The Enterococcus faecalis MSCRAMM ACE Binds Its Ligand by the Collagen Hug Model*§

Qing Liu‡,1, Karthe Ponnura§,1, Yi Xu, Vannakambadi K. Ganesh‡, Jouko Sillanpää§, Barbara E. Murray‡, Sthanam V. L. Narayana, and Magnus Höök‡2

From the ‡Center for Extracellular Matrix Biology, Texas A&M University System Health Science Center, Albert B. Alkek Institute of Biosciences and Technology, Houston, Texas 77030, the §Division of Infectious Diseases, Department of Internal Medicine, and Center for the Study of Emerging and Re-emerging Pathogens, University of Texas Medical School at Houston, Houston, Texas 77030, and the ¶Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai 600025, India

We have determined the crystal structure of the ligand binding segment of the Enterococcus faecalis collagen binding MSCRAMM ACE (microbial surface components recognizing adhesive matrix molecules adhesin of collagen from enterococci). This segment is composed of two subdomains, N1 and N2, each adopting an IgG-like fold and forming a putative collagen binding surface at the interface between the two subdomains. This structure is very similar to that recently reported for CNA, the collagen binding MSCRAMM of Staphylococcus aureus, for which a unique ligand binding mechanism called the Collagen Hug was proposed. We suggest that ACE binds collagen by a similar mechanism and present the first biochemical evidence for this binding model. Replacing residues in the putative collagen binding trench of ACE N2 with Ala residues affected collagen binding. A closed conformation of ACE stabilized by an engineered disulfide bond is unable to bind collagen. Finally, the importance of the residues in the N2 extension in stabilizing the MSCRAMM-ligand complex is demonstrated by selected point and truncation mutations.

During the last decade, Enterococcus faecalis has emerged as a common cause of nosocomial infections, and the organism has been associated with infections such as septicemia, endocarditis, and urinary tract infections (1). The ability of E. faecalis to readily exchange DNA by conjugation allows for the rapid spread of genetic elements and is probably the reason for the observed increase in multidrug resistance among clinical enterococcal isolates (2–4).

The molecular pathogenesis of different enterococcal infections has not been elucidated but presumably involves multiple sets of virulence factors (5). As with most other bacterial infections, the adherence of E. faecalis to host tissues likely represents an early critical step in the infection process (6, 7). Extracellular pathogens, such as enterococci, staphylococci, and streptococci, often target extracellular matrix components for attachment and colonization. The adhesins on the surface of microbes mediating these interactions collectively have been named MSCRAMMs§ (8). Most MSCRAMMs on Gram-positive pathogens belong to a family of structurally related cell wall-anchored proteins. These MSCRAMMs contain an N-terminal signal peptide followed by a non-repetitive region called the A region, which in most cases is responsible for ligand binding. The A regions are composed of two or more subdomains each adopting an immunoglobulin G-like (IgG-like) fold. Following the A region is often a segment composed of repeated sequences or motifs that is referred to as the B region. The C-terminal segment has features required for cell wall attachment including an LPXTG-like motif and a hydrophobic transmembrane region followed by a short positively charged cytoplasmic tail at the end (8).

ACE, a collagen binding MSCRAMM, was the first MSCRAMM identified on E. faecalis (9–11). ACE has a structural organization similar to that of the Staphylococcus aureus collagen adhesin CNA, which has been characterized in some detail (12–16). The CNA A region consists of three subdomains: N1, N2, and N3. The N2 subdomain, a 168-amino-acid-long segment, was earlier characterized as the minimum collagen binding region (12). The crystal structure of the N2 subdomain showed that it adopts an IgG-like jelly-roll fold, composed of two antiparallel β-sheets and two short α-helices. Collagen docking studies by molecular modeling identified a shallow groove on one of the β-sheets of N2 as a putative binding interface (17). Binding studies later showed that the N1N2 protein segment binds collagen with a considerably higher affinity than the N2 subdomain alone or the intact A region (18). The crystal structures of the N1N2 segment of CNA, both as an

* This work was supported by National Institutes of Health Grants AI 20624 (to M. H.) and R37 AI 47923 (to B. E. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† The atomic coordinates and structure factors (code 2Z1P) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ The on-line version of this article (available at http://www.jbc.org) contains a supplemental table.

