **1996**, *37*, 1347–1350. (b) Ottl, J.; Moroder, L. *J. Am. Chem. Soc.* **1999**, *121*, 653–661.

(8) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3028–3033.

Table S1. Breakdown of E_{strain} into Force-Field Energy Components

Linker type	CH_2 groups on linker, $n_{Xaa}+n_{Yaa}$	E_{strain} components (kcal/mol)				
		Total	Bonds	Angles	Torsions	Van der Waals
h–c	3	3.8	0.2	0.7	2.3	0.6
t–t	6	7.9	0.2	1.6	2.6	3.5
t–c	4	8.0	0.4	1.2	2.5	3.9
t–h	5	8.7	0.3	1.2	5.6	1.7
h–h	4	9.2	0.2	1.2	5.8	2.0
c–t	4	10.7	0.3	1.6	6.7	2.3
h–t	5	11.0	0.5	1.9	5.0	3.6
c–h	3	11.9	0.6	4.1	5.3	1.9
c–c	2	15.9	1.2	9.1	4.5	1.2

Table S2. Side-chain[a] and Disulfide[b] Torsion Angles Along the Linker as Obtained from Computational Models Reported in Degrees

	c–c	*h–c*	*c–h*	*h–h*	*c–t*	*h–t*	*t–t*	*t–h*	*t–c*
Xaa	Cys	Hcy	Cys	Hcy	Cys	Hcy	Tnv	Tnv	Tnv
C^{α}									
|	42.5	49.8	42.2	65.6	53.8	43.1	81.9	49.2	68.1
C^{β}									
|	81.1	73.8	**112.5**	**102.8**	**113.2**	160.1	159.4	152.4	–175.4
C^{γ}									
|		–160.9		-168.3		–76.2	**–146.9**	–83.2	–165.0
C^{δ}									
|							–69.7	**–90.1**	–60.5
S									
|	**–146.5**	129.6	–114.4	81.5	**–153.3**	–90.9	126.6	94.7	106.3
S									
|					60.7	**92.9**	–87.5		
C^{δ}									
|			–49.2	**–123.7**	**–104.3**	**–113.9**	–62.0	**–122.9**	
C^{γ}									
|	69.3	–79.2	**104.4**	**129.7**	**135.7**	**149.1**	169.1	**149.2**	**–132.5**
C^{β}									
|	**–127.6**	–71.4	–154.0	–158.9	–167.0	–170.1	–176.4	–168.2	–44.2
C^{α}									
Yaa	Cys	Cys	Hcy	Hcy	Tnv	Tnv	Tnv	Hcy	Cys

[a] Side-chain torsion angles across C–C and C–S bonds that are closer to the eclipsed rather than the staggered conformation are shown in bold typeface.

[b] Torsion angles across the disulfide bond are underlined for each linker, and those angles closer to energy maxima (0° and 180°) rather than minima (±90°) are shown in bold typeface.

Table S3. Values of T_m (°C) for Hetero- and Homotrimers as Determined by Thermal Denaturation Experiments Monitored by Circular Dichroism Spectroscopy

Linker type	*x–y·s2*	Strands that constitute *x–y·s2*	*s1x·s2·s3y (calculated)*[a]	*s1x + s2 + s3y mixture*
		s1x , *s2* , *s3y*		
c–c	28	25 , 37 , 29	30	30
h–c	35	22 , 37 , 29	29	30
c–h	28	25 , 37 , 33	31	31
h–h	27	22 , 37 , 33	31	32

[a] The estimate of the T_m for "reduced" *x–y·s2* was calculated as the average of the homotrimer T_m values for the related *s1x*, *s2* and *s3y* strands.

