Domain Structure of the *Staphylococcus aureus* Collagen Adhesin†

R. L. Rich,*,‡ B. Demeler,§ K. Ashby,|| C. C. S. Deivanayagam,⊥ J. W. Petrich,|| J. M. Patti,‡ S. V. L. Narayana,⊥ and M. Höök‡

Center for Extracellular Matrix Biology, Institute of Biosciences and Technology, Texas A&M University, Houston, Texas 77030, Department of Chemistry, Iowa State University, Ames, Iowa 50011, Department of Biochemistry, University of Texas Health Science Center, San Antonio, Texas 78284, and Center for Macromolecular Crystallography, University of Alabama at Birmingham, Birmingham, Alabama 35294

Received July 22, 1998; Revised Manuscript Received September 8, 1998

ABSTRACT: Sequence analysis of surface proteins from Gram-positive bacteria indicates a composite organization consisting of unique and repeated segments. Thus, these proteins may contain discrete domains that could fold independently. In this paper, we have used a panel of biophysical methods, including gel permeation chromatography, analytical ultracentrifugation, circular dichroism, and fluorescence spectroscopy, to analyze the structural organization of the *Staphylococcus aureus* collagen adhesin, CNA. Our results indicate that the structure, function, and folding of the ligand-binding domain (A) are not affected by the presence or absence of the other major domain (B). In addition, little or no interaction is observed between the nearly identical repeat units within the B domain. We propose that CNA is indeed a mosaic protein in which the different domains previously indicated by sequence analysis operate independently.

Bacterial adherence to the host tissue is the first critical step in the pathogenic process of most infections. Organisms that cause infections in the extracellular space colonize the extracellular matrix. This adherence is mediated by a subfamily of bacterial surface proteins called MSCRAMMs[1] (*1*−*3*). During the past decade several MSCRAMMs on Gram-positive bacteria have been identified and partially characterized (for reviews, see references *4*−*6*). Sequence analyses of the MSCRAMMs revealed a common structural theme composed of unique and repeated sequences, and a multidomain organization of these proteins was therefore proposed. One such MSCRAMM, the collagen adhesin, CNA (Figure 1a), is present on some strains of *Staphylococcus aureus* and is composed of a N-terminal signal sequence followed by a 55 kDa segment of unique sequence called the A domain. This domain is followed by the B domain, which is composed of a 23 kDa subunit that is repeated 1−4 times in CNA from different *S. aureus* strains (*7*). The C-terminal portion of the protein consists of a cell wall attachment region, that contains the cell wall anchoring site LPKTG, a hydrophobic transmembrane region and a short cytoplasmic tail composed primarily of positively charged residues. It is believed that the protein is cleaved at the cell wall anchoring site and becomes covalently linked to the cell wall (*8*).

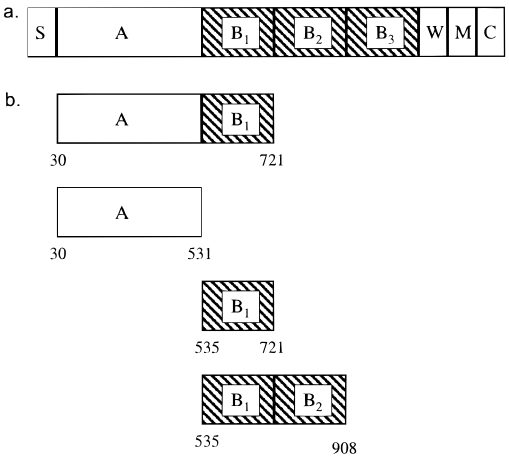

FIGURE 1: (a) Schematic of the *S. aureus* FDA 574 collagen-binding MSCRAMM, CNA. The putative signal peptide (S), nonrepetitive binding domain (A), domain of three subunits ($B_1B_2B_3$), cell wall domain (W), membrane spanning domain (M), and charged carboxyl-terminal domain (C) are indicated. (b) Recombinant CNA proteins used in this study that mimic portions of the collagen-binding MSCRAMM, with the inclusive residues indicated. The A domain subfragment (M19) for which the crystal structure has been solved spans residues 151−318 (*11*). MRGSHHHHHHGS is the amino acid sequence of the unstructured His_6-tag required for purification.

CNA mediates bacterial adherence to collagenous tissues such as cartilage, a process that is important in the pathogenesis of septic arthritis caused by *Staphylococci*. An isogenic mutant in which the *cna* gene has been inactivated is considerably less virulent than the wild-type CNA-expressing bacteria in a mouse model of septic arthritis (*9*). Furthermore, a *S. aureus* strain that normally lacks the *cna*

†This research was supported by grants from the Arthritis Foundation (awarded to R.L.R. and J.M.P.) and the National Institutes of Health (AR44415 awarded to M.H.)

* To whom correspondence should be addressed: 713-677-7557 (phone); 713-677-7576 (fax); rrich@ibt.tamu.edu (e-mail).

‡ Texas A&M University.

§ University of Texas Health Science Center.

|| Iowa State University.

⊥ University of Alabama at Birmingham.

[1] Abbreviations: MSCRAMMs, microbial surface components recognizing adhesive matrix molecules; CNA, *Staphylococcus aureus* collagen adhesin; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; ESI/MS, electrospray ionization mass spectrometry; UV/vis, ultraviolet/visible; CD, circular dichroism; SPR, surface plasmon resonance spectroscopy; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; BSA, bovine serum albumin; amino acids are represented by standard one-letter codes.

10.1021/bi981773r CCC: $15.00
Published on Web 10/16/1998

Table 1: Synthetic Oligonucleotides Used in Plasmid Constructions

construct	base pairs	amino acids	oligonucleotide[a]
A	238−1744	30−531[b]	*Bam*HI primer #5: 5′-GC**GGATCC**CGCACGAGATATTTCA-3′
			JP3 primer: 5′-CG**GTCGAC**TTATTCAGTATTAGTAACCAC-3′
AB_1	238−2314	30−721	*Bam*HI primer #5: 5′-GC**GGATCC**CGCACGAGATATTTCA-3′
			B_1 (3′) primer: 5′-CGC**GTCGAC**TTATGGCGTATATTTATTCGT−3′
B_1	1753−2314	535−721	B_1 (5′) primer: 5′-CGC**GGATCC**GAAACAACATCAATTAGT−3′
			B_1 (3′) primer: 5′-CGC**GTCGAC**TTATGGCGTATATTTATTCGT−3′
B_1B_2	1753−2875	535−908	B_1 (5′) primer: 5′-CGC**GGATCC**GAAACAACATCAATTAGT−3′
			B_1 (3′) primer: 5′-CGC**GTCGAC**TTATGGCGTATATTTATTCGT−3′

[a] Restriction endonuclease sites are in boldface type. [b] Not including N-terminal histidine tag.

gene becomes more virulent in this model when the *cna* gene is introduced (*9*). Recent studies suggest that the recombinant CNA A domain, when used as a vaccine, provides protection in a mouse model of staphylococcal-induced sepsis (*10*).

The collagen-binding function of CNA has been located to the A domain, and a recombinant protein containing a 19 kDa subfragment of the A domain has collagen-binding activity (*11*), although a recombinant version of the full-length A domain appears to bind collagen more effectively than the shorter subfragment (*12*). We recently solved the crystal structure of the 19 kDa collagen-binding subfragment (*13*). This polypeptide folds as a jellyroll composed of two β-sheets connected by two short α-helices. One of the β-sheets has a noticeable "trench" transversing it. Molecular modeling studies showed that this trench can accommodate a collagen triple helix, and analyses of single point mutations in the trench suggest this does in fact represent the collagen-binding site and several residues making contact along the extended collagen macromolecule. Such a binding scheme has been suggested recently for other collagen-binding proteins (*14, 15*).

The B domain of CNA in *S. aureus* strain FDA 574 is a triplet of repeat units. The first and third repeat units of the B domain are sequentially identical; the second varies by a single residue: residue N566, which is located in the central part of the B repeat unit in B_1, is replaced by D753 in B_2. The importance of the CNA B repeat units is unclear. Recent work by Hartford et al. showed that expression of a CNA construct in which the B domain repeat units were deleted resulted in presentation of a truncated, but functional, collagen-binding MSCRAMM on the bacterial surface (*16*). These B domain repeat units could affect the function or the stability of the A domain, they may serve as a "stalk", projecting the A domain away from the bacterial surface and positioning the A domain for binding to collagen, or they may possess binding sites for ligands other than collagen.

