Supporting Information

Heim et al. 10.1073/pnas.1413018111

SI Methods

Design, Expression, and Purification of Recombinant P1 and P1 Polypeptides. DNA encoding the N terminus and first alanine-rich repeat of P1 (NA1, amino acids 39–308) (Fig. 1*A*) was PCR-amplified from *spaP* of *S. mutans* NG8 using CGC**CATATG**-CACCATCATCATCATCATGATGAAACGACCACTACTAGT (with NdeI restriction site in bold and noncleavable 6-his tag underscored) and CGG**AAGCTT**TCAGTCAGTCATGCTTCG-TTTGCGGCATTAGC (HindIII restriction site in bold) as forward and reverse primers, respectively, cloned into pET-30(c) (EMD Millipore), and transformed into *E. coli* BL21 (DE3) cells (Life Technologies) as described previously (1). DNA encoding the third proline-rich repeat through the C terminus of P1 (P3C, amino acids 921–1486) (Fig. 1*A*) was generated similarly using CGC**CATATG**CACCATCATCATCATCATAATAAACCCAC-ACCGCCGACC (NdeI restriction site in bold and noncleavable 6-his tag underscored) and CGC**GGATCC**TCAGTCAGTCAT-GAACTGTAAGTTACCCCATT (BamHI restriction site in bold) as forward and reverse primers, respectively (1). NA1 and P3C were expressed separately and then copurified after culturing. Cultures were grown at 37 °C in terrific broth to an OD_{600} of 0.7–0.8, induced with 1 mM isopropyl β-D-1-thiogalacto-pyranoside, and incubated overnight at 22 °C. Owing to the relatively low expression of P3C, the cells expressing either NA1 or P3C were mixed together at a ratio of 1:8 (500 mL NA1 with 4 L P3C) and harvested the next morning.

The mixed NA1/P3C cell pellets were stored at −20 °C overnight. The next day, the cell pellets were suspended in TALON equilibration buffer (50 mM Tris·HCl and 300 mM NaCl, pH 7.4) supplemented with 20 μL of Dnase I (Thermo Scientific), 10 mg of lysozyme, and 10 mM $MgCl_2$, and then passed three times through an Avestin EmulsiFlex-C5 high-pressure homogenizer at a pressure of 15,000–20,000 psi. Cell lysates were centrifuged at 45,000 × *g* for 30 min, and the supernatant (soluble) fractions were filtered through a 0.22-μm syringe-driven filter (Millipore). Filtered samples were applied to 10 mL of TALON Cobalt Metal Affinity Resin (Clontech), and bound protein was eluted with 150 mM imidazole.

Because uncomplexed NA1 and P3C are inherently unstable, the purified NA1/P3C complex was incubated at 37 °C for 48 h to promote degradation of any residual NA1 or P3C fragments. Then, to remove any impurities remaining after the 37 °C degradation step, the purified complex was dialyzed overnight into TALON equilibration buffer to remove imidazole, again applied to the TALON resin, and eluted with 150 mM imidazole. Finally, the NA1/P3C complex was polished over a HiLoad 16/600 Superdex 200 prep-grade size-exclusion column (GE Healthcare). SDS/PAGE and Western blots of the purified NA1/P3C complex are shown in Fig. S1*A*.

DNA encoding full-length recombinant P1 lacking the signal sequence (rP1, amino acids 39–1566) was cloned into the pQE-30 vector and used to transform *Escherichia coli* M15-pREP4 cells as described previously (2, 3). In a previous study, the P1 mutant PC967, in which Cla1 restriction sites had been introduced immediately upstream and downstream of DNA encoding the P region, was found to have unusual properties, including diminished stability and altered antigenicity (4). The engineered Cla1 sites resulted in the introduction of additional isoleucine and aspartic acid residues at positions 826–827, as well as at positions 999–1000 (Fig. 1 *A* and *C*). Polypeptides containing one or both of these mutations were re-engineered to facilitate expression and purification of the variant polypeptides for further evaluation. DNA encoding the PC967 P1 derivatives, referred to herein as rP1-$Cla1_{Upstream}$ and rP1-$Cla1_{Up/Downstream}$, was cloned into pET-30(c) and used to transform *E. coli* BL21 (DE3) cells. rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ sample preparation was performed as described above for the NA1/P3C complex, except that all proteins were kept separate with no mixing, and a 37 °C degradation step was not applied.

