Supporting Information

Larson et al. 10.1073/pnas.0912293107

SI Methods

Construction of Expression Plasmids. The primary sequence features of AgI/II and extents of the recombinant fragments used in this study are shown in Fig. 1. Plasmids pDC20 and pRR2 that encode full-length and aa 84–1223 of AgI/II (1, 2), respectively, pCK1 that encodes A_1VP_3 linked to an amino-terminal MBP tag (2), and pSBR (3) that encodes the alanine-rich repeats A_{123}, were used as templates for PCR. Construction and expression of the recombinant his-tagged full-length AgI/II (CG14) was described previously (1). DNA primer sequences for PCR, with appropriate restriction sites for cloning underlined, are listed in Table S2. PCR reactions were carried out using the PTC-200 thermocycler (MJ Research) and the DeepVent proofreading polymerase (New England Biolabs). DNA ligations were carried out with T4 DNA ligase (New England Biolabs). MBP–A_1VP_1, A_1VP_3, V, VP_3, and C-term fragments were subcloned into pET23d (Novagen) using endonuclease sites BamHI and XhoI, whereas A_{1-3} was subcloned into pET20b (Novagen) using endonuclease sites NcoI and XhoI. The resulting recombinant plasmids were used to transform *E. coli* XL1-blue cells (Stratagene). Colonies were screened for plasmids with the correct size inserts, and these were confirmed by sequencing (University of Alabama at Birmingham Sequencing Center). These plasmids were used subsequently to transform the *E. coli* BL21 (λ*DE3*) expression strain.

Expression and Purification of A_3VP_1: MBP–A_1VP_1 protein was expressed in *E. coli* BL21 (λ*DE3*) according to the manufacturer's instruction. The cells were lysed using a Sonic Dismembranator model 500 (Fisher Scientific) on ice, and the lysate was purified with a HisPrep Nickel column (Amersham Biosciences), followed by purification over an amylose resin column (New England Biolabs). MBP–A_1VP_1-containing fractions were pooled and dialyzed overnight in 20 mM hepes pH 7.5, 100 mM sodium chloride, and 2 mM $CaCl_2$. The next day, Factor Xa protease (Sigma-Aldrich) was added (50 ng Factor Xa per mg MBP–A_1VP_1) and digestion was continued over 4 days at 22 °C. Factor Xa digestion resulted in a 54.5-kDa protein fragment (size determined by mass spectroscopy). Western blot analysis with mouse monoclonal anti-penta-histidine (Novagen) or rabbit anti-MBP (New England Biolabs) antibodies, followed by alkaline phosphatase-labeled goat anti-mouse or anti-rabbit IgG (Novagen), revealed that overdigestion had occurred at the amino terminus, such that the resultant product comprises the third alanine-rich region (A_3), the variable region (V), and the first proline-rich region (P_1) (Fig. S2). Edman degradation analysis at the University of Florida Protein Core Facility confirmed Ala386 of AgI/II as the amino-terminal residue. This stable fragment was further purified with ion exchange over a MonoQ HR10/10 (Amersham Biosciences) and size exclusion over a Superdex 200pg 26/60 (Amersham Biosciences) column (Fig. S2). The A_3VP_1 fragment was concentrated to 15 mg/mL by an Amicon concentrator (Amicon) under nitrogen gas.

Expression and Purification of A_{1-3}, V, VP_3, A_1VP_3, and C-Term Polypeptides. A_{1-3}, V, VP_3, A_1VP_3, and C-term fragments of AgI/II were expressed in *E. coli* BL21 (λ*DE3*) similar to A_3VP_1. The recombinant *E. coli* cells were collected and lysed and AgI/II polypeptides were purified using a three-step protocol; lysates were sequentially subjected to a HisPrep affinity column (Amersham) followed by ion-exchange chromatography over a MonoQ/S column (Amersham) and then polished over a size exclusion column (Superdex 75/200pg 26/60, Amersham). The A_{1-3} purification used the MonoS and Superdex 75 columns whereas all other proteins were purified using the MonoQ and Superdex 200 columns. Proteins were concentrated to ∼5–15 mg/mL.