The structural organization of this MSCRAMM, the interaction between domains, and the role of the B domain in collagen binding (or possible alternative function in *S. aureus* pathogenicity) of the MSCRAMM have not been previously examined. In this report, we have used a panel of biophysical methods to characterize recombinant forms of the A domain, a B repeat unit, a heterodimeric AB construct, and a homodimeric BB construct (Figure 1b) with the intent to determine their structural and functional relationships.

EXPERIMENTAL PROCEDURES

Construction of Expression Plasmids. The expression plasmids pQE-A, pQE-AB_1, pQE-B_1, and pQE-B_1B_2 were constructed based on the vector pQE-30 (Qiagen Inc., Chatsworth, CA) using standard molecular biology protocols (*17, 18*). Recombinant proteins expressed from this vector contain an N-terminal tail of six histidine residues. DNAs encoding the various gene fragments were obtained by polymerase chain reaction using *S. aureus* FDA 574 genomic DNA (*19*) as a template and the oligonucleotides shown in Table 1 as primers. The amplified gene fragments, containing *Bam*HI and *Sal*I restriction sites at the 5′- and 3′-ends, respectively, were introduced into the oligonucleotides, treated with proteinase K, phenol/chloroform extracted, and ethanol-precipitated as described by Crowe et al. (*20*). Each of the fragments was then digested with the *Bam*HI and *Sal*I restriction endonucleases, purified by agarose gel electrophoresis (Geneclean kit, ISC BioExpress), and ligated to the vector pQE-30 (previously linearized by digestion with the same endonucleases). Ligation mixtures were transformed into *E. coli* strain JM101. Plasmids from isolated transformants were analyzed by restriction digestions and automated DNA sequencing analysis (Molecular Genetics Core Facility, UT−Houston Medical School) to confirm the expected open reading frame. The published sequence (*8, 11*) is identical to the now-determined sequence except for the following: Arg399→Ala, Ile665→Thr, and Ile852→Thr.

Bacterial Growth Conditions. Recombinant *E. coli* (JM101) were grown at 37 °C in Luria Broth (Life Technologies) containing 100 μg/mL ampicillin. Overnight cultures of stationary-phase bacteria were used to inoculate fresh media (1:1000), and the cells were allowed to grow until $OD_{600\ nm}$ = 0.6−0.8 (approximately 6 h). Protein expression was induced by addition of isopropylthio-β-D-galactoside to a final concentration of 0.2 mM, and the culture was incubated for an additional 3 h. Bacteria were collected by centrifugation and resuspended in a minimal volume of 5 mM imidazole, 100 mM NaCl, 4 mM Tris-HCl, pH 7.9, and frozen at −70 °C.

Isolation of Recombinant Proteins. Cell pellets were thawed, passed through a French press (11 000 lb/in^2) 3 times to lyse the cells, and centrifuged at 14 krpm for 20 min to remove cell debris and insoluble proteins. The supernatants were filtered through a 0.45 μm membrane, and recombinant proteins were purified using metal-chelating affinity chromatography. Bacterial lysates were applied to a HR10/10 (Pharmacia) column of nickel-charged iminodiacetic acid Sepharose (Sigma), and bound protein was eluted with a 200 mL linear gradient of 0−200 mM imidazole in 4 mM Tris-HCl, 100 mM NaCl, pH 7.9. In general, this purification step yielded proteins that were >95% pure as estimated by SDS−PAGE. The recombinant AB_1 preparation, however, required an additional purification step. Fractions containing

Table 2: Summary of Physical Parameters

species	molecular mass (kDa)[a] sequence	mass spec	sed vel	$\bar{\nu}$ (cm^3/g)	$s_{20,w}$ (S)[b]	$D_{20,w}$ (cm^2/s) $\times 10^7$ [b]	effective radius (Å)[c]	F/F_o
A	56586	56627	58189	0.7263	3.27	4.98	39.1 (31.6)	1.68
AB_1	78120	78130	77722	0.7248	4.28	4.88	42.8 (35.6)	1.56
B_1	22612	22587	24846	0.7168	2.30	7.93	26.3 (22.5)	1.42
B_1B_2	43807	43754	40582	0.7186	3.46	7.35	31.3 (28.8)	1.29

[a] Values represent molecular mass calculated from the published sequence, mass spectrometry, and sedimentation velocity measurements, respectively. [b] Sedimentation coefficients obtained from the finite element analysis and calculated using the molecular weight determined by sedimentation velocity. [c] Assuming a hydrated sphere. The data shown here are an average of values obtained by gel-permeation chromatography and sedimentation velocity measurements. Radii expected for globular proteins of these molecular masses are shown in parentheses and calculated from formulas outlined by Uversky (*46*).

recombinant AB_1 (as determined by SDS−PAGE) were pooled and dialyzed against 20 mM tris-HCl, pH 8.0, before being applied to a HR10/10 Q Sepharose column (Pharmacia). Bound protein was then eluted using a 100 mL gradient of 0−0.5 M NaCl, 20 mM Tris-HCl, pH 8.0. During all chromatographic procedures, the eluate was monitored by absorbance at 280 nm and collected fractions were examined for purity by SDS−PAGE. Fractions containing only the desired product were pooled and concentrated using an Amincon ultrafiltration system. Concentrated proteins were then dialyzed extensively against 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, fast-frozen, and kept at −70 °C until use.

Verification of Protein Identity. To verify the protein's identities, electrospray mass spectrometry measurements of each were performed at the University of Texas−Houston Health Science Center. Samples were dialyzed 3 × 1 L against 50 mM EDTA in water to remove Ni^{2+}, and then dialyzed 3 × 1 L against water to remove excess EDTA. All aqueous stock sample solutions were diluted with ESI/MS solvent, consisting of 100/100/10 (v/v/v) methanol/water/glacial acetic acid. Instrument setup and calibration were performed using a mixture of horse skeletal muscle apomyoglobin and BSA (2.5 pmol/μL each) in ESI/MS solvent. Analyses were performed using a Finnigan TSQ70 (upgraded to TSQ700) mass spectrometer fitted with a modified Vestec ESI source. Duplicate spectra were acquired for each sample. The instrument was calibrated before each run using a series of standards; deviation between theoretical and measured molecular masses of the standards was consistently less than 0.02%. Table 2 lists the molecular mass of each recombinant protein as determined from the amino acid sequence, measured by mass spectrometry, and calculated from sedimentation velocity parameters.

Gel Permeation Chromatography. Effective radii were measured at room temperature using gel-permeation chromatography. Protein samples were diluted in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, applied to a HR10/30 Superdex 200 column, and eluted with the same buffer at a flow rate of 0.5 mL/min. The column was calibrated using a standard globular protein kit (Pharmacia). Radii were then calculated using sample elution volumes and standard curves as described in the calibration kit.

Sedimentation Velocity Measurements. Prior to analysis, protein solutions were dialyzed 3 × 1 L against 10 mM Na_2HPO_4, pH 7.0, to remove NaCl that may damage the instrument's centerpieces. Experiments were performed using a Beckman XL-A analytical ultracentrifuge, equipped with an AN-60 Ti four-position rotor and absorption optics. All samples were centrifuged at 20 °C and 60 krpm in double-sector aluminum centerpieces. Sedimentation data were analyzed with the UltraScan software package[2] using the van Holde−Weischet (*21*), second moment (*22*), and finite element analysis (*23*, *24*) methods. The van Holde−Weischet method was used to confirm sample homogeneity and identify the possible concentration dependence of the sedimentation coefficient, *s*. The second moment method was used to ensure that the scans included in the analysis did not exhibit time-dependent variations such as aggregation, degradation, or pelleting. Finally, to determine the sedimentation and diffusion coefficients, the experimental data were fit directly to a finite element solution of the Lamm equation as described by Demeler and Saber (*25*). The partial specific volume, $\bar{\nu}$, of each protein was estimated from the amino acid sequence by the method of Sober (*26*) and was temperature-corrected. To model the overall geometry of the molecules, axial ratios were estimated for prolate and oblate ellipsoids from the experimentally determined sedimentation and diffusion coefficients as described in ref *27*.

Absorption Spectroscopy. Absorption measurements were taken at ambient temperature (23 ± 2 °C) using a Beckman DU-70 UV/vis spectrophotometer and a 1.0 cm path length cuvette. All spectra were corrected for background signal, and the reported absorbance maxima are ±1.0 nm. Since the proteins in this study all contain chromophoric amino acids, the concentration of each protein solution was determined spectrophotometrically using the Edelhoch method of UV-light absorption (*28*). Extinction coefficients were experimentally determined as described by Gill and von Hippel (*29*), employing Pace et al.'s values for the extinction coefficients of the individual residues (*30, 31*).