After cell lysis, the soluble fractions were applied to 10 mL of TALON Cobalt Metal Affinity Resin, and bound proteins were eluted with 150 mM imidazole. Each eluted protein sample was subsequently polished using a HiLoad 16/600 200 PG size-exclusion column. SDS/PAGE and Western blots of purified rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ are shown in Fig. S1*B*.

SDS/PAGE and Western Blot Analysis of Purified Proteins. After purification of recombinant P1 polypeptides, 10 μg of each protein was loaded onto either a 10% polyacrylamide gel for NA1/P3C or a 7.5% polyacrylamide gel for rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ to analyze the purity of each protein by SDS/PAGE and Western blot analysis. After SDS/PAGE, each gel was either stained with Coomassie blue or electroblotted onto a Protran nitrocellulose membrane (Whatman). Replicate membranes blotted with the NA1/P3C complex were reacted with either mouse anti-P1 MAb 3–8D or 6–8C ascites fluid (5), which recognizes the alanine-rich region and the C terminus, respectively (2). The membrane blotted with rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ was reacted with mouse anti-P1 MAb 1–6F ascites fluid (5), which maps to the globular region at the apex of the protein (2) (Fig. S1*A*). All anti-P1 primary antibodies were diluted 1:500 in PBS-0.3% Tween 20 (PBST). After incubation with primary antibody, the membranes were washed with PBST and reacted with an HRP-labeled goat anti-mouse secondary antibody (MP Biomedicals) at a dilution of 1:1,000 in PBST and developed as described previously (6) (Fig. S1*B*).

Differential Scanning Calorimetry. Purified rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ were dialyzed overnight into PBS. The next day, each protein was diluted to a final concentration of 1 mg/mL in PBS dialysis buffer. Samples were analyzed using a MicroCal VP-DSC calorimeter, with 0.525 μL of each loaded into the sample cell and 0.525 μL of PBS dialysis buffer loaded into the reference cell. Samples were scanned from 30 °C to 85 °C at a scan rate of 1 °C/min. A buffer-only run was also performed with PBS in the sample cell, and background values were subtracted from each dataset. Data analysis and curve fitting were performed using Origin software (Microcal). Each experiment was performed in triplicate.

Circular Dichroism.

Initial circular dichroism measurements. All circular dichroism (CD) measurements were performed on an AVIV model 202 CD spectrometer using a quartz cuvette with a 0.2-cm path length. Purified rP1, rP1-$Cla1_{Upstream}$, and rP1-$Cla1_{Up/Downstream}$ were dialyzed extensively into 5 mM sodium phosphate buffer (pH 7.4), and far-UV CD spectra were measured at a concentration of 0.25 mg/mL from 260 to 185 nm over an average of four scans. All spectra were converted to units of mean residue ellipticity (MRE) and analyzed using the deconvolution algorithm CDSSTR from the CDPro software package (7–12) to estimate the secondary structure using the reference set SDP42. A buffer-only control scan

was performed, and background values were subtracted from the CD spectra.

Thermal denaturation. CD measurements were performed as described above for rP1, rP1-Cla1$_{Upstream}$, and rP1-Cla1$_{Up/Downstream}$, but with the temperature of the jacket increased progressively to 70 °C. The sample cuvette was allowed to equilibrate for ~1 min per 1 °C temperature increase before data collection. Three scans were measured for each temperature, and the dynode voltage remained below 800 V for all wavelengths over the entire temperature range. Protein refolding efficiency after thermal denaturation was determined by allowing each protein to refold overnight at 4 °C before data collection the next day. After refolding, no visible protein precipitate was observed, and four scans were averaged for each refolded sample. All spectra were converted to MRE and analyzed using CDSSTR. A buffer-only scan was performed separately at each temperature and subtracted from the CD spectra.