Analytical Ultracentrifugation. A_3VP_1, A_1VP_3, and CG14 polypeptides were dialyzed into 20 mM Tris pH 8.0, 150 mM sodium chloride, 1 mM EDTA to a final protein concentration of 0.6 mg/mL. Sedimentation velocity experiments were carried out using a Beckman Optima XL-A. During velocity runs at 40,000 rpm and 20 °C, frames of the absorbance at 280 nm across the cell were collected every 5 min. Sedimentation velocity data were fit using the program Sedfit (4) with the first seven frames of data removed and then every third frame of data used in the analysis. The partial specific volume (vbar) of 0.731 mL/g and buffer density of 1.00517 as well as hydration values for A_3VP_1, A_1VP_3, and CG14 of 0.45 g/g, 0.44 g/g, and 0.44 g/g, respectively, were calculated in Sednterp and then used in Sedfit to fit with a resolution of 100 (s-grid increments of 0.1 S) using a continuous c(s) distribution to model the velocity sedimentation and obtain the sedimentation and frictional coefficients. The coefficients of the largest peak (which contributed to 94% or more of the total protein concentration) were used to determine oblate and prolate ratios in Sednterp.

CD Spectroscopy. CD Spectra were recorded at 22 °C on the Olis DSM 100 circular dichroism spectrophotometer using a quartz cell (Starna Cells) with a path length of 0.2 mm. Proteins were diluted to 0.3–0.6 mg/mL in a buffer of 20 mM Tris at pH 8.0, 150 mM sodium chloride, and 1 mM EDTA. Spectra were recorded over 200–260 nm in 1-nm steps and averaged over 10 complete scans. The millidegree data were converted to $\Delta\varepsilon$ (liter m^{-1} cm^{-1}) and then analyzed using the CONTIN/LL algorithm implemented in CDPRO (5) with the SDP48 reference set of proteins.

Preparation of Salivary Agglutinin. Salivary agglutinin was prepared as previously described (6, 7). Unstimulated saliva was collected over a 1-h period from healthy volunteers into 50 mL conical tubes on ice. Saliva was clarified by centrifugation (8,000 × g) for 20 min at 4 °C. Clarified saliva was mixed with potassium-PBS (KPBS; 2.7 mM potassium chloride, 1.5 mM potassium phosphate, 137 mM sodium chloride, 6.5 mM sodium phosphate, pH 7.2) and *S. mutans* cells (at OD_{600} = 1.0) at a 1:1:1 (v:v:v) ratio and incubated end-over-end at 37 °C for 30 min. Cells were pelleted at 2,000 × *g* for 15 min, washed once with KPBS, and the adsorbed agglutinin eluted from the *S. mutans* cells in one volume of KPBS containing 1 mM EDTA. The cells were removed by centrifugation at 4,000 × *g* for 20 min and the supernatant-containing agglutinin was filter sterilized with a 0.2-μm filter, dialyzed overnight against KPBS, and stored at −20 °C.

Evaluation of AgI/II Epitopes by Competition ELISA. ELISA plate wells were coated with *S. mutans* whole cells (∼10^5 cfu/well) in carbonate-bicarbonate buffer (pH. 9.6), washed, and blocked with PBS (pH 7.2) containing 0.03% Tween-20 (PBS-Tween) as described (8). One hundred microliters of serial 2-fold dilutions beginning at 5 μM of each polypeptide diluted in PBS-Tween were incubated for 2 h at 37 °C with 100 μL of anti-AgI/II monoclonal antibodies (MAbs) 1–6F, 4–9D, 4–10A, or 3–10E (9) diluted in the same buffer and titrated to result in OD_{450} ∼1.0 in the absence of inhibitor. The source of MAbs was murine ascites fluids. After washing, MAb binding to the plate wells was traced with 100 μL of peroxidase-labeled goat anti-mouse IgG (MP Biomedicals) diluted 1:1,000 in PBS-Tween and incubated for 2 h at 37 °C. Plates