Steady-State Fluorescence Spectroscopy. Fluorescence excitation and emission spectra were collected at ambient temperature using an Aminco SLM 4800C fluorimeter equipped with a 450 W xenon lamp, employing excitation and emission band-passes of 2 nm. The excitation source was set at 295 nm to preferentially select tryptophyl emission. Protein samples, in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, were stirred throughout each scan to prevent photobleaching. All scans were background-subtracted, and all reported maxima were reproducible within ±1.0 nm. The relative fluorescence quantum yields, $\phi_F(x)$, were determined using the calculations outlined by Petrich and co-workers (*32*):

[2] A sedimentation data analysis software package developed by one of the authors (B. Demeler).

$$\frac{\phi_F(x)}{\phi_F(B_1)} = \frac{1 - 10^{-OD(B_1)}}{1 - 10^{-OD(x)}} \frac{\int_0^\infty I_{em}(x,\upsilon)\, d\upsilon}{\int_0^\infty I_{em}(B_1,\upsilon)\, d\upsilon} \quad (1)$$

where OD(x) is the optical density of recombinant protein x at the excitation wavelength and $I_{em}(x,\nu)$ is the emission intensity at the excitation wavelength. For these studies, fluorescence quantum yields were determined relative to that of CNA B_1. Chaotropic protein denaturation was induced by equilibrating the samples overnight with 6.0 M guanidine hydrochloride (ultrapure grade, USBiochem), 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, at ambient temperature. Emission spectra of guanadine hydrochloride in buffer were subtracted from the corresponding sample scans.

Time-Resolved Fluorescence Emission Anisotropy. Proteins were labeled with Texas Red Sulfonyl Chloride (Pierce Chemical) to measure rotational tumbling ratios. Each protein was diluted into 10 mM Na_2HPO_4, 100 mM NaCl, pH 9.0, to a final concentration of 0.2 mg/mL. The labeling was performed as previously described (*33, 34*), with the following modifications. One milligram of Texas Red ($\lambda_{abs}^{max} = 596$ nm in buffered aqueous solution) was added to the protein solution, which was then mixed gently and incubated at room temperature for 1 h. To remove excess dye, the protein solution was dialyzed exhaustively against 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0.

Fluorescence lifetimes, $K(t)$, and fluorescence anisotropy decays, $r(t)$, were obtained using the time-correlated, single-photon counting technique (*35, 36*). A time scale of approximately 15 ns was used to collect the fluorescence decays. All measurements were performed at 20.3 °C and with $\lambda_{ex} = 570$ nm and $\lambda_{em} > 610$ nm. The anisotropy decays of the proteins were fit to the function $r(t) = r_1(0) \exp(-t/\tau_{r1}) + r_2(0) \exp(-t/\tau_{r2})$, with the short time constant fixed at the value obtained for the probe alone. Calculations for the determination of the order parameter and hypothetical cone semiangle are described in ref *32*.

Circular Dichroism (CD). Far-UV circular dichroism data were collected using a Jasco J720 spectropolarimeter calibrated with *d*-10-camphorsulfonic acid, employing a band-pass of 1 nm and integrated for 4 s at 0.2 nm intervals. All sample concentrations were less than 15 μM in 1 mM Na_2HPO_4, 10 mM NaCl, pH 7.0. Spectra were recorded at ambient temperature in cylindrical 0.5 mm path length cuvettes.

Twenty scans were averaged for each spectrum, the contribution from buffer was subtracted, and quantitation of secondary structural components was performed using software provided by N. J. Greenfield (UMDNJ−Robert Wood Johnson Medical School, Piscataway, NJ) and D. Greenwood (Softwood Co., Brookfield, CT). We initially used four deconvolution programs (SELCON, VARSLC1, CONTIN, and CD estima) that were derived from databases of known protein structures (*37*−*40*). The reliability of these programs in predicting the secondary structure of CNA domains was assessed by analyzing the spectra obtained for M19, a truncated form (19 kDa, residues 151−318) of the A domain that encompasses the collagen-binding region. The structure of M19 has been solved using X-ray crystallography and consists of 8% α-helix, 53% β-sheet, and 39% coil (*13*). The CONTIN analysis consistently yielded exaggerated β-sheet content, while the CD estima program dramatically underreported the percentage of β-sheet structure. Averaging data from all four programs yielded the same results as comparison of SELCON and VARSLC1 only, but with significant error (up to 50%). As a result, we have chosen to report an average of results obtained from the SELCON and VARSLC1 analyses only. For M19, this yielded secondary structure percentages that are comparable to the data obtained from the crystallographic analysis: 12 ± 4% α-helix, 49 ± 4% β-sheet, and 39 ± 13% other components.

Thermal Denaturation. Protein unfolding was induced by heat and monitored by far-UV CD. The thermal denaturation (ramp rate of 20 °C/h) of each protein in 1 mM Na_2HPO_4, 10 mM NaCl, pH 7.0, was followed at 208 nm by far-UV CD using the parameters described above. The reversibility of the thermal denaturation was determined immediately after the unfolding analyses, and the samples were cooled at a ramp rate of 20 °C/h.

Surface Plasmon Resonance Spectroscopy (SPR). Analyses were performed using the BIAcore system. Bovine type I collagen (Collagen Corp., Fremont, CA) was immobilized on a CM5 sensor chip as described previously (*11*). Recombinant CNA proteins in 150 mM NaCl, 50 mM HEPES, 0.05% P-20 surfactant, pH 7.4, were flowed over three flow cells containing different amounts of immobilized collagen. All data were corrected for the response obtained using a flow cell that was not coated with collagen. Analytical conditions and data analysis were as described previously (*12, 41*). No mass transport effects were observed in these measurements.

ELISA Assays. Microtiter plates were coated with 10 μg of type I collagen per well in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, overnight at 4 °C. Collagen-coated wells were then washed with the same buffer and incubated for 1 h in the presence of 5% BSA in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0. Wells were again washed with buffer, and varying concentrations of recombinant CNA protein, in buffer containing 0.1% BSA, were added. After 3 h, unbound protein was removed by washing the wells with the same buffer. To detect bound protein, the wells were incubated with a monoclonal antibody that recognizes the histidine tag of the recombinant proteins (6xHis Monoclonal Antibody, Clontech), followed by alkaline phosphatase-conjugated goat anti-mouse IgG antibody. Relative binding was measured by monitoring the sample absorbance at 405 nm following the addition of *p*-nitrophenyl phosphate in 1 M diethanolamine, 0.5 mM $MgCl_2$, pH 9.0.

RESULTS

Expression and Purification of Recombinant Proteins Mimicking Regions of CNA. A model of the collagen-binding MSCRAMM from *S. aureus* FDA 574 has been described previously, in which a multidomain nature of this protein was proposed (*8*). We have now analyzed recombinant proteins containing individual (A and B_1), homodimeric (B_1B_2), and heterodimeric (AB_1) domains of the MSCRAMM with the goal to evaluate this model. SDS−PAGE, followed by staining with Coomassie Brilliant Blue, revealed that each preparation of recombinant protein migrated as a single band, but at unexpectedly high molecular masses due to their acidic natures (Figure 2). No degradation

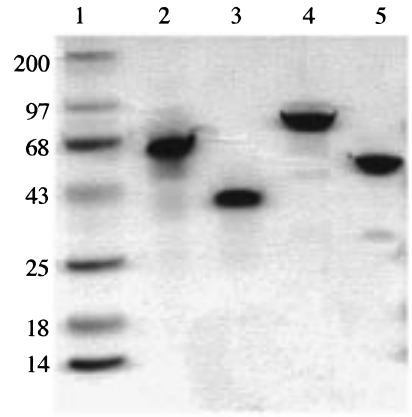

FIGURE 2: SDS−PAGE analysis of recombinant His_6-CNA proteins depicted in Figure 1b. Affinity-purified fusion proteins were fractionated by electrophoresis on a 12% polyacrylamide gel, and the proteins were stained with Coomassie Brilliant Blue. Lane 1, molecular mass standards; lane 2, A; lane 3, B_1; lane 4, AB_1; lane 5, B_1B_2.