§ Both authors contributed equally to this work.

1 To whom correspondence should be addressed: Center for Extracellular Matrix Biology, TX A&M University System Health Science Center, Institute of Biosciences and Technology, 2121 West Holcombe Blvd., Suite 603, Houston, TX 77030, Tel.: 713-677-7552; Fax: 713-677-7576; E-mail: mhook@ibt.tamhsc.edu.

‡1 The abbreviations used are: MSCRAMM, microbial surface components recognizing adhesive matrix molecules; ACE, adhesin of collagen from enterococci; SPR, surface plasmon resonance.
apo-protein and in complex with a synthetic collagen triple helical peptide, were reported recently. Based on the analyses of these two structures, a mechanism for the binding of CNA to collagen called the Collagen Hug (Fig. 1) was postulated (18). In the Collagen Hug model, the N1N2 subdomains of CNA are predicted to adopt an open conformation where a binding trench on the N2 subdomain is accessible and allows a collagen triple helix ligand to dock. As a result, the MSCRAMM appears to wrap around the collagen triple helix where the N1 and N2 subdomains of CNA create a “tunnel-like” structure that “locks” the collagen ligand in between the two subdomains. In a final step, the C-terminal extension of the N2 subdomain acts as a “latch” by inserting into a trench present on the N1 subdomain by β-sheet complementation. This event is expected to stabilize the structure (18).

Sequence comparison and structural modeling suggest a structural similarity of the ligand binding regions of ACE and CNA. In fact, there is a 56% sequence similarity between the A (N1N2) regions of the two proteins (9). In the current communication, we report a crystal structure of a segment of the ACE A region that is very similar to the CNA N1N2 apoprotein structure. Furthermore, we report a series of biochemical experiments to examine different aspects of the Collagen Hug binding mechanism for ACE.

**EXPERIMENTAL PROCEDURES**

**Bacterial and Growth Conditions—**Escherichia coli strains JM101 and XL1-Blue (Stratagene, La Jolla, CA) were used as the bacterial host for plasmid cloning and protein expression. E. coli strains were grown at 37 °C in Luria Bertani (LB) broth or on LB agar with appropriate antibiotics.

**Construction of Expression Plasmids—**A series of expression plasmids was constructed using the vector pQE-30 (Qiagen Inc., Chatsworth, CA) essentially as described previously (9). Recombinant proteins expressed from this vector contain an N-terminal tail of 6 histidine residues.

**Site-directed Mutagenesis—**Site-directed mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Stratagene). The introduction of specific mutations into ACE32–367 was achieved by the use of pQE-30 ACE32–367 (pQE-30 with a 1-kb insert containing ACE32–367 from E. faecalis strain OG1RF) as template DNA with the mutagenic primers listed in supplemental Table 1. After initial denaturation at 95 °C for 30 s, the cycling parameters were 30 s at 95 °C followed by 1 min at 55 °C and 5 min at 68 °C (18 cycles).

The reaction mixtures were placed on ice for 2 min. To minimize the chance of the supercoiled double-stranded template DNA being transformed, the mixtures were digested with DpnI at 37 °C for 1 h before being transformed into competent E. coli XL1-Blue cells. Cloned sequences were confirmed with automated DNA sequencing analysis (ABI Prism DNA sequencer; Applied Biosystems) at the University of Texas Medical School DNA sequencing facility, using pQE-30 sequencing primers.

**Expression and Purification of Recombinant Proteins—**Large scale expression and purification of the recombinant fragments (Fig. 2) by a nickel-charged HiTrap chelating column and an anion-exchange Sepharose column (Amersham Biosciences) were as described previously (19). Protein-containing fractions were analyzed by SDS-PAGE, appropriately pooled. Finally, except for ACE32–367N124C, N328C, protein samples were dialyzed extensively against phosphate-buffered saline, pH 7.4, stored at 4 °C, and generally used within 2 weeks. Protein con-
centrations were determined by absorption spectroscopy at 280 nm using calculated molar absorption coefficient values (20). Molecular masses of the expressed proteins were confirmed with matrix-assisted laser desorption/ionization-time of flight mass spectrometry (Tufts University mass spectrometry facility) from protein samples in H$_2$O.