were washed and developed with *o*-phenylenediamine dihydrochloride substrate solution and the absorbance at 450 nm was recorded using an MPM Titertek model 550 ELISA plate reader (Bio-Rad). All assays were performed in triplicate. Percent inhibition was calculated as 100 − [(mean OD_{450} of MAb + test polypeptide/mean OD_{450} of MAb alone) × 100].

Surface Plasmon Resonance. Adherence of AgI/II polypeptides to human salivary agglutinin was assessed by surface plasmon resonance using the BIAcore 2000 (BIAcore AB). Purified SAG was immobilized to an RU value of 3,800 on a CM5 sensor chip via amine coupling, and the experiment or control surfaces were blocked with 1.0 M of ethanolamine. A total of 80 μL of each AgI/II polypeptide, in concentrations ranging from 0.25 to 4 μM in 10 mM hepes pH 7.4 and 150 mM sodium chloride (HBS) with 2.5 mM $CaCl_2$, were flowed at a rate of 20 μL/min over the sensor chip. An 8-min dissociation time was used to measure the slow dissociation rates observed. Between experiments, 10 μL of HBS augmented with 3 mM EDTA, 0.0005% Tween-20 (HBSEP) and 10 mM sodium hydroxide was flowed over the chip surface. This was followed by flowing 5 μL of 10 mM HCl over the surfaces to regenerate the cells. Each binding experiment was repeated three times in different series of analyte concentrations on the same chip surface as well as on a separately prepared CM5 chip surface that showed similar results. Signal from the noncoated control cell FC1 was subtracted from that of the SAG-coated FC2 surface to obtain the sensograms (ΔRU). The collected data were fitted using a 1:1 Langmuir kinetic model and the BIAEvaluation software version 4.2 (Biacore AB).

Isothermal Titration Calorimetry. Calorimetric titrations were carried out with a MicroCal ITC calorimeter (MicroCal). The VP_3 protein was loaded into the 1.342 mL stirred cell (at 400 rpm) with a concentration ranging from 7 to 11 μM and the A_{1-3} protein at an 8-fold higher concentration was loaded into a 250-μL syringe. Experiments were carried out at different temperatures (15, 20, 25, and 30 °C) with the proteins in 20 mM hepes (pH 7.0), 150 mM sodium chloride, and 1% (vol/vol) glycerol. After the initial injection of 1 μL (which was excluded in data fitting), 24 injections of 10 μL each were delivered at 300-s intervals. Data were fit to a single binding site model using Origin software (version 7.0383 with MicroCal ITC-analysis module) from which the ΔG, K_a, K_d, ΔH, ΔS, and stoichiometry (n) of the binding were obtained. The enthalpy of dilution of the A_{1-3} protein was measured by titration of A_{1-3} into buffer alone to be ∼4 kcal/mol, small relative to the large enthalpy changes obtained in titrations of A_{1-3} into VP_3. Within these experimental temperatures, ΔH is well approximated as a linear function of T, where $\Delta C_p = (\Delta H_{T2} - \Delta H_{T1}) / (T_1 - T_2)$ and ΔC_p is the change in heat capacity at constant pressure upon binding. This relationship was used to calculate the ΔC_p as -7.95 ± 0.58 kJ K^{-1} mol^{-1}.