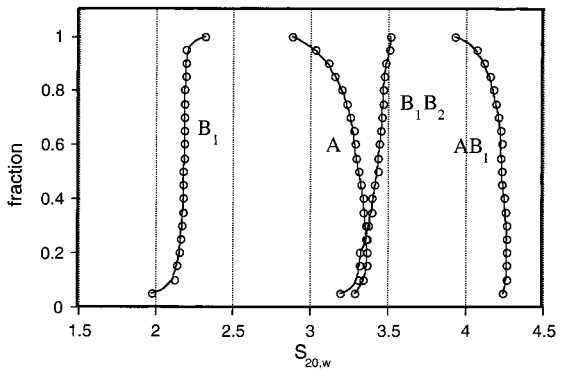

FIGURE 3: Integral distribution plots, $G(s)$, for the recombinant CNA protein as obtained from the van Holde−Weischet method. The data shown refer to diffusion, temperature, and buffer-corrected s-value distributions for 100% of the boundary analyzed. A fraction of 0.5 refers to the midpoint in the sedimentation velocity boundary. The negative slope of the distribution of A and AB_1 indicates slight solution nonideality due to concentration dependence of the sedimentation coefficient, s. All distributions indicate sample homogeneity.

or multimerization was observed for any of these recombinant proteins over a period of several weeks when the proteins were stored at −70 °C. However, attempts to express the full-length CNA protein ($AB_1B_2B_3$) from a pQE-based vector yielded no detectable amounts of soluble intact protein.

Structure of the CNA A Domain. Analysis of the recombinant CNA A domain by gel-permeation chromatography showed that the protein eluted significantly earlier than what was expected for a globular protein with a molecular mass of 57 kDa as compared with the elution positions of globular protein standards. From these measurements, we calculated that the effective radius (the protein's radius assuming it folds as a fully hydrated sphere) of the A domain was 39.1 Å, compared with 31.6 Å for a globular protein of similar mass (Table 2). Sedimentation velocity measurements using a Beckman XL-A analytical ultracentrifuge did not indicate aggregation of the A domain over a wide range of concentrations and yielded a F/F_o ratio of 1.68 (Figure 3, Table 2). If we assume a prolate ellipsoid model for this protein, the calculated axial ratio would be 12.8. Hence, the recombinant CNA A domain is significantly nonglobular.

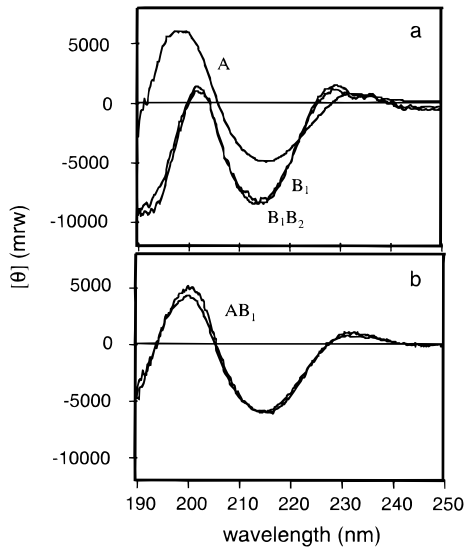

FIGURE 4: Circular dichroism of recombinant CNA proteins. (a) Far-UV spectra of A, B_1, and B_1B_2 illustrating the unique secondary structure of the A domain and the similarity between the monomeric and dimeric B repeat units. (b) Far-UV spectra of AB_1, measured and calculated. The mathematical construction of the calculated spectrum is described in the text. Secondary structural compositions are listed in Table 3. Mean residue weight molar ellipticity reported in deg·cm^2/dmol.

Table 3: Summary of Secondary Structural Components[a]

species	α-helix	β-sheet	other
A	0.09 ± 0.04	0.49 ± 0.02	0.42 ± 0.09
AB_1	0.08 ± 0.07	0.50 ± 0.02	0.42 ± 0.10
B_1	0.08 ± 0.06	0.41 ± 0.03	0.52 ± 0.11
B_1B_2	0.07 ± 0.05	0.44 ± 0.02	0.49 ± 0.10

[a] Values reported here are an average of results obtains using SELCON and VARSLC1 deconvolution programs. Refer to Experimental Procedures for further details.

The far-UV CD spectrum of this protein contains a minimum at 215 nm and relative maxima at 198 and 232 nm (Figure 4a). Deconvolution of this spectrum revealed that the full-length A domain recombinant is a primarily β-sheet protein that contains a small percentage α-helix (Table 3). This composition of secondary structure elements is similar to that for M19, the CNA A domain truncate that encompasses the collagen-binding motif (*11*), for which the crystal structure was recently solved (*13*). Thermal denaturation, as followed by intensity of the far-UV CD spectrum, revealed that the A domain is quite thermostable and unfolds reversibly over a range of 30 °C (60−90 °C), undergoing a multiphasic thermal denaturation with a highly populated intermediate state apparent at approximately 74 °C (Figure 5).

The recombinant A domain protein exhibits an absorbance maximum of 278 nm and has an experimentally determined extinction coefficient of 58 740 M^{-1} cm^{-1}. The fluorescence quantum yield of A is 15% of that for B_1, and the five tryptophans of the A domain reside in an overall nonpolar environment ($\lambda_{em}^{max} = 330$ nm when $\lambda_{ex} = 295$ nm) as shown by steady-state emission spectroscopy. The A domain's tryptophyl emission spectrum red-shifts upon denaturation in 6 M guanidine hydrochloride ($\lambda_{em}^{max} = 352$ nm) as all tryptophans become fully solvated.

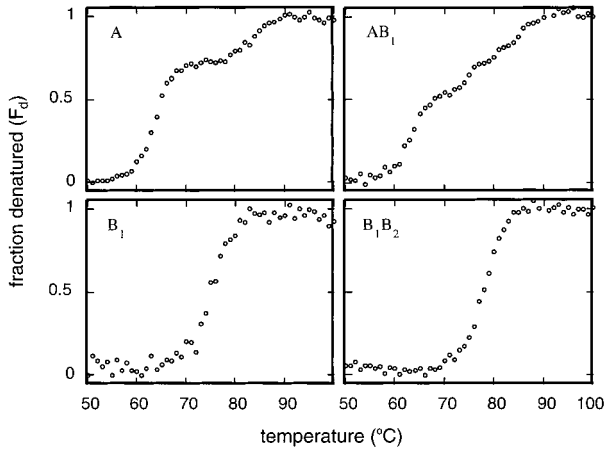

FIGURE 5: Thermal denaturation profiles of recombinant CNA proteins as monitored by far-UV CD.The midpoints of thermal unfolding (T_m) are 65 ± 2 °C and 86 ± 5 °C for A, 70 ± 2 °C for AB_1, 74 ± 1 °C for B_1, and 77 ± 1 °C for B_1B_2.

Structure of a CNA B Domain Repeat Unit. We propose that the structural data obtained for the B_1 subunit are valid for any single repeat unit of the B domain since the amino acid sequences of the repeat units are so similar (>99% of the residues are identical). Gel-permeation chromatography measurements of B_1 revealed that this protein eluted earlier than would be expected for a protein of this molecular mass, yielding an effective radius of 26.3 Å (versus 22.5 Å for a globular protein of the same mass, Table 2). Sedimentation velocity analyses demonstrated that a single repeat unit is monomeric in solution (Figure 3), and attempts to induce dimerization at high protein concentrations (>200 μM) in order to mimic the B_1B_2 recombinant were unsuccessful. In addition, the sedimentation velocity measurements gave a F/F_o ratio of 1.42. When modeled as a prolate ellipsoid, this would represent a shape with an axial ratio of 7.7. Together, the gel-permeation chromatography and sedimentation velocity results indicate that a CNA B domain repeat unit is nonglobular.

The far-UV CD spectrum of B_1 contains a relative minimum at 214 nm, maxima at 203 and 228 nm, and significant negative molar ellipticity at 190 nm (Figure 4a). These characteristics are typical of a structure rich in β-sheet. Deconvolution of this spectrum yielded secondary structural content that resembles that of the A domain (Table 3). However, the arrangement of the secondary structural motifs in the two domains is clearly not the same since the far-UV CD spectra of the two proteins are quite different. Heat denaturation of the recombinant single B domain repeat unit followed by changes in the far-UV CD spectrum demonstrated that this single repeat unit is particularly stable with respect to heat, denaturing reversibly in the range 68−82 °C via an apparently single, sharp transition (Figure 5).