To remove multimers formed through intermolecular disulfide bonds, post anion-exchange ACE$_{32–367}$N124C, N328C was pooled and applied to size-exclusion chromatography. Size-exclusion chromatography was performed on a 80-ml Superdex 200 column (Amersham Biosciences) equilibrated and eluted with 50 mM Tris, 150 mM NaCl (pH 8.0) using a fast protein liquid chromatography system (Amersham Biosciences). One-ml fractions were collected. A flow rate of 1 ml/min was maintained throughout this procedure. Protein-containing fractions were analyzed by SDS-PAGE, appropriately pooled, and dialyzed against 50 mM NaH$_2$PO$_4$, 0.5M NaCl, 1 mM EDTA (pH 7.5). Finally, to remove the small amount of protein in which the introduced Cys residues had not formed disulfide bonds, the protein sample was passed through a 5-ml activated thiol-Sepharose 4B column previously equilibrated with 50 mM NaH$_2$PO$_4$, 0.5 mM NaCl, 1 mM EDTA (pH 7.5).

Surface Plasmon Resonance (SPR) Analysis—SPR analysis was performed at 25 °C using a BIAcore 3000 system (BIAcore AB, Uppsala, Sweden) based on the method described previously (21). Analysis of the association and dissociation rates was performed using the BIAevaluation 3.0 software (BIAcore). The $K_D$ values for the binding of recombinant MSCRAMMs to the immobilized collagen were calculated by Scatchard analysis based on the responses at the steady state portion of the sensorgrams using a one-site binding non-linear regression model (19).

**Structure Determination of ACE$_{32–367}$**—The crystallization of the ACE$_{32–367}$ segment was reported previously (22). The structure of ACE$_{32–367}$ was determined by the molecular replacement method using the crystal structure of the N$_2$ subdomain of ACE as search model (23). Briefly, ACE$_{32–367}$ crystals were grown by the hanging-drop method using 2-methyl-2,4-pentanediol as a precipitant. Crystals diffracted to 2.5 Å resolution, and the data were collected at 100 K using an R-AXIS IV image plate detector with Rigaku rotating anode generator. Data processing and scaling were performed using DENZO and SCALEPACK (24). The unit cell parameters and data reduction statistics are presented in Table 1. Phase improvement was performed using density modification by solvent flipping and density truncation without phase extension. Experimental amplitudes (20–2.5 Å data) and phases calculated from the model were used in a density modification by solvent flipping and density truncation without phase extension. Experimental amplitudes (20–2.5 Å data) and phases calculated from the model were converged to 22.9%/26.5%. Final refinement statistics are listed in Table 1. The 41 C-terminal residues and the 9-residue linker residues that connect N$_1$ and N$_2$ subdomains could not be identified from the electron density maps. In addition, breaks in the main chain were observed at three places (residues 68, 122–123, and 142). The side chains for some residues in a few loop regions appeared to be genuinely disordered in the crystal. The final model contains 282 residues out of a total of 336 residues.
Molecular Modeling—Modeling of ACE-Collagen peptide complex was performed using CNA-collagen complex (Protein Data Bank ID: 2F6A) as template. N1 and N2 subdomains of CNA were superimposed on the corresponding subdomains of ACE. The transformation matrix used for the N2 domain superposition was used for the rigid body transformation of the collagen triple helix from the CNA-collagen complex. A stereo view of sections of the N2 subdomain and the collagen triple helix is shown in Fig. 5B. The N1 and N2 subdomains of ACE could adopt a similar interdomain orientation as observed in CNA-collagen complex without steric clashes. Building of the linker residues that were not observed in the ACE crystal structure was not attempted. Superposition was performed using the secondary structure matching method (26). Figs. 1, 3, 4, 5B, and 7 were made using RIBBONS (27).

RESULTS

Crystal Structure of ACE32–367—ACE32–367 is folded into two subdomains, N1 and N2, as expected from its sequence similarity to CNA (Fig. 3). Both subdomains are similar in size (N1, 146 residues; N2, 135 residues) and adopt the DEv-IgG (DEvariant-IgG) fold, a variant of the C-type immunoglobulin fold (28). The structure of ACE32–367 is very similar to the previously reported structure of CNA31–344. In ACE, the N2 subdomain has two principal β-sheets (βI and βII), and each sheet consists of five antiparallel β-strands. βI and βII stack face-to-face and form a “β-sandwich” with jelly roll topology. The N2 subdomain exhibits a shallow groove on the βI sheet, representing a putative collagen binding region. The N1 subdomain is also composed of 10 β-strands grouped into two β-sheets forming a β-sandwich.