1. Oli MW, et al. (2012) Functional amyloid formation by *Streptococcus mutans*. *Microbiology* 158(Pt 12):2903–2916.
2. McArthur WP, et al. (2007) Characterization of epitopes recognized by anti-*Streptococcus mutans* P1 monoclonal antibodies. *FEMS Immunol Med Microbiol* 50(3):342–353.
3. 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(9):4274–4282.
4. Crowley PJ, et al. (2008) Requirements for surface expression and function of adhesin P1 from *Streptococcus mutans*. *Infect Immun* 76(6):2456–2468.
5. 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(11):2759–2767.
6. Brady LJ, Piacentini DA, Crowley PJ, Bleiweis AS (1991) Identification of monoclonal antibody-binding domains within antigen P1 of *Streptococcus mutans* and cross-reactivity with related surface antigens of oral streptococci. *Infect Immun* 59(12):4425–4435.
7. Yang JT, Wu CS, Martinez HM (1986) Calculation of protein conformation from circular dichroism. *Methods Enzymol* 130:208–269.
8. 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(2):252–260.
9. Sreerama N, Woody RW (2004) Computation and analysis of protein circular dichroism spectra. *Methods Enzymol* 383:318–351.
10. Sreerama N, Venyaminov SY, Woody RW (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. *Protein Sci* 8(2):370–380.
11. Compton LA, Johnson WC, Jr (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem* 155(1): 155–167.
12. Hennessey JP, Jr, Johnson WC, Jr (1981) Information content in the circular dichroism of proteins. *Biochemistry* 20(5):1085–1094.

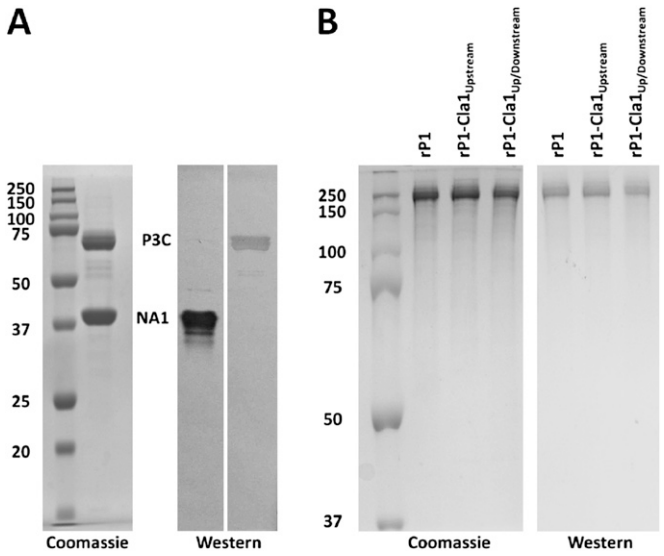


Fig. S1. Coomassie blue-stained SDS/PAGE gel and Western blots of purified proteins used in this study. (*A*) Coomassie blue-stained denaturing SDS/PAGE gel and Western blots of NA1/P3C probed with the anti-P1 mAbs 3–8D and 6–8C, which recognize the A region and C terminus, respectively (2). (*B*) Coomassie blue-stained gel and Western blot of the purified rP1, rP1-Cla1$_{Upstream}$, and rP1-Cla1$_{Up/Downstream}$ polypeptides probed with mAb 1–6F, which maps to the globular region at the apex of the protein (5).

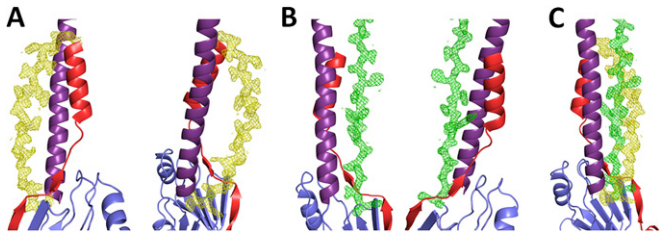
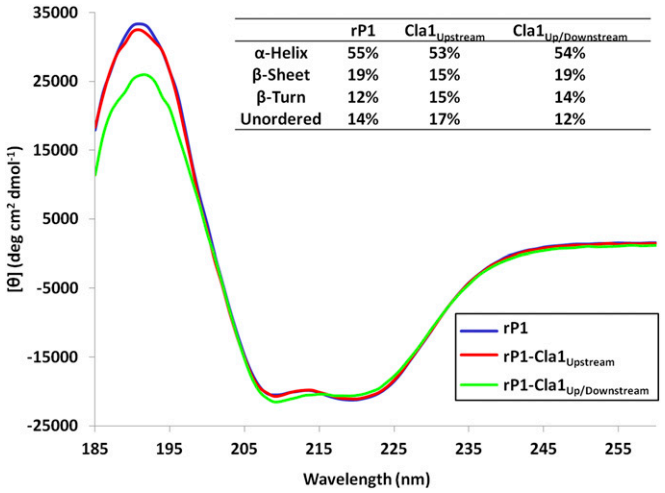