The absorbance maximum of the recombinant B_1 repeat unit is 278 nm, and the extinction coefficient is 37 070 M^{-1} cm^{-1}. The five tryptophans of a B domain repeat unit reside in an environment that is much more polar than that of the A domain tryptophans. The maximum of the emission spectrum for B_1 occurs at 340 nm (λ_{ex} = 295 nm), 12 nm red-shifted from that for the A domain. In 6 M guanidine hydrochloride, all hydrophobicity of the tryptophans' local environments is removed and the tryptophyl emission maximum shifts to 352 nm.

No Structural Rearrangements Occur upon B Repeat Unit Dimerization. The recombinant B_1B_2 protein is a dimer of two B repeat units covalently linked at the C-terminus of the first and the N-terminus of the second. Since the sequence of the second repeat unit of the CNA B domain mimics exactly that of the first except for one residue (and the third repeat unit is identical to the first), we suggest that this dimer, B_1B_2, is a model for any two covalently joined B domain repeat units. The far-UV CD spectrum of B_1B_2 overlays that obtained for B_1 (Figure 4a), implying that the secondary structures of the B repeat units do not change dramatically when they are covalently linked in series. Deconvolution of the far-UV CD spectrum produced the same composition of structural components as for B_1, as was expected (Table 3). In addition, the secondary structural elements in B_1B_2 denature upon heating apparently as they do in B_1, for the completely reversible thermal denaturation of B_1B_2 occurs at 70−82 °C and is also a simple, two-state transition (Figure 5).

Gel-permeation chromatography measurements indicate that the effective radius of B_1B_2 is somewhat greater than that expected for a 44 kDa globular protein (31.3 vs 28.8 Å, Table 2). Sedimentation velocity experiments determined that neither degradation nor multimer formation of B_1B_2 occurs at high protein concentrations and the protein has a F/F_o ratio of 1.29 (Figure 3, Table 2). B_1B_2, therefore, also is nonglobular, yet not to the extent that B_1 is.

The absorption spectrum of the recombinant B_1B_2 heterodimer has a maximum of 278 nm, and the protein's extinction coefficient is 74 450 M^{-1} cm^{-1}. The 10 tryptophans of B_1B_2 are in an environment comparable with those of B_1 (λ_{em}^{max} = 341 nm, λ_{ex} = 295 nm). B_1B_2's fluorescence quantum yield is twice that of B_1. In fact, a theoretical construction of the B_1B_2 emission spectrum shown in Figure 6a closely duplicated the measured spectrum of B_1B_2 (the theoretical construction of the emission spectrum from eq 3 is discussed in detail below for AB_1). The emission maximum of B_1B_2 in 6 M guanidine hydrochloride shifts to 352 nm as all tryptophans become completely solvated upon protein denaturation.

We would expect that if no significant structural reorganization occurs among the B repeat units upon joining to form the dimer, B_1B_2 would have a significantly greater diameter than B_1. One method to compare relative protein sizes employs time-resolved fluorescence anisotropy. This technique follows the rotation of molecules in solution. Since large proteins tumble slower than small ones, we would expect the anisotropy decay (and, therefore, the rotational rate) of B_1B_2 to be much longer lived than that of B_1 if the individual repeats fold independently and are simply joined by a polypeptide tether sequence. For this analysis, B_1 and B_1B_2 were labeled with the visible light fluorescent probe Texas Red, and emission greater than 610 nm was collected.

The fluorescence lifetime of the Texas Red probe fit well to a single exponential of 4.2 ns. The fluorescence lifetime of this label did not change dramatically upon attachment to the B_1 (or B_1B_2) recombinant protein. The fluorescence lifetime decays of B_1 and B_1B_2 both fit to an exponential of the form: $K(t) = A_1 \exp(-t/\tau_{f1}) + A_2 \exp(-t/\tau_{f2}) = 0.80 \exp(-t/4.5\text{ ns}) + 0.20 \exp(-t/1.0\text{ ns})$, indicating that the local environment of the probe was similar in these two proteins.

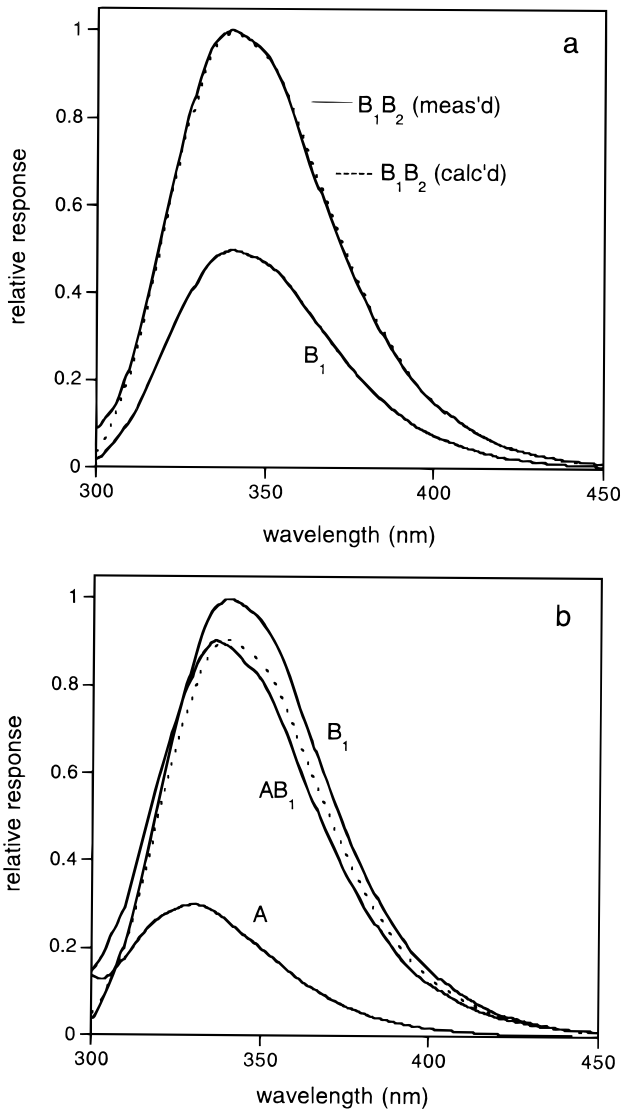

FIGURE 6: Comparison of the fluorescence emission spectra of recombinant CNA proteins. (a) B_1 and B_1B_2. The dashed line is the construction of a B_1B_2 spectrum using the B_1 spectrum, where B_1B_2(calcd spectrum) = 2 × B_1(measd spectrum). (b) A, B_1, and AB_1. The dashed line is the construction of the AB_1 spectrum using eq 3. All sample concentrations were 1.0 μM, and λ_{ex} = 295 nm.

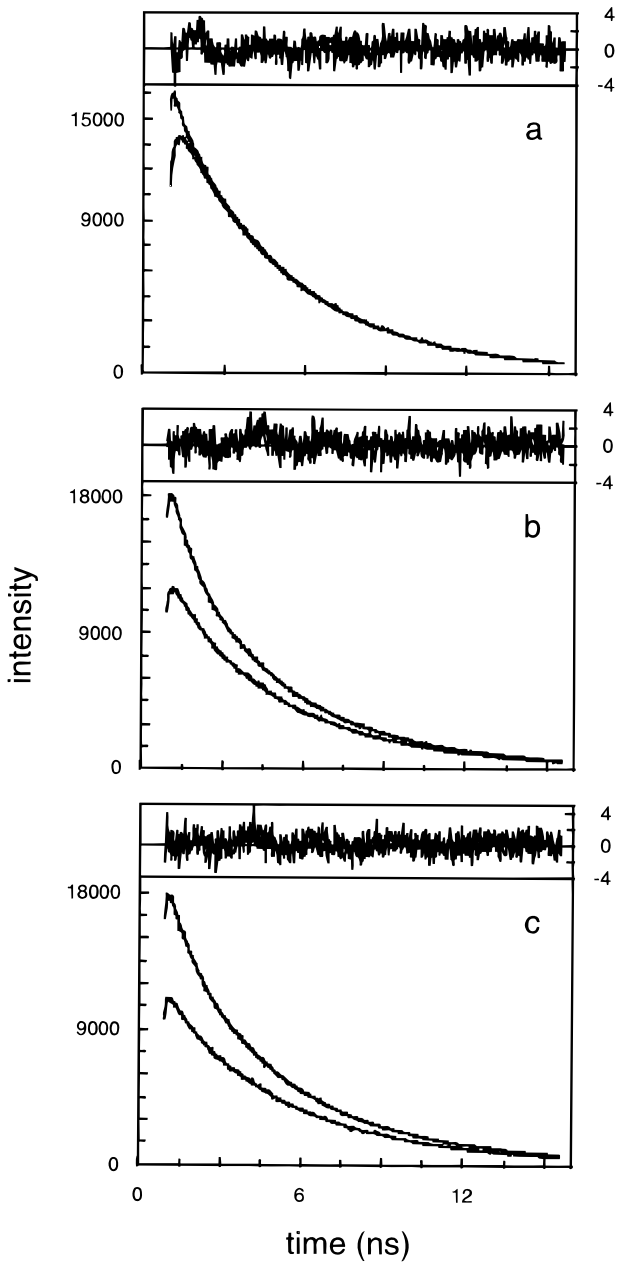

FIGURE 7: Parallel and perpendicular fluorescence intensity profiles of the recombinant CNA proteins from which the anisotropy decay is calculated (λ_{ex} = 570 nm, λ_{em} > 610 nm, 20.3 °C). (a) The fluorescent probe Texas Red alone in solution; (b) the labeled single B domain repeat unit, B_1; (c) the labeled B domain repeat unit dimer, B_1B_2. The residuals corresponding to emission polarized parallel and perpendicular to the excitation source are overlaid above each panel. All time-resolved measurements have an error of approximately 10%.