As observed in the structure of apo-CNA31–344 (18), the G′ strand is not formed by the residues of the N1 subdomain. Instead, this strand is created by the C-terminal extension of the N2 subdomain corresponding to residues 320–325, which complements the βI sheet in N1. In particular, the association of N1 and N2 generates a deep cleft at their interface, lined by the putative collagen binding trench and two loop segments of the N1 subdomain. The linker residues 165–173 that connect the two subdomains are disordered and therefore not included in the refinement. However, modeling of the linker based on the observed discrete densities in the 2F0 − F map suggests that the linker forms a wall in the tunnel-like structure as observed in the CNA-collagen complex and could play a similar role in locking the collagen ligand in place. Because of the structural similarities between the structures of CNA31–344 and ACE32–367, we hypothesize that ACE binds its ligand by the
Collagen Hug mechanism, which was previously proposed for the CNA collagen interaction (18).

Structural Differences between CNA and ACE—The individual N1 and N2 subdomains of ACE fold similarly to those of CNA, with root mean square deviations of 0.81 and 1.32 Å, respectively, for the corresponding Cα atoms forming the core β-sheet regions of the N1 and N2 subdomains. However, superposition of either one of the corresponding subdomains on CNA from the CNA-collagen complex clearly showed a significant deviation in the position of the other subdomain. Superposition of the N1 subdomains of ACE and CNA is shown in Fig. 4. Analysis of these crystal structures using the program DynDom (29) showed that the interdomain orientation of the N1 and N2 subdomains in ACE is altered by about 36° as compared with that of CNA. Due to this difference in the interdomain orientation, the apo-ACE structure is in a partially open conformation as compared with the closed conformation of the CNA-collagen crystal structure.

The electron density for the C-terminal residues (326–331) of the latch (320–331) is not observed in the ACE structure, and only two pairs of hydrogen bonds between the β-strands G′ and E of apo-ACE are weaker than those observed between the corresponding strands in the crystal structures of CNA as an apoprotein and in complex with the collagen.

Examining the Putative Ligand Binding Trench by Mutagenesis Analysis—The structure of CNA31–344 in complex with a synthetic collagen-like peptide showed that the collagen peptide binds at a shallow binding cleft formed on the surface of the N2 subdomain and makes additional contacts with a few residues in the linker connecting N1 and N2. Arg-15 and Hyp-12 of the middle strand in the collagen helix make hydrogen bonds with a backbone oxygen and a side chain oxygen (OG), respectively, of Thr-169 in one copy of the CNA molecule in the asymmetric unit. In another copy of the molecule in the asymmetric unit, the backbone conformation of the linker is slightly different, and therefore, the interactions between the linker residues and the collagen triple helix are different but significant. Examination of the CNA-collagen complex suggests that bulky, hydrophilic residues would be more favorable to make interactions with the linker by reaching out to the linker residues. The interactions with the collagen triple helix at the N2 subdomain is conserved in all the CNA molecules in the asymmetric unit. Therefore, it could be hypothesized that a high affinity binding site for MSCRAMMs would be 3–4 GPO repeats flanked by hydrophilic residues containing GXY repeats, where GPO repeats are primarily recognized by the N2 subdomain. Residues that make contact with the collagen peptide in the CNA N2 subdomain (Tyr-175, Thr-177, Gly-178, Leu-181, Glu-183, Arg-189, Phe-191, Asn-196, Asn-223, and Tyr-223) are largely conserved in ACE (Fig. 5A). Mutational analysis was used to evaluate the importance of these residues in ACE32–367 for collagen binding. Four residues (Tyr-180, Arg-193 Phe-195, and Asn-197) within the binding region and one residue (Lys-237) outside the groove were individually replaced with Ala residues. The binding of recombinant wild type ACE32–367 and mutant proteins to collagen type I immobilized on a sensor chip was analyzed by SPR using a BIAcore 3000 system. Mutations at positions Tyr-180, Arg-193, and Phe-195 resulted in proteins with much lower affinity for collagen as compared with wild type ACE32–367 (Fig. 6, Table 2), suggesting that the mutated residues participate in collagen binding as
observed in the CNA-collagen complex. Examination of the CNA-collagen complex shows that the corresponding residues of Phe-195 and Tyr-180 in CNA (Phe-191 and Tyr-175, respectively) make stacking interactions with proline residues at position \( X \) in the \((G\alpha\chi\chi)\) collagen peptide repeats. In the CNA-collagen structure, these residues are a part of the collagen binding groove in the N2 subdomain. Replacing the aromatic side chain of these residues with a single carbon atom side chain would severely affect the stacking interaction. The decreased affinity for collagen exhibited by the ACE F195A and Y180A mutants suggested that similar stacking interactions may also occur in ACE. Surprisingly, a significant increase in apparent affinity for collagen as compared with wild type ACE was observed when Asn-197 in ACE was replaced with Ala (Fig. 6, Table 2). A Closed Conformation of ACE Does Not Bind Collagen—In the Collagen Hug model, only the open form of the MSCRAMM would have its ligand binding trench accessible to collagen. To examine this part of the proposed binding model, we attempted to "freeze" the closed conformation of ACE by introducing a disulfide bond between the latch (G/H)11 strand and the E strand of the N1 subdomain. Two amino acid positions (Asn-124 and Asn-328) were identified by molecular modeling to be at a suitable distance and orientation to form disulfide bonds. These residues were replaced with Cys residues (Figs. 2 and 7A). The resulting protein ACE32–367N124C,N328C migrated somewhat faster than the wild type protein on SDS-PAGE under non-reducing conditions, whereas under reducing conditions, the migration of the mutant protein was found to be very similar to that of the wild type protein (Fig. 7B, panel 1). This observation suggests that the mutated recombinant protein forms an internal disulfide bond and maintains a compact closed conformation of the protein during SDS-PAGE.