Modeling PMS into the Structure of A_3VP_1: Residues 495–828 of the A_1VP_3 structures were highly similar to the solved AgI/II V-region structure (10) (average rmsd of 0.431 Å), however the residue Trp816, located inside a cleft of the V region, has a large side chain movement in the A_3VP_1 molecules compared to the AgI/II V-region structure. This tryptophan residue neighbors a large continuous difference density that suggested the presence of a single covalently linked molecule. As all crystal forms of A_3VP_1 were supplemented with PMSF (phenyl-methyl-sulfonyl fluoride) during purification, we modeled this site with a hydrolyzed form (PMS) of the inhibitor PMSF. The sulfate group of PMS is positioned near the positively charged side chain of Arg824, while the PMS benzyl group stacks to form a hydrophobic interaction with Trp816.

Solvent-Accessible Surface Area (ASA) of A_3VP_1 X-Ray Structure: ASA values were calculated from atomic positions of the Ag I/II structures using the software NAccess (ver. 2.1.1) with a probe size of 1.4 Å and Z-slices of 0.05 Å. To estimate the change in ASA (ΔASA) upon A–P association, the ASA changes were calculated using the structure of A_3VP_1 from the $P2_12_12$ crystal. Polar (ASA_{polar}) and nonpolar (ASA_{apolar}) surface areas were calculated for the associated A_3 and P_1 present in the structure, and then separately for only the A_3 residues (386–475) or the P_1 residues (836–874). The ΔASA for polar and nonpolar surfaces were then calculated as: $\Delta ASA_{A3P1} = ASA_{A3} + ASA_{P1} - ASA_{A3P1}$, resulting in 2515.2 $Å^2$ to 920.0 $Å^2$ nonpolar:polar contact surface.

1. Brady LJ, et al. (1998) Deletion of the central proline-rich repeat domain results in altered antigenicity and lack of surface expression of the *Streptococcus mutans* P1 adhesin molecule. *Infect Immun* 66:4274–4282.
2. McArthur WP, et al. (2007) Characterization of epitopes recognized by anti-*Streptococcus mutans* P1 monoclonal antibodies. *FEMS Immunol Med Microbiol* 50: 342–353.
3. Harokopakis E, Hajishengallis G, Greenway TE, Russell MW, Michalek SM (1997) Mucosal immunogenicity of a recombinant *Salmonella typhimurium*-cloned heterologous antigen in the absence or presence of coexpressed cholera toxin A2 and B subunits. *Infect Immun* 65:1445–1454.
4. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78:1606–1619.
5. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287:252–260.
6. Brady LJ, Piacentini DA, Crowley PJ, Oyston PC, Bleiweis AS (1992) Differentiation of salivary agglutinin-mediated adherence and aggregation of mutans streptococci by use of monoclonal antibodies against the major surface adhesin P1. *Infect Immun* 60:1008–1017.
7. Oli MW, McArthur WP, Brady LJ (2006) A whole cell BIAcore assay to evaluate P1-mediated adherence of *Streptococcus mutans* to human salivary agglutinin and inhibition by specific antibodies. *J Microbiol Methods* 65:503–511.
8. Crowley PJ, et al. (2008) Requirements for surface expression and function of adhesin P1 from *Streptococcus mutans*. *Infect Immun* 76:2456–2468.
9. Ayakawa GY, et al. (1987) Isolation and characterization of monoclonal antibodies specific for antigen P1, a major surface protein of mutans streptococci. *Infect Immun* 55:2759–2767.
10. Troffer-Charlier N, Ogier J, Moras D, Cavarelli J (2002) Crystal structure of the V-region of *Streptococcus mutans* antigen I/II at 2.4 A resolution suggests a sugar preformed binding site. *J Mol Biol* 318:179–188.
11. Larkin MABG, et al. (2007) ClustalW and ClustalX version 2. *Bioinformatics* 27: 2947–2948.