The anisotropy decay of the Texas Red dye alone in solution fit well to a single exponential of 240 ps ($\chi^2 = 1.38$) (Figure 7a). If the data were fit to two exponentials, a long component [τ_{r2} = 15.9 ns; $r_1(0) = 0.13$, $r_2(0) = 0.01$] was observed. This results in a hypothetical cone semiangle, θ_O (within which the transition dipole moment can diffuse), of approximately 90°, as was expected for the probe free in solution.

The fluorescence anisotropy decay of B_1 tagged with Texas Red fit well to two exponentially decaying components: $r(t) = r_1(0)\exp(-t/\tau_{r1}) + r_2(0)\exp(-t/\tau_{r2}) = 0.09\exp(-t/0.24\text{ ns}) + 0.12\exp(-t/8.8\text{ ns})$, $\chi^2 = 1.25$ (Figure 7b). This indicates that we were probing the rapid librational motion of the probe with respect to the recombinant protein as well as the overall tumbling motion of the protein macromolecule itself. The rotational anisotropy data for B_1B_2 were also resolved into two components: $r(t) = 0.06\exp(-t/0.24\text{ ns}) + 0.15\exp(-t/11.9\text{ ns})$, $\chi^2 = 1.26$ (Figure 7c). The second component of the anisotropy decay (τ_{r2}) of B_1B_2 is substantially longer-lived than that of B_1, as would be expected for a molecule of greater length. The hypothetical cone semiangle was determined to be approximately 34° for Texas Red attached to B_1 and 27° for B_1B_2. The fluorescent probe's significant cone semiangle of rotation was characteristic of a loosely tethered moiety, as would be expected for the dye molecule attached at the N-terminus of the unstructured histidine tag of B_1. The cone semiangle of rotation of the probe varied little between B_1 and B_1B_2, implying that no dramatic conformational change occurred at the N-terminal protein face upon dimerization.

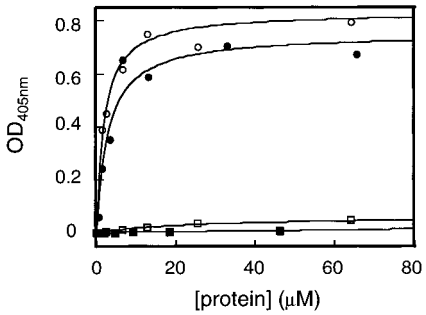

FIGURE 8: Relative binding of immobilized type I collagen (100 μg/mL) by A (○), AB_1 (●), B_1 (□), and B_1B_2 (■) as a function of recombinant CNA protein concentration. BSA bound type I collagen at levels comparable to those of recombinant CNA B_1B_2. Half-maximal binding of recombinant CNA A and AB_1 was calculated to be 1.9 ± 0.4 and 2.9 ± 0.9 μM, respectively.

No Structural Rearrangements Occur upon Forming the AB_1 Heterodimer. The heterodimeric construct AB_1 consists of a B_1 repeat unit linked via its N-terminus to the C-terminus of the A domain. The far-UV CD spectrum of AB_1 is indicative of a protein having significant β-sheet character, with an intensity minimum at 214 nm and relative maxima at 200 and 229 nm. The composition of secondary structural components of AB_1 is similar to that of the other constructs (Table 3), yet its profile (Figure 4b) does not superimpose that of the recombinant A or B domain shown in Figure 4a. However, if the percentages and arrangements of α-helix, β-sheet, and other components of A and B_1 do not undergo any change upon forming AB_1, a calculated far-UV CD spectrum may be constructed from the weighted sum of A and B_1's spectra that duplicates the observed AB_1 spectrum in Figure 4b:

$$[\theta]_{AB_1}(\text{calc'd}) = \frac{(residues)_A([\theta])_A + (residues)_{B_1}([\theta])_{B_1}}{(residues)_{AB_1}} \quad (2)$$

where *residues* is the number of amino acids in each protein. This is indeed the case: the theoretical construction and measured curve are nearly superimposable (Figure 4b). Hence, the organization of secondary structure components in A and B_1 is the same in the individual recombinants and the heterodimer.

The effective radius (42.8 Å) of AB_1 is significantly greater than expected for a globular 78 kDa protein. This protein is monomeric by sedimentation velocity analysis (Figure 3), and its F/F_o ratio of 1.56 is intermediate between those of A and B_1, indicating a shape less globular than B_1, but more so than A alone.

The thermal unfolding of AB_1 is complex: no obvious intermediate state is observed, and the transition from native to denatured protein is diffuse, with a T_m of 70 °C and a dynamic range of more than 25 °C (Figure 5). This would indicate a denaturation via an unresolved intermediate state. Unlike the unfolding of the other recombinants, the denaturation of AB_1 is not completely reversible: only approximately 70% of the original AB_1 conformation is regained upon cooling.

The absorbance maximum of the recombinant AB_1 heterodimer is 278 nm, and its extinction coefficient is 92 750 M^{-1} cm^{-1}. From the steady-state emission spectrum of AB_1, we conclude that its 10 tryptophans reside in an overall environment of intermediate polarity (λ_{em}^{max} = 337 nm, λ_{ex} = 295 nm) in relation to the other 3 recombinant proteins in this study. The fluorescence quantum yield of AB_1 is 0.85 with respect to B_1. As is expected, the fluorescence emission spectrum of AB_1 in 6 M guanadine hydrochloride indicates all 10 tryptophans become fully solvated upon denaturation.

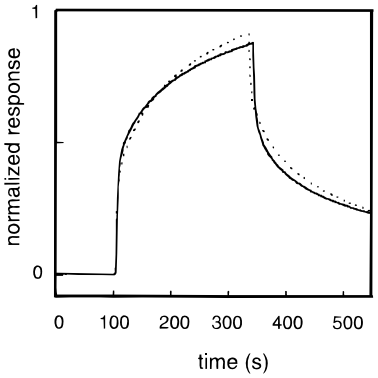

FIGURE 9: Representative profiles of the normalized SPR responses for the binding of 1.0 μM recombinant CNA A (—) and AB_1 (···) to immobilized type I collagen. In the analyses shown here, the 4-min injection (begun at 100 s) employed a flow rate of 5 μL/min.

A theoretical construction similar to that described above for the secondary structure of AB_1 (eq 2) may be used to determine the change, if any, in tertiary structure upon forming the mixed-domain protein. A theoretical fluorescence emission spectrum of (A + B_1) may be constructed based on the relative quantum yields (A has a fluorescence quantum yield of 0.15 with respect to B_1) and spectra shape and positions of the individual A and B_1 units:

$$I_{AB_1}(\text{calc'd}) = \frac{(\phi_F^{rel})_A(I_A) + (\phi_F^{rel})_{B_1}(I_{B_1})}{(\phi_F^{rel})_A + (\phi_F^{rel})_{B_1}} \quad (3)$$

where I_{AB1}(calc'd) is the calculated emission spectrum of AB_1, $(\phi_F^{rel})_A$ is the quantum yield of A relative to B_1, and I_A is the emission intensity of the A domain protein. If the tryptophan environments of both A and B_1 are not altered upon complexation (for example, surface-exposed tryptophans of the A and B_1 proteins becoming buried in the A−B_1 interface), this calculated spectrum should duplicate the observed AB_1 spectrum in Figure 6b. The observed and calculated spectra are nearly identical, providing additional evidence that (1) the local environments about the tryptophans in the recombinant heterodimer are no different than their environments in the individual domain recombinants and (2) no tryptophans reside at the A−B_1 interface.