Collectively, mutations of residues in the proposed ligand binding site in ACE showed either positive or negative effects on collagen binding, indicating that these residues in ACE, identified through comparative analyses with CNA, are involved in ligand binding. The K237A mutation, which involves a residue distant from the putative binding region, did not affect the affinity of ACE32–367 for collagen. (Fig. 6, Table 2).

A Closed Conformation of ACE Does Not Bind Collagen—In the Collagen Hug model, only the open form of the MSCRAMM would have its ligand binding trench accessible to collagen. To examine this part of the proposed binding model, we attempted to "freeze" the closed conformation of ACE by introducing a disulfide bond between the latch (G/H)11 strand and the E strand of the N1 subdomain. Two amino acid positions (Asn-124 and Asn-328) were identified by molecular modeling to be at a suitable distance and orientation to form disulfide bonds. These residues were replaced with Cys residues (Figs. 2 and 7A). The resulting protein ACE32–367N124C,N328C migrated somewhat faster than the wild type protein on SDS-PAGE under non-reducing conditions, whereas under reducing conditions, the migration of the mutant protein was found to be very similar to that of the wild type protein (Fig. 7B, panel 1). This observation suggests that the mutated recombinant protein forms an internal disulfide bond and maintains a compact closed conformation of the protein during SDS-PAGE.

**TABLE 2**

| Construct          | \( K_D \) (app) /M |
|--------------------|-------------------|
| ACE32–367          | 45.6              |
| ACE32–320          | 196               |
| ACE32–325          | 185               |
| ACE32–340          | 53                |
| ACE32–367G304P     | 119               |
| ACE32–367G304R     | 245               |
| ACE32–367Y149A     | 248               |
| ACE32–367R162A     | 90.3              |
| ACE32–367F164A     | 317               |
| ACE32–367N166A     | 42.7              |
| ACE32–367N193A     | 26.6              |
under non-reducing conditions. Multimers of the protein presumably linked by intermolecular disulfide bonds were also observed in the gel (Fig. 7B, panel 1). These multimers were removed by size-exclusion chromatography (Fig. 7B, panel 2). A shadow of mutated ACE32–367, migrating as a wild type protein in SDS-PAGE under non-reducing conditions was also detected. It is likely that in this population of the protein, the Cys residues have not formed disulfide bonds. Proteins with free Cys residues were removed by passing the protein through a column of activated thiol-Sepharose. Structural analyses of the purified disulfide bond stabilized protein by near- and far-UV CD spectroscopy suggested that no significant changes in secondary structure had occurred as a result of the mutations (data not shown). A BIACore sensogram for the disulfide bond stabilized closed form of ACE32–367 (10 μM) run over immobilized collagen type I showed that ACE32–367 N124C, N328C did not bind collagen type I (Fig. 8A).