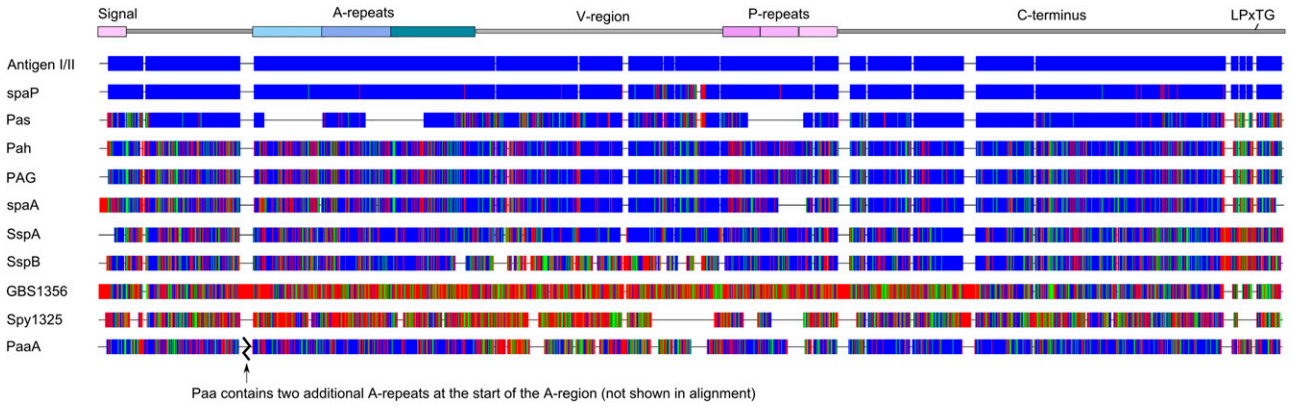

Fig. S1. The Clustalx2 (11) alignment of antigen I/II with well-characterized streptococcal AgI/II family proteins was transformed into the above graphical representation using AlignmentDraw (supplementary Perl script). *S. mutans* NG8 AgI/II (GenBank no. ACV69919.1) was aligned with *S. mutans* spaP (AAN58348.1), *S. intermedius* Pas (BAA96878.2), *S. downei* Pah (AB207813.1), *S. sobrinus* PAG (D90354.1), *S. sobrinus* spaA (X57841.1), *S. gordonii* SspA (AAC44099.1), *S. gordonii* SspB (AAC44100.1), *S. agalactiae* GBS1356 (CAD47015.1), *S. pyogenes* M28 Spy1325 (AAX72435.1), and *S. criceti* PaaA (AB042239.3). Positions with identical sequence are colored in blue, similar residues are colored in green, and dissimilar residues are shown red. Gaps are illustrated in white. A high level of variation in the V region is apparent in both Paa and SspB. Pas has deletions resulting in fewer A repeats and P repeats. Paa has a total of five A repeats, with two extra A repeats entailing a total of 164 residues occurring at the break (≀), which were omitted from the Paa sequence of this diagram for clarity on the alignment. Paa, however, only has three P repeats, which results in two unmatched A repeats. The *S. agalactiae* and *S. pyogenes* AgI/II family members show most dissimilarity with *S. mutans* AgI/II.