Collagen Binding Is Not Influenced by the B Domain. An ELISA demonstrating the four CNA recombinant proteins binding to immobilized type I collagen is shown in Figure 8. Neither B_1 nor B_1B_2 exhibited collagen-binding activity. In this assay, the affinities of A and AB_1 for collagen were comparable. Figure 9 illustrates the normalized SPR responses of A and AB_1 flowed over a BIAcore sensor chip coated with type I collagen (the responses of B_1 and B_1B_2 were undetectable). It is apparent from these SPR profiles

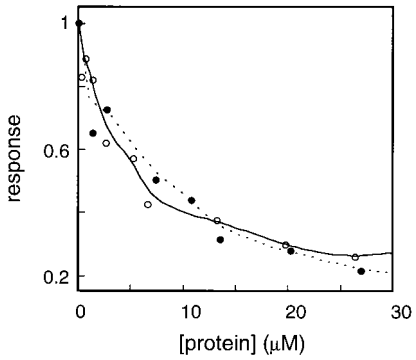

FIGURE 10: Inhibition of recombinant CNA AB_1 to immobilized type I collagen by biotinylated recombinant CNA A. The biotinylated and unbiotinylated species were allowed to bind simultaneously to the immobilized collagen. The dark circles represent self-inhibition (i.e., inhibition of A binding to immobilized type I collagen by biotinylated A), measured as a control. The data points represent average values of duplicate assays.

that addition of the B_1 repeat unit to the A domain does not significantly affect the association and dissociation rates of the protein/collagen interaction. In addition, the CNA recombinant A and AB_1 proteins compete for sites within the collagen macromolecule. Figure 10 demonstrates the inhibition of AB_1 by biotinylated A, in which the collagen binding of AB_1 was completely blocked by A. The reverse experiment yielded similar results: the inhibition of A by AB_1 is complete (data not shown). Data in Figures 8−10 illustrate that the presence of a B domain repeat unit does not alter the collagen-binding characteristics of the A domain.

DISCUSSION

The Organization of the A Domain. The recombinant A domain encompasses the motif of the binding domain truncate, M19 [which includes a trench where, in modeling studies, a triple-helical peptide mimicking collagen was found to fit (*13*)]. However, the full-length A domain of CNA expressed as a recombinant binds collagen with a higher apparent affinity and specificity: fewer high-affinity sites are found for A compared to M19, and the K_D's for the high-affinity binding sites of these two species are 0.3 and 3.0 μM, respectively (*12*). These results may suggest a model where amino acid residues outside the sequence encompassing M19 continue the structural fold determined by M19. Perhaps the collagen-binding trench in the native protein is extended and involves residues outside the sequence encompassed by M19. An alternative model would suggest that the A domain is composed of subunits that fold independently but may interact to make up the structure of this protein. The presence of an intermediate state in the complex denaturation of the A domain (Figure 5) may support this suggestion, since subunits having different stabilities with respect to heat could yield such a unfolding profile. In this second model, M19 could represent such a subunit. At the present time, we cannot differentiate between these two models.

A and B Domains Behave Independently. Analytical ultracentrifugation measurements demonstrated that the AB_1 heterodimer is more spherical than the A domain, suggesting that the two subunits do not have simply end-to-end joining. To determine whether this behavior is due to reorganization of the two subunits upon joining, we compared the far-UV CD and emission spectra, thermal denaturation profiles, and collagen-binding capabilities of A and AB_1.

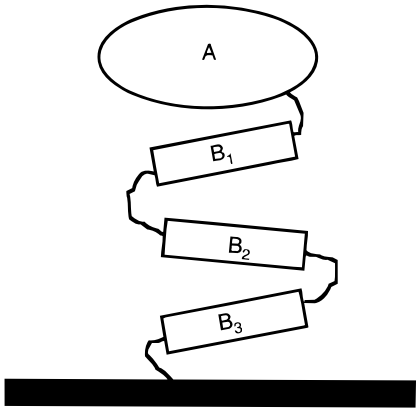

FIGURE 11: Cartoon drawing of the proposed organization of *S. aureus* CNA anchored to the cell surface. For simplicity, only the A and B domains are shown.

Theoretical far-UV CD and tryptophyl emission spectra of AB_1, calculated using eqs 2 and 3, are shown in Figures 4b and 6b. These spectra closely match those obtained experimentally; hence, neither the secondary nor the tertiary structural elements within the A and B_1 subunits undergo detectable reorganization when the two are joined.

Is the complex thermal denaturation observed for AB_1 unique to the heterodimer, a consequence of an interaction between the two domains, or merely a composite of A and B_1 unfolding simultaneously? If the denaturation observed for AB_1 is due only to a sum of the individual subunits (A and B_1) unfolding independently, we should be able to model the curve for AB_1 in Figure 5 by averaging the denaturation curves of A and B_1. The calculated and observed denaturation curves overlay (data not shown), indicating that (1) each subunit unfolds independently and (2) no changes in stability occur as the heterodimer is formed.

This comparison of theoretical and measured profiles of the far-UV CD, emission, and denaturation confirms that the overlap between A and B_1 in the heterodimer does not alter the structure or stability of the two subunits. These results, in conjuction with the collagen-binding measurements shown in Figures 8−10, suggest that the CNA A and B domains do not interact. A plausible construction of the CNA MSCRAMM is shown in Figure 11. In this arrangement, the AB_1 heterodimer would be more globular than the A domain alone, and the two domains make little, if any, contact.

A Model for the CNA B Domain. The far-UV CD spectra of B_1 and B_1B_2 are very similar, indicating that little, if any, alteration in the secondary structure of the individual repeat units occurs upon formation of the B_1B_2 dimer. Not only do B_1 and B_1B_2 have the same percentages of α-helix and β-sheet structural motifs as resolved by a variety of spectral deconvolution programs (Table 3), but also their far-UV CD spectra are almost identical (Figure 4a). The loss of secondary structural organization upon heating occurs similarly for the monomeric and dimeric B domain proteins. No intermediate states in the denaturation were observed for either protein, and the temperature ranges over which the two proteins unfold were alike (Figure 5). A slight increase in T_m is apparent in the dimer, perhaps due to the enhanced

stability arising from hydrogen bonds and/or the interaction of hydrophobic surfaces at the B_1-B_2 interface (*42*).

It is unlikely that the difference in T_m for B_1B_2 compared to B_1 is due to an increased hydrophobicity of tryptophans at the B_1-B_2 interface, however. From a comparison of emission spectra (Figure 6a), it is evident that the overall tryptophyl environments of B_1 and B_1B_2 are quite similar. No obvious rearrangement of subunit tertiary structure occurs nor do any of the five tryptophans that may be surface-exposed in a single B repeat unit become buried in the interface as two B repeat units are joined.

Forming the B_1B_2 homodimer causes an decrease in the sedimentation velocity frictional ratios (Table 2), implying that the axial ratios and frictional ratio of B_1B_2 are less than those of B_1. This observation suggests that the interface is not a simple end-to-end stacking of the repeat units. The time-resolved fluorescence anisotropy measurements provide supporting evidence for the more globular nature of B_1B_2 compared to B_1. The anisotropy decay is longer-lived for B_1B_2 than B_1, but not tremendously so. If the two repeat units stacked end-to-end, making the dimer a significantly more elongated molecule, we would expect the anisotropy decay of the dimer to be even longer-lived. We propose that the B domain repeat units in CNA join as depicted in Figure 11, where at each unit juncture the protein folds somewhat back onto itself. Since no structural or functional interaction is observed between the repeat units of B_1B_2, we conclude that there is little significant contact among any of the B domain repeat units.

That (1) no observable reorganization of the secondary or tertiary structure occurs and (2) no dramatic differences upon denaturation are detected upon forming the B_1B_2 dimer confirms that the two repeat units of B_1B_2 fold independently and in the same manner as the single B_1 unit does. We propose that this autonomy is preserved regardless of the number of repeat units present within a particular strain of CNA-containing *S. aureus*. Consequently, we limited the mixed-domain analyses to the AB_1 recombinant and suggest that this construct is representative of all constructs that include the A domain and any number of B repeat units. Since the A and B domains (and the repeat units within the B domain) are independent, extrapolation of characteristics of the CNA recombinants in vitro to those of the *S. aureus* MSCRAMM in vivo is possible. It must be kept in mind that if domain cooperativity in protein folding, stability, or collagen binding had been observed, this would not be the case.