The Latch in ACE32–367 Is Required for Effective Collagen Binding—In the Collagen Hug binding model, the insertion of the C-terminal latch sequence into the neighboring N1 subdomain is predicted to stabilize the closed conformation of the protein-ligand complex. To evaluate the role of the latch sequence, three truncated versions of ACE32–367 were generated in which the latch sequence is completely removed (ACE32–320), partially truncated (ACE32–325), or intact (ACE32–340) but with the C-terminal latch extension removed (Figs. 2 and 7A). Binding of the truncation mutants to collagen was analyzed by SPR. Consistent with the predicted role of the latch sequence in collagen binding, the ACE32–325 and ACE32–320 polypeptides showed a markedly reduced affinity for the ligand. The calculated $K_D$ of these ACE forms with truncations extending into the latch sequence were 185 (ACE32–325) and 196 μM (ACE32–320) respectively. On the other hand, the affinity of ACE32–340 ($K_D = 53$ μM) was similar to that of the intact ACE32–367 ($K_D = 45.6$ μM) (Fig. 8B, Table 2).

Mutations in the Latch Sequence Affect Collagen Binding—Two critical steps in the Collagen Hug model are the redirection of the latch sequence and insertion of the sequence into the latching trench in the N1 subdomain. Modeling of amino acid replacements in the latch sequence suggested that a G323P mutation (Figs. 2 and 7A) should prevent the redirection of the latch sequence on ligand binding and that a G323R mutation (Figs. 2 and 7A), due to the introduction of a bulky residue at this position, might result in the latch not being able to fit optimally in the latching trench. Both mutations were predicted to disrupt the latching event and affect ligand binding. The proposed mutants of Ace32–367, were produced, and their bindings to collagen were examined by SPR analysis. Both recombinant proteins, ACE32–367 G323P and G323R, showed significantly reduced binding affinities for collagen with calculated $K_D$, of 119 and 245 μM, respectively, as compared with a $K_D = 46$ μM determined for the wild type protein (Fig. 8C, Table 2).

DISCUSSION

Enterococci are part of the normal human intestinal flora but have also emerged as important opportunistic pathogens, ranked third among the organisms isolated from nosocomial infections. In particular, *E. faecalis* can cause many clinical infections and is responsible for 5–15% of cases of bacterial endocarditis. Little is known about the molecular mechanisms that allow enterococci to colonize host tissues, translocate across epithelial barriers, survive in different host environments, and cause disease. It is likely that multiple sets of virulence factors are involved and that proteins on the surface of the bacteria play key roles in these processes. ACE is the first MSCRAMM identified on *E. faecalis*. The now proposed Collagen Hug model of ACE binding to collagen provides a molecular basis for the formation of the adhesin-ligand complex.

The Collagen Hug model was derived from the recently published crystal structures of the ligand binding region of CNA both as an apoprotein and in complex with a synthetic collagen-like triple helix. Analysis of the structure of the now reported
structure of ACE_{32–367} revealed that the two individual subdomains of ACE have very similar structures to the CNA subdomains and contain all the structural features required for a Collagen Hug mechanism of ligand binding. We have now for the first time experimentally examined this hypothetical dynamic binding mechanism. Structural rearrangement is one of the critical parts of the Collagen Hug mechanism, and we predicted that the MSCRAMM needs to adopt an open conformation to allow the docking of the collagen triple helix to the binding trench of N2 (Fig. 1). However, in the structures of the apo-form of CNA N1N2 and ACE N1N2, the MSCRAMMs appear in a closed or partially closed conformation with the N2 extension, inserted into a latching trench on N1 and with apparent contacts between the N1 and the N2 subdomains. This closed conformation could be a snapshot of a conformation in equilibrium between open and closed conformation, stabilized by crystal packing. The partial closed (or partial open) conformation, observed in the ACE crystal structure, shows some evidence for the molecule in equilibrium (Fig. 1.). Therefore, the crystal structures could represent an inactive closed conformation, and so we predicted that only an open conformation of the MSCRAMM is capable of binding collagen. To examine this prediction, we stabilized the closed conformation of ACE N1N2 by introducing a disulfide bond between a Cys residue introduced at the bottom of the N1 subdomain. With these two point mutants, a disulfide bond is formed spontaneously. Consistent with the Collagen Hug model, the closed conformation of ACE stabilized by a disulfide bond did not bind type I collagen.