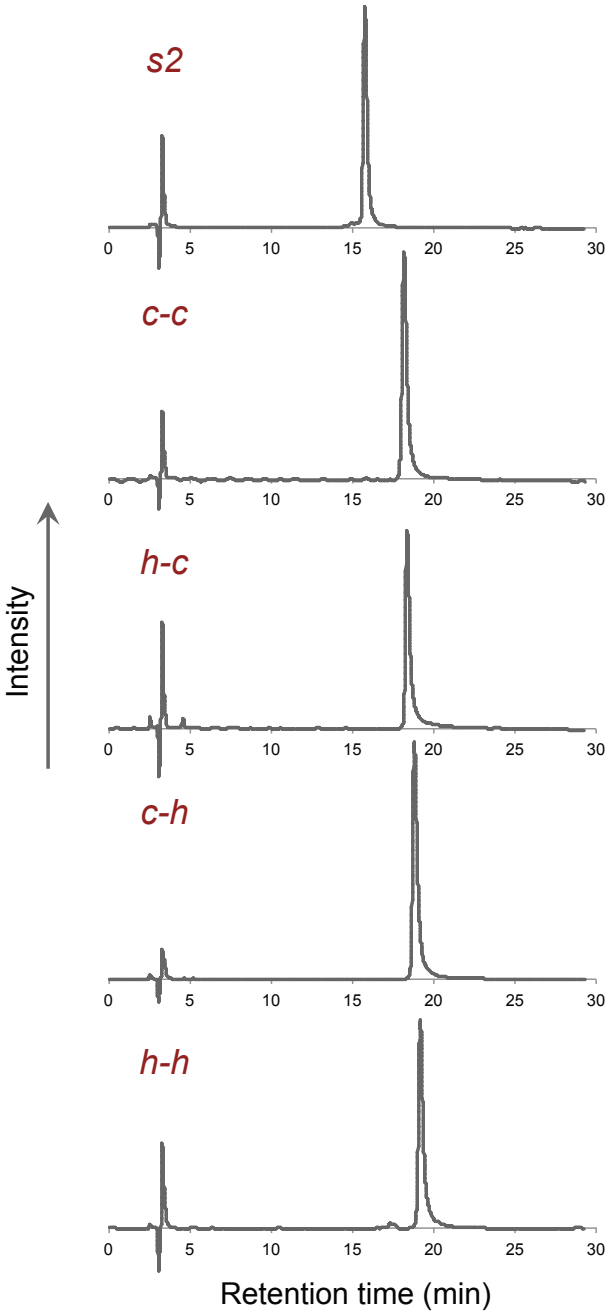

Figure S1. Analytical HPLC chromatograms for linked-dimers and the *s2* peptide in 50 mM HOAc, as acquired after purification.

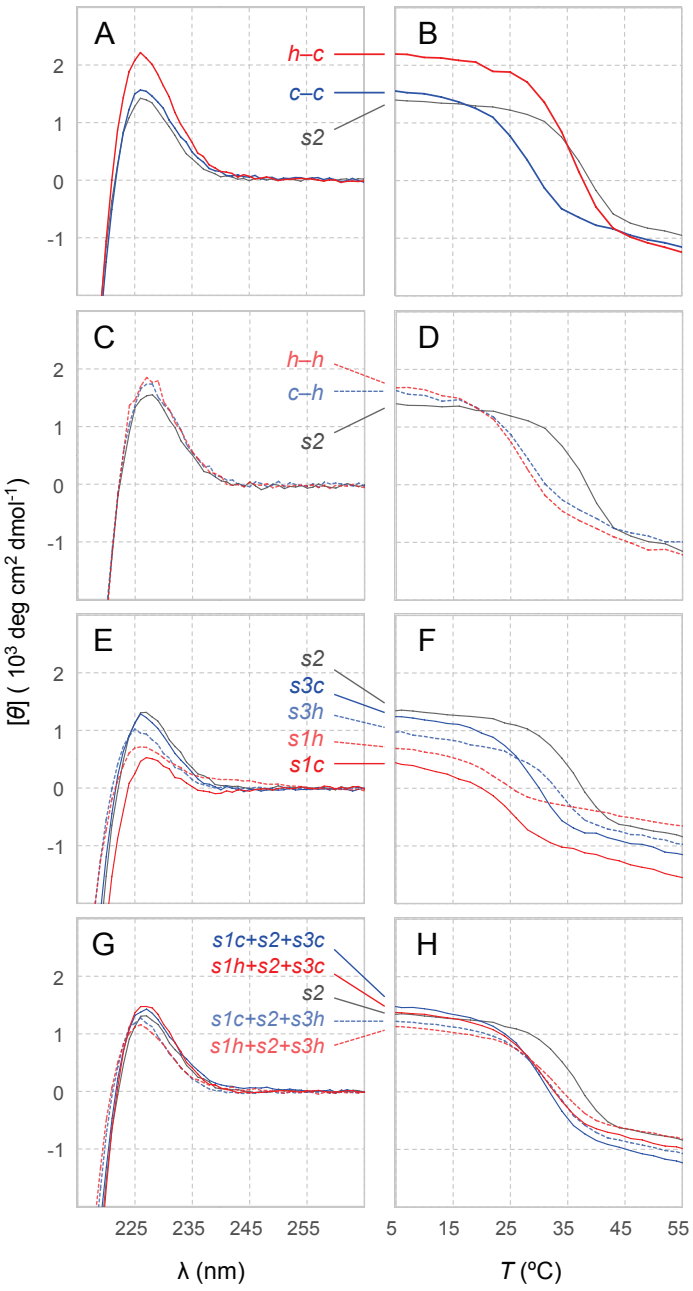

Figure S2. Circular dichroism spectra and melting profiles for linked-dimer – *s2* mixtures and individual strands. Each panel displays data acquired on the same instrument: (A) CD spectra and (B) temperature melts for *x–y·s2* featuring *c–c* and *h–c* linkers, (C) CD spectra and (D) temperature melts for *x–y·s2* featuring *c–h* and *h–h* linkers, (E) CD spectra and (F) temperature melts for homotrimers of individual strands, and (G) CD spectra and (H) temperature melts for *s1x+s2+s3y* mixtures.