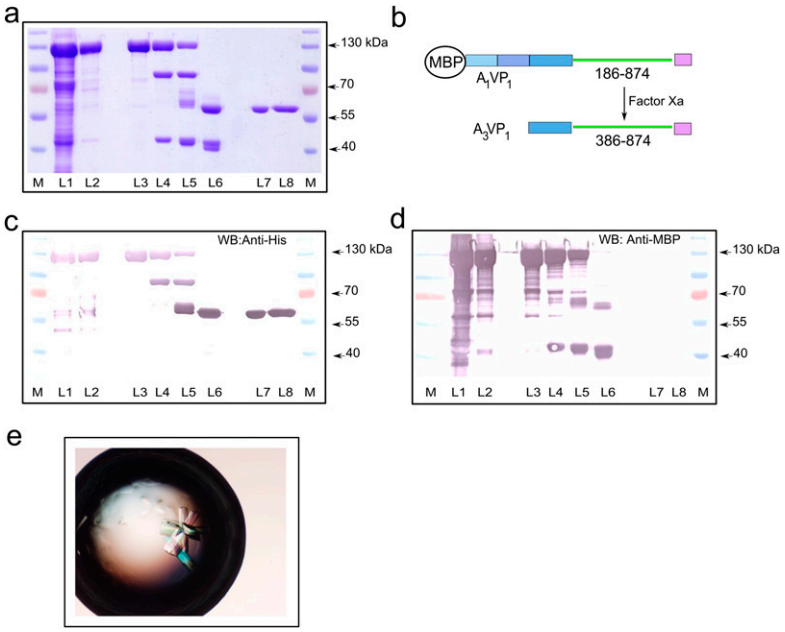

Fig. S2. Purification of A_3VP_1. (*A*) MBP–A_1VP_1 was purified from the initial lysate (L1) by nickel affinity (L2) followed by MBP affinity chromatography (L3), and the MBP moiety sequentially cleaved for 24 h (L4), 48 h (L5), and 72 h (L6) with Factor Xa. After cleavage, the remaining A_3VP_1 was further purified over a MonoS cationic affinity column (L7) and polished by size exclusion chromatography (L8). The final product of Factor Xa cleavage of MBP–A_1VP_1 is diagrammed in *B*. Western blots (WB) probed with an anti-histidine antibody (*C*) verified the intact carboxy terminus, whereas reactivity with anti-MBP (*D*) indicated that cleavage occurred at the amino terminus. The extent of proteolysis was evaluated by mass spectroscopy and Edman degradation confirmed the N-terminal sequence to correspond to A_3VP_1. A_3VP_1 migrates with an apparent molecular mass of ~60 kDa on an SDS/PAGE, and the molecular mass was determined to be 54.5 kDa by mass spectroscopy. (*E*) A_3VP_1 crystals formed in a 1:1 μL hanging drop by vapor diffusion with the reservoir solution containing 30% (vol/vol) polyethylene-glycol monomethyl ether 2000, 200 mM ammonium sulfate, and 50 mM sodium cacodylate at pH 4.6.

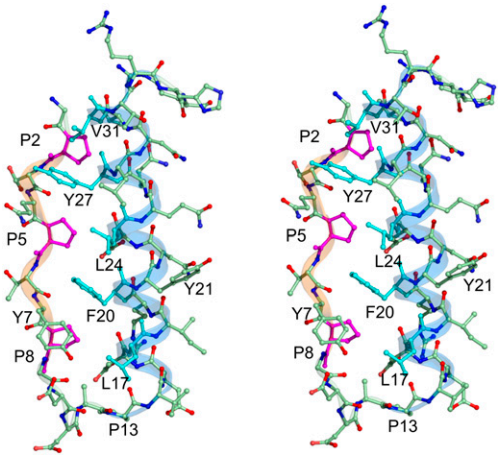

Fig. S3. Stereo representation of the crystal structure of the Avian pancreatic peptide (APP) resolved at 0.99 Å resolution (PDB 2BF9). Figure shows the short α-helix (highlighted in blue) and the short PPII helix (highlighted in orange). The color schemes adopted are similar to Fig. 3*B* in the main text to allow for comparisons. The residues colored in cyan namely, F20, Y26, L17, L24, and V31 in this structure, are homologous to the heptad repeats observed in A_3VP_1 crystal structure. Similarly, the PxxP motifs (colored in magenta) are homologous to the A_3VP_1 crystal structure.