Using a mutagenesis approach, we also demonstrated that the putative collagen binding cleft on N2 is in fact the ligand binding site. We identified residues that were predicted to make contact with a docked collagen molecule by molecular modeling (Fig. 5B). Mutations of some of these residues resulted in a loss of collagen binding activity, and in one case, a gain of function. These residues are conserved in CAN, and molecular modeling based on the CNA-collagen complex structure could explain the effects of the mutations on ACE/collagen binding. The residues in the shallow ligand binding trench of CNA N2 identified to make contact with collagen and the solved structure of CNA N2N3 in complex with a collagen peptide suggest that the N2 trench can accommodate a generic (GPO)n triple helical peptide. This is consistent with our earlier studies that demonstrate that CNA can bind to multiple sites along a collagen monomer with varying affinities (12). It is not clear what specific residues in collagen and the CNA or ACE contribute to this varying binding affinity for different sites. MSCRAMM can interact to define a high affinity intermediate or low affinity site. In this context, it is also possible that some residues in collagen are not compatible with MSCRAMM binding and define a non-binding site. The dynamic nature of the collagen triple helix on binding makes it difficult to use a molecular modeling approach to define the importance of different residues. We demonstrated that introduction of a proline or an arginine residue in the N2 extension or truncation of the latch sequence affected the latching function of the segment, resulting in MSCRAMM proteins with reduced affinity for collagen.

The structure presented in this study shows a significant difference from CNA in the interdomain orientation, possibly due to a weak latch. A weak latch would allow the molecule to adopt an intermediate conformation between the open and closed form and provide a snapshot of a stage in the binding process. Therefore, the structure determined for ACE and the biochemical analyses of the selected ACE mutants provide strong evidence in support of the predicted multistep model of ligand binding and strand complementation as proposed in the Collagen Hug model.

Although the previously determined CNA-collagen structure and ACE crystal structure and biochemical analysis presented here provide evidence for a Collagen Hug model at an atomic level, a key biological question is the relevance of this mechanism of binding to supramolecular fibrillar arrangement in vivo. A simplistic explanation for MSCRAMM binding to collagen would be the presence of disrupted, broken collagen molecules in damaged tissues. These collagen molecules could serve as anchoring templates for the pathogens involved in wound infections. The broad specificity of CNA or ACE MSCRAMMs would help identify binding sites in any monomeric collagen extending off the fiber. It is possible that the collagen binding MSCRAMM could use the Collagen Hug binding mechanism to target any other molecules that contain a triple helical collagenous domain. CNA has been shown to contribute to the virulence of _S. aureus_ in a septic arthritis (30) model, and this property is due to its collagen binding activity. Studies have shown that an ACE mutant is significantly less virulent in a rat endocarditis model as compared with the parent wild type model.4 Thus, the two MSCRAMMs appear to play important roles in the disease processes, but the collagenous host molecules targeted by the adhesions remain to be defined.

The Collagen Hug model and the Dock, Lock, and Latch mechanism previously proposed for the binding of a linear fibrinogen peptide to the SdrG MSCRAMM (31) have several similarities. The two MSCRAMMs are structurally very similar. In both models, the ligand binding region is situated in between the two subdomains: N2 and N3 in SdrG and N1 and N2 in ACE (or CNA). The ligand-MSCRAMM complex is apparently stabilized by insertion of the extension of the second subdomain into an incomplete β-sheet in the first subdomain. The Collagen Hug model also has some unique features. The linker region (165–173) connecting the N1 and N2 subdomains of ACE_{32–367}, which is required for wrapping around the triple helix, is much longer than the one connecting the N2 and N3 subdomains of SdrG. Thus, the ACE N1 and N2 subdomains are far apart from each other, making it possible for the bulky collagen triple helix molecule to fit into the binding cleft before the insertion of latch sequence into the latching trench. It seems that various MSCRAMMs can utilize similar structural building blocks (IgG-like modules) to bind to different ligands with similar mechanisms that are modified to fit the structural specificities of the ligands.