Fig. S2. An omit map contoured at 2σ displaying regions of unassigned electron density upon removal of the topological loop formed by the N terminus around the postproline-rich region. The electron density where the main chain of the N terminus was modeled into the loop is shown in yellow, and the electron density where the main chain of the postproline-rich region was modeled is shown in green. (*A*) Electron density of the N terminus clearly displays a single peptide chain forming a loop that leads into the base of the elongated α-helix of the alanine-rich region. Two views are shown, rotated 180°. (*B*) Electron density of the postproline-rich region clearly displays a single peptide chain coming up directly from the domain C1 of the C terminus. Two views are shown, rotated 180°. (*C*) Composite of the N-terminal and postproline-rich electron density, showing the N-terminal electron density wrapping around the postproline-rich electron density to form the topological loop represented within our crystal structure.



	rP1	$Cla1_{Upstream}$	$Cla1_{Up/Downstream}$
α-Helix	55%	53%	54%
β-Sheet	19%	15%	19%
β-Turn	12%	15%	14%
Unordered	14%	17%	12%

Fig. S3. CD measurements and predicted secondary structure content of rP1-$Cla1_{Upstream}$, rP1-$Cla1_{Up/Downstream}$, and rP1. Far-UV CD spectra of purified polypeptides measured from 260 to 185 nm. Units are in MRE. (*Inset*) The predicted secondary structure contents of the purified polypeptides were determined by the deconvolution algorithm CDSSTR (7–12). The protein reference set used was SDP42.