No ligand binding has yet been identified for the B domain repeat units, and a B domain is apparently not necessary for presenting a functional binding A domain on the surface of *S. aureus cells*. Furthermore, the studies presented here suggest that the B domain does not influence the collagen-binding activity of the A domain and the function of the B domain remains unclear.

CNA May Be a Model for Other MSCRAMM Proteins. Surface proteins on Gram-positive bacteria are often composed of unique and repeated sequences. The repeated sequences in a particular protein can occur in different copy numbers depending on the strain of origin. For example, the B domain in CNA may contain 1, 2, 3, or 4 repeats of a 23 kDa unit in the protein from different isolates. This structural organization, which is derived from sequence analysis, suggests that the MSCRAMMs are mosaic proteins. In this study, we demonstrate that CNA is an example of a protein in which genetic motifs translate into individual structural and functional domains in the protein. Furthermore, the CNA A and B domains are independent structurally and functionally, as are the repeat units within the B domain. Thus, CNA is a true mosaic protein. We suggest that these characteristics may be valid for MSCRAMMS in general. Bottomley et al. concluded from their studies of *S. aureus* protein A, SpA, that the joining the domains in SpA does not affect the structure of the individual domains (*43*). Also, preliminary analyses of the relationship between domains within other MSCRAMMs support our hypothesis (unpublished results).

The occurrence of different numbers of repeat units in the same protein expressed by nonrelated isolates may have arisen from intragenic recombination, where homologous sequences on each side of the repeat unit represent target sites for crossover. Intronless splicing has recently been suggested to be important in the evolution of prokaryotic protein families. These *recer* sequences (*re*combination spa*cer*) are located as structureless links between domains and represent target sites for recombination events. Björck and co-workers described a specific *recer* sequence presumably involved in the shuffling of modules in surface proteins among *Peptostreptococcus magnus* strains (*44, 45*). The *recer* sequence identified in *P. magnus* is not found in *S. aureus* FDA 574 *cna*, but perhaps other *recer* classes have evolved among *S. aureus* strains. Future studies will undoubtedly address these possibilities.

ACKNOWLEDGMENT

We appreciate the assistance of Amy Schneider in the production and isolation of the recombinants and Agneta Höök in the SPR measurements (IBT, Texas A&M University, Houston, TX). Nick Pace (Department of Medical Biochemistry, Texas A&M University, College Station, TX) and Rick Owens (IBT, Texas A&M University, Houston, TX) provided invaluable insight in the interpretation of results.

REFERENCES

1. Switalski, L. M., Patti, J. M., Butcher, W., Gristina, A. G., Speziale, P., and Höök, M. (1993) *Mol. Microbiol. 7*, 99−107.
2. McDevitt, D., Francois, P., Vaudaux, P., and Foster, T. J. (1994) *Mol. Microbiol. 11*, 237−248.
3. Joh, H. J., House-Pompeo, K., Patti, J. M., Gurusiddappa, S., and Höök, M. (1994) *Biochemistry 33*, 6086−6092.
4. Patti, J. M., Allen, B. L., McGavin, M. J., and Höök, M. (1994) *Annu. Rev. Microbiol. 48*, 585−617.
5. Patti, J. M., and Höök, M. (1994) *Curr. Opin. Cell Biol*. 6, 752−758.
6. Foster, T. J., and McDevitt, D. (1994) *Infections Associated with Indwelling Medical Devices* (Bisno, A. L., and Walvogel, F. A., Eds.) 2nd ed., pp 31−44, ASM Press, Washington, D.C.
7. Gillaspy, A. F., Patti, J. M., Pratt, F. L., Jr., Iandolo, J. J., and Smeltzer, M. S. (1997) *Gene 196*, 239−248.
8. Patti, J. M., Jonsson, H., Guss, B., Switalski, L. M., Wiberg, K., Lindberg, M., and Höök, M. (1992) *J. Biol. Chem. 267*, 4766−4772.
9. Patti, J. M., Bremell, T., Krajewska-Pietrasik, D., Abdelnour, A., Tarkowski, A., Rydén, C., and Höök, M. (1994) *Infect. Immun. 62*, 152−161.

10. Nilsson, I.-M., Patti. J. M., Bremell, T., Höök, M., and Tarkowski, A. (1998) *J. Clin. Invest. 101*, 2640−2649.
11. Patti, J. M., Boles, J. O., and Höök, M. (1993) *Biochemistry 32*, 11428−11435.
12. House-Pompeo, K., Boles, J. O., and Höök, M. (1994) *Methods: Compan. Methods Enzymol. 6*, 134−142.
13. Symersky, J., Patti, J. M., Carson, M., House-Pompeo, K., Teale, M., Moore, D., Jin, L., DeLucas, L. J., Höök, M., and Narayana, S. V. L. (1997) *Nat. Struct. Biol. 10*, 833−838.
14. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) *J. Biol. Chem. 272*, 28512−28517.
15. Perona, J. J., Tsu, C. A., Craik, C. S., and Fletterick, R. J. (1997) *Biochemistry 36*, 5381−5392.
16. Hartford, O., McDevitt, D., and Foster, T. J. (1998) *Infect. Immun.* (submitted for publication).
17. Sambrook, J., Fritsch, E. F., and Maniatas, T. (1989) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1991) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York.
19. Marmur, J. (1961) *J. Mol. Biol. 3*, 397−401.
20. Crowe, J. S., Cooper, H. J., Smith, M. A., Sims, M. J., Parker, D., and Gewert, D. (1991) *Nucleic Acids Res. 19*, 184.
21. van Holde, K. E., and Weischet, W. O. (1978) *Biopolymers 17*, 1387−1403.
22. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, W. H. Freeman and Co., New York.
23. Claverie, J.-M., Dreux, H., and Cohen, R. (1975) *Biopolymers 14*, 1685−1700.
24. Todd, G. P., and Haschemeyer, R. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6739−6743.
25. Demeler, B., and Saber, H. (1998) *Biophys. J. 74*, 444−454.
26. Sober, H. (1968) *The Handbook of Biochemistry and Molecular Biology*, Chemical Rubber Co., Cleveland, OH.
27. van Holde, K. E. (1985) *Physical Biochemistry*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.
28. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
29. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
30. Pace, C. N., Vajdos, F., Fee, L., Grmsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
31. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
32. Rich, R. L., Gai, F., Lane, J. W., Petrich, J. W., and Schwabacher, A. W. (1995) *J. Am. Chem. Soc. 117*, 733−739.
33. Titus, J. A., Haugland, R., Sharrow, S. O., and Segal, D. M. (1982) *J. Immunol. Methods 50*, 193−204.
34. Sélo, I., Négroni, L., Créminon, C., Grassi, J., and Wal, J. M. (1996) *J. Immunol. Methods 199*, 127−138.
35. Chen, Y., Rich, R. L., Gai, F., and Petrich, J. W. (1993) *J. Phys. Chem. 97*, 1770−1780.
36. Rich, R. L., Chen, Y., Neven, D., Négrerie, M., Gai, F., and Petrich, J. W. (1993) *J. Phys. Chem. 97*, 1781−1788.
37. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
38. Manavalan, P., and Johnson, W. C., Jr. (1987) *Anal. Biochem. 167*, 76−85.
39. Provencher, S. W., and Glockner, (1981) *J. Biochem. 20*, 33−37.
40. Chang, C. T., Chuen-Shang, C. W., and Yang, J. T. (1978) *Anal. Biochem. 91*, 13−31.
41. Patti, J. M., House-Pompeo, K., Boles, J. O., Garza, N., Gurusiddappa, S., and Höök, M. (1995) *J. Biol. Chem. 270*, 12005−12011.
42. Freire, E., Murphy, K. P., Sanchez-Ruiz, J. M., Galisteo, M. L., and Privalov, P. L. (1992) *Biochemistry 31*, 250−256.
43. Bottomley, S. P., Popplewell, A. G., Scawen, M., Wan, T., Sutton, B. J., and Gore, M. G. (1994) *Protein Eng. 7*, 1463−1470.
44. de Château, M., and Björck, L. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8490−8495.
45. de Château, M., Holst, E., and Björck, L. (1996) *J. Biol. Chem. 271*, 26609−26615.
46. Uversky, V. N. (1993) *Biochemistry 32*, 13288−13298.

BI981773R