---

4 B. Murray, personal communication.
REFERENCES

1. Murray, B. E. (1990) Clin. Microbiol. Rev. 3, 46–65
2. Murray, B. E., and Weinstock, G. M. (1999) Proc. Assoc. Am. Physicians 111, 328–334
3. Horodniceanu, T., Bougueleret, L., El-Solh, N., Bieth, G., and Delbos, F. (1979) Antimicrob. Agents Chemother. 16, 686–689
4. Clewell, D. B. (1990) Eur. J. Clin. Microbiol. Infect. Dis. 9, 90–102
5. Jett, B. D., Huycke, M. M., and Gilmore, M. S. (1994) Clin. Microbiol. Rev. 7, 462–478
6. Patti, J. M., and Hook, M. (1994) Curr. Opin. Cell Biol. 6, 752–758
7. Foster, T. J., and Hook, M. (1998) Trends Microbiol. 6, 484–488
8. Patti, J. M., Allen, B. L., McGavin, M. J., and Hook, M. (1994) Annu. Rev. Microbiol. 48, 585–617
9. Rich, R. L., Kreikemeyer, B., Owens, R. T., LaBrenz, S., Narayana, S. V., Weinstock, G. M., Murray, B. E., and Hook, M. (1999) J. Biol. Chem. 274, 26939–26945
10. Nallapareddy, S. R., Qin, X., Weinstock, G. M., Hook, M., and Murray, B. E. (2000) Infect. Immun. 68, 5210–5217
11. Patti, J. M., Boles, J. O., and Hook, M. (1993) Biochemistry 32, 11428–11435
12. Switalski, L. M., Speziale, P., and Hook, M. (1989) J. Biol. Chem. 264, 21080–21086
13. Patti, J. M., Jonsson, H., Guss, B., Switalski, L. M., Wiberg, K., Lindberg, M., and Hook, M. (1992) J. Biol. Chem. 267, 4766–4772
14. Patti, J. M., House-Pompeo, K., Boles, J. O., Garza, N., Gurusiddappa, S., and Hook, M. (1995) J. Biol. Chem. 270, 12005–12011
15. Rich, R. L., Demeler, B., Ashby, K., Deivanayagam, C. C., Petrutch, J. W., Patti, J. M., Narayana, S. V., and Hook, M. (1998) Biochemistry 37, 15423–15433
16. Symersky, J., Patti, J. M., Carson, M., House-Pompeo, K., Teale, M., Moore, D., Jin, L., Schneider, A., DeLucas, L. J., Hook, M., and Narayana, S. V. (1997) Nat. Struct. Biol. 4, 833–838
17. Zong, Y., Xu, Y., Liang, X., Keene, D. R., Hook, A., Gurusiddappa, S., Hook, M., and Narayana, S. V. (2005) EMBO J. 24, 4224–4236
18. Visai, L., Xu, Y., Casolini, F., Rindi, S., Hook, M., and Speziale, P. (2000) J. Biol. Chem. 275, 39837–39845
19. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) Protein Sci. 4, 2411–2423
20. Xu, Y., Gurusiddappa, S., Rich, R. L., Owens, R. T., Keene, D. R., Mayne, R., Hook, A., and Hook, M. (2000) J. Biol. Chem. 275, 38981–38989
21. Ponnuraj, K., Xu, Y., Moore, D., Deivanayagam, C. C., Boque, L., Hook, M., and Narayana, S. V. (2002) Biochim. Biophys. Acta 1596, 173–176
22. Ponnuraj, K., and Narayana, S. V. L. (2007) Proteins, in press
23. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
24. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
25. Krissinel, E., and Henrick, K. (2004) Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2256–2268
26. Deivanayagam, C. C. (2002) EMBO J. 21, 6660–6672
27. Hayward, S., and Berendsen, H. J. (1998) Proteins 30, 144–154
28. Patti, J. M., Bremell, T., Krajewska-Pietrasik, D., Abdelnour, A., Tarkowski, A., Ryden, C., and Hook, M. (1994) Infect. Immun. 62, 152–161
29. Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., Xu, Y., Hook, M., and Narayana, S. V. (2003) Cell 115, 217–228