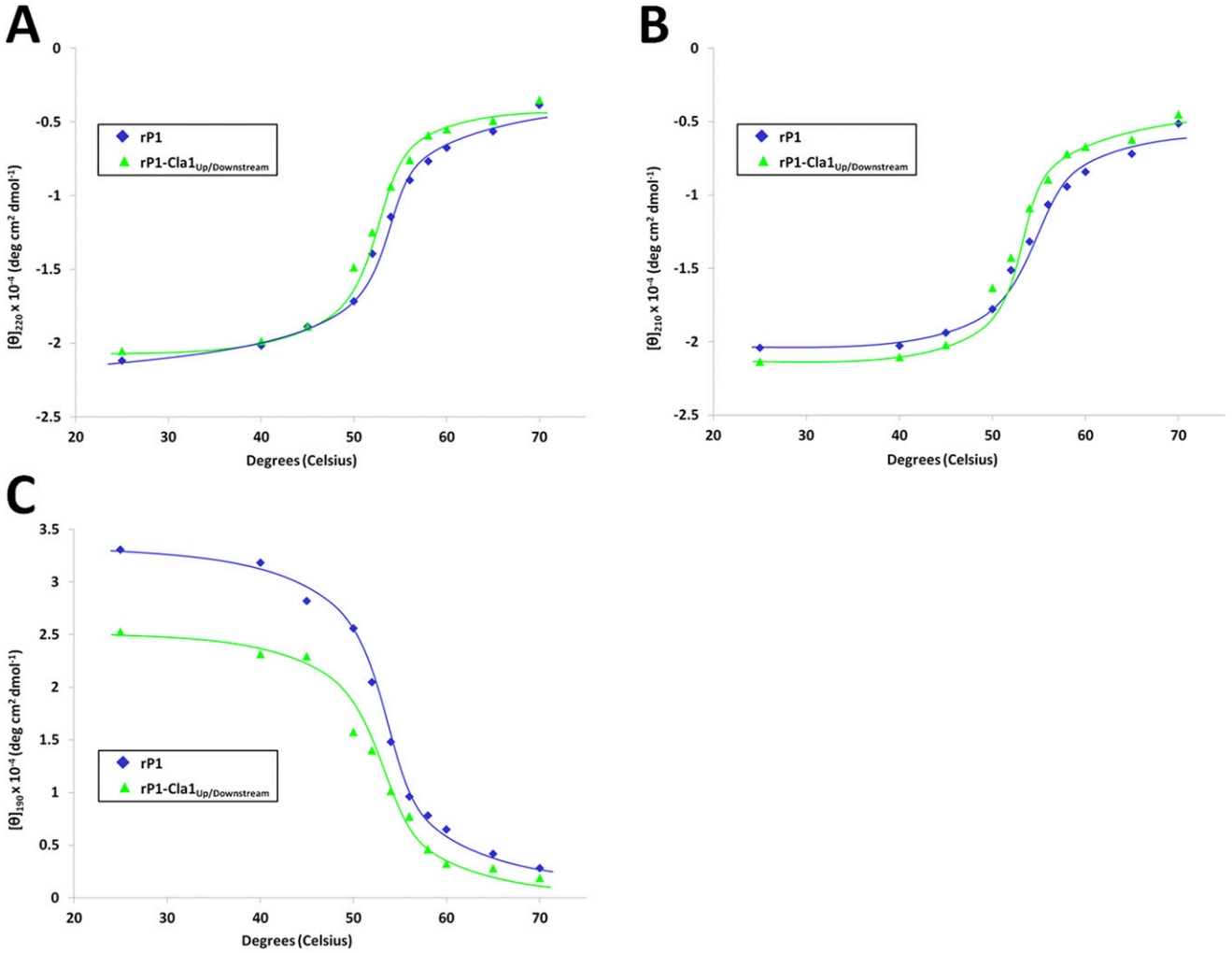


Fig. S4. Comparison of rP1 and rP1-Cla1$_{Up/Downstream}$ changes in mean residue ellipticity at fixed wavelengths—220 nm (*A*), 210 nm (*B*), and 190 nm (*C*)—measured over a temperature range of 25–70 °C.

Table S1. Summary of secondary structure of the N terminus

β-strands (amino acids)	α-helices
83–86 (anti-parallel to 1134–1131)	72–79
88–90 (anti-parallel to 1000–998)	95–106
110–113 (parallel to 990–993)	125–172
117–122/ (parallel to 996–1001)	

Corresponding residues forming hydrogen bonds between β-strands are shown in parentheses.

Table S2. Predicted secondary structure content of rP1-Cla1$_{Upstream}$, rP1-Cla1$_{Up/Downstream}$, and rP1 during thermal denaturation

Secondary structure content	25 °C	40 °C	45 °C	50 °C	52 °C	54 °C	56 °C	58 °C	60 °C	65 °C	70 °C
rP1											
α-helix content	55	52	52	47	43	34	25	19	16	9	5
β-sheet content	19	17	14	16	16	15	18	19	20	23	28
β-turn content	12	14	15	15	17	18	17	17	17	16	18
Unordered content	14	17	19	22	24	32	40	45	47	52	47
rP1-Cla1$_{Upstream}$											
α-helix content	53	53	52	48	42	35	23	19	16	8	5
β-sheet content	15	17	16	18	18	17	19	19	20	21	30
β-turn content	15	14	15	16	19	18	20	18	17	16	19
Unordered content	17	17	17	18	21	31	38	44	47	54	46
RP1-Cla1$_{Up/Downstream}$											
α-helix content	54	51	51	41	35	23	18	10	8	7	6
β-sheet content	19	19	18	17	17	19	20	25	22	25	29
β-turn content	14	15	16	16	18	19	18	17	16	18	19
Unordered content	12	15	15	26	30	39	44	48	53	50	46

Percentage secondary structure estimates were calculated using the deconvolution algorithm CDSSTR (7–12). The protein reference set was SDP42.

Table S3. Predicted secondary structure content of recombinant rP1-Cla1$_{Upstream}$, rP1-Cla1$_{Up/Downstream}$, and rP1 after thermal denaturation and subsequent refolding

Secondary structure content	25 °C	25 °C renatured
rP1		
α-helix content	55	53
β-sheet content	19	17
β-turn content	12	15
Unordered content	14	15
RP1-Cla1$_{Upstream}$		
α-helix content	53	53
β-sheet content	15	19
β-turn content	15	13
Unordered content	17	15
RP1-Cla1$_{Up/Downstream}$		
α-helix content	54	41
β-sheet content	19	17
β-turn content	14	17
Unordered content	12	25

Percentage secondary structure estimates were calculated using the deconvolution algorithm CDSSTR (7–12). The protein reference set was SDP42.