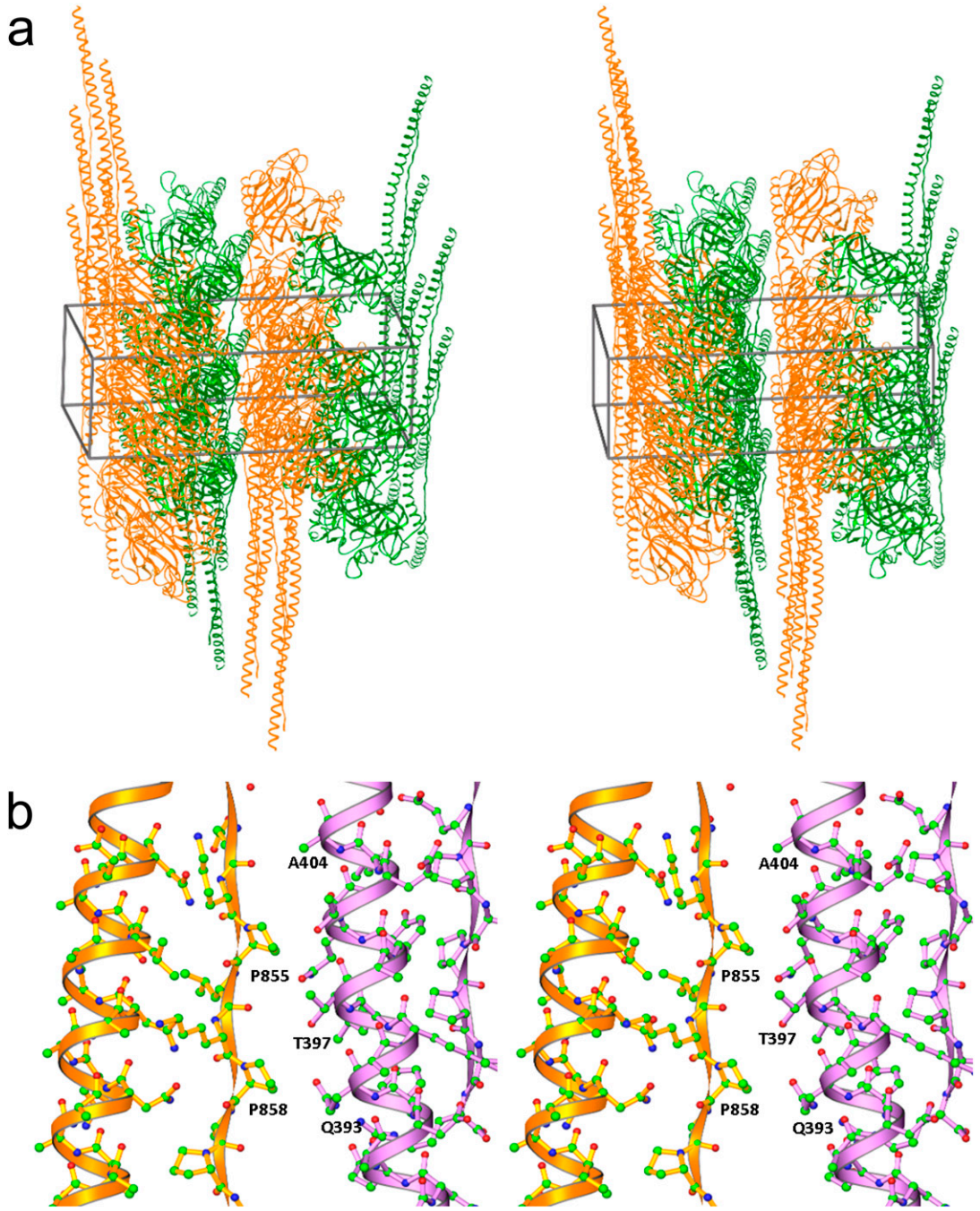

Fig. S4. Crystal packing diagram. (*A*) The unit cell for the $P2_12_12$ spacegroup crystals is delineated by the box. Molecules in the crystal pack with the V regions stacking upon each other, whereas the long helical A_3/P_1 units lie parallel and in a staggered alignment between the symmetry related molecules. (*B*) Two prolines point outward from the P_1's PPII helix and interact with a neighboring symmetry-related molecule A_3 helix.