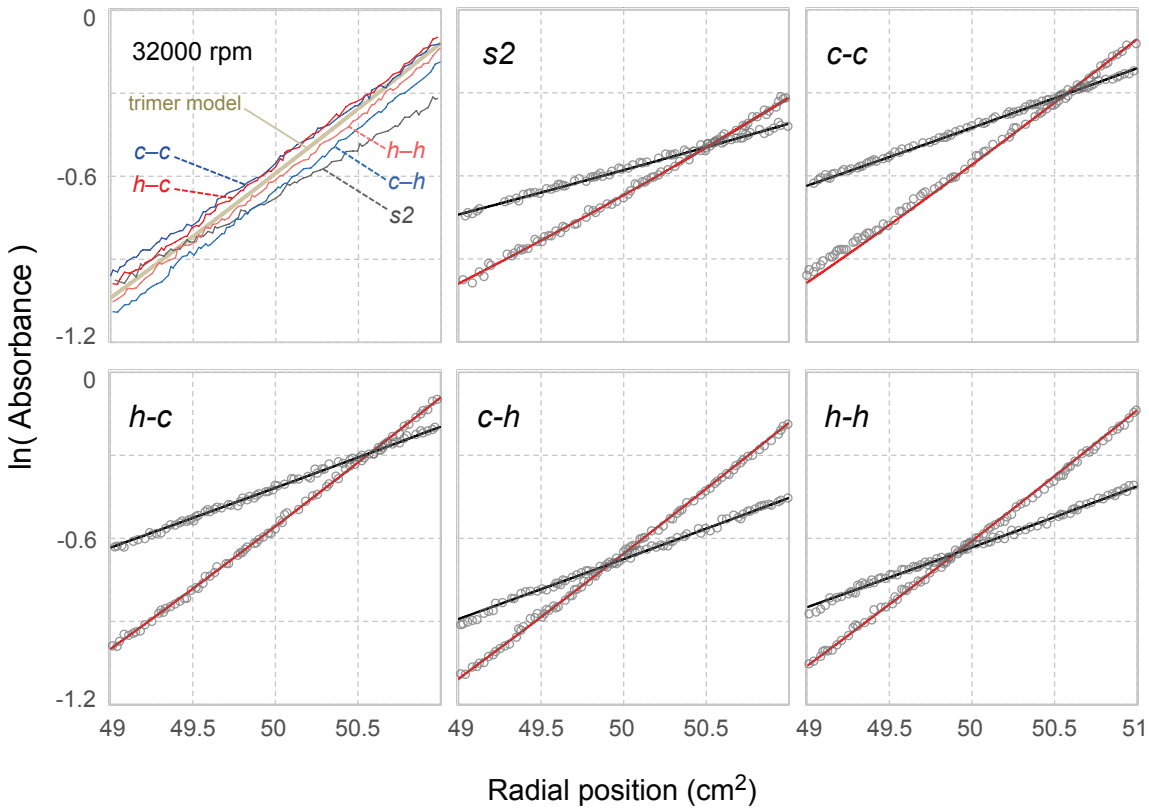

Figure S3. Sedimentation equilibrium analysis of *x–y·s2* constructs. Equilibrium gradients for linked-dimer – *s2* mixtures are shown with models that provide optimal fits. The first panel compares raw data collected at 32 k rpm for all samples with the theoretical model of a trimer, and highlights the difference between gradients for *s2* and *x–y·s2* constructs. The slope of the computed model represents the gradient expected from a pure, three-stranded unit, whereas its placement on the y-axis is arbitrary. All *x–y·s2* constructs form gradients that are similar and match the model predicted for a triple-helical unit. The remaining five panels show gradients formed at 22 and 32 k rpm (gray circles) and models that best fitted the data (black and red lines for 22 and 32 k rpm, respectively). The *s2* data are best described by a mixture of monomers and triple helices, whereas the remaining sets are modeled as single triple-helical species. Although the *c–c* data suggest the presence of some lower molecular-weight species in solution, constructs featuring the *h–c*, *c–h*, and *h–h* linkers are in good agreement with a triple-helical model.