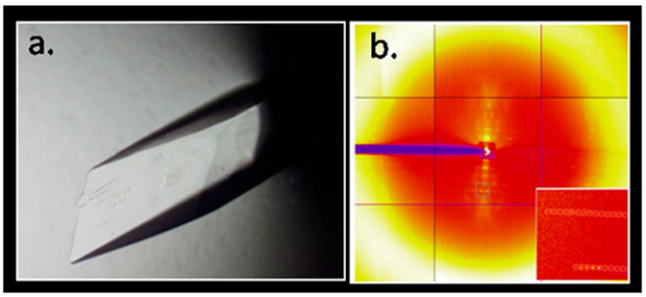

Fig. S5. A_{1-3} was purified as described in *SI Methods*. (*A*) Large crystals of the A_{1-3} were grown using the hanging drop vapor diffusion method. A droplet containing 2 μL of protein (7.2 mg/mL) was mixed with an equal volume of the reservoir solution that contained, 9% PEG 8000, 100 mM Tris pH 8.0, 50 mM L-Arginine, and 10 mM EDTA. (*B*) However, when exposed to X-rays, we observed asymmetric diffraction, and the crystal parameters were estimated to be: *a* = 69.96, *b* = 671.43, and *c* = 74.69 with β = 91.8. One of the cell dimensions was >600 Å, and most frames contained a fiber-like diffraction as seen in *B*. The asymmetric diffraction pattern, and the quick radiation damage to the crystals, made it impossible to collect a full data set, and they only diffracted at best to 6 Å. Solvent content estimations using the statistical Matthews coefficient, indicate that there could be at least 12 monomers in the asymmetric unit. From the ultracentrifugation studies, the estimated length of 51.56 Å for A_{1-3} >50 nm compares well with the estimated elongated cell dimensions of >60 nm observed in the A_{1-3} crystals, indicating that they could pack in an elongated fashion along the *b*-axis.

Table S1. Predicted secondary structures from CD (CONTIN/LL)

	A_{1-3}	V	VP_3	A_3VP_1	A_1VP_3	C-term
Total residues	248	356	516	517	812	516
α-Helical content % total (no. residues)	97.9 (243)	8.8 (31)	9.2 (47)	38.5 (199)	61.7 (501)	5.8 (30)
β-Sheet content % total (no. residues)	1.9 (5)	34.2 (122)	27.1 (140)	18.6 (96)	4.2 (34)	53.8 (278)
β-Turn content % total (no. residues)	0.2 (0)	22.1 (79)	19.0 (98)	18.7 (96)	16.1 (131)	17.6 (91)
Other % total (no. residues)	0.0 (0)	34.9 (124)	44.6 (230)	24.3 (126)	18.1 (147)	22.9 (118)

Table S2. PCR primers

Construct	Forward primer	Reverse primer	Template vector
MBP-A_1VP_1	GCAAATGGATCCCCTCGGCTCTCGCC	GATAACTCGAGCCTTGTCGGCGGTGTTGGCTCC	pCK1
A_1VP_3	CGAGTGGATCCCCAGTAAAACAGCTTATGAAGC	GCAATACTCGAGATGAACAGTTGGTACAGATGG	pRR2
VP_3	CGAGTGGATCCATAAAAATGAAGACGGAAACTTAACAGAACC	GCAATACTCGAGATGAACAGTTGGTACAGATGG	pRR2
A_{1-3}	ATATCCATGGGGGCTTATGAAGCTAAATTGG	ATATCTCGAGGTAATCAGCTTTAGCATCCGC	pSBR
V	CGAGTGGATCCATAAAAATGAAGACGGAAACTTAACAGAACC	GCAATACTCGAGGGTAACTTTAGGAACATTAACCGCACGG	pRR2
C-term	TATAAGGATCCATTTCCATTACTTTAAACTAGC	TTATACTCGAGTGAACTGTAAGTTACCCCATTGACAG	pDC20