Crystal structure of the von Willebrand factor A domain of human capillary morphogenesis protein 2: An anthrax toxin receptor

D. Borden Lacy^{*†}, Darran J. Wigelsworth^{*†}, Heather M. Scobie[‡], John A. T. Young[‡], and R. John Collier^{*§}

*Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115; and [†]Infectious Disease Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037

Contributed by R. John Collier, March 3, 2004

Anthrax toxin is released from *Bacillus anthracis* as three monomeric proteins, which assemble into toxic complexes at the surface of receptor-bearing host cells. One of the proteins, protective antigen (PA), binds to receptors and orchestrates the delivery of the other two (the lethal and edema factors) into the cytosol. PA has been shown to bind to two cellular receptors: anthrax toxin receptor/tumor endothelial marker 8 and capillary morphogenesis protein 2 (CMG2). Both are type 1 membrane proteins that include an ≈ 200 -aa extracellular von Willebrand factor A (VWA) domain with a metal ion-dependent adhesion site (MIDAS) motif. The anthrax toxin receptor/tumor endothelial marker 8 and CMG2 VWA domains share $\approx 60\%$ amino acid identity and bind PA directly in a metal-dependent manner. Here, we report the crystal structure of the CMG2 VWA domain, with and without its intramolecular disulfide bond, to 1.5 and 1.8 Å, respectively. Both structures contain a carboxylate ligand-mimetic bound at the MIDAS and appear as open conformations when compared with the VWA domains from α -integrins. The CMG2 structures provide a template to begin probing the high-affinity CMG2-PA interaction (200 pM) and may facilitate understanding of toxin assembly/internalization and the development of new anthrax treatments. The structural data also allow molecular interpretation of known CMG2 VWA domain mutations linked to the genetic disorders, juvenile hyaline fibromatosis, and infantile systemic hyalinosis.

Anthrax toxin is comprised of three nontoxic proteins; protective antigen (PA), edema factor (EF), and lethal factor (LF), which are released from the anthrax bacillus as monomers and assemble into toxic complexes on the surface of host cells (1). PA binds cell-surface receptors and mediates delivery of the two catalytic factors, EF and LF, into the host cell cytosol. EF, an 89-kDa calmodulin-dependent adenylate cyclase, elevates levels of cAMP, which benefits the bacillus by impairing macrophage functions (2, 3). LF, named for its lethal effect in rodent model systems, is a 90-kDa zinc protease that has been shown to cleave mitogen-activated protein kinase kinases (4, 5). This activity may impair dendritic cell function (6) and stimulate apoptosis in activated macrophages (7) and endothelial cells (8), thereby promoting bacterial survival and pathogenesis.

The current model for cellular intoxication involves a multi-step mechanism. In the first step, the full-length 83-kDa form of PA binds a host cell-surface receptor. Binding is followed by cleavage of the full-length 83-kDa form of PA by furin or a furin-like protease, resulting in removal of a 20-kDa fragment from the N terminus (9). The remaining 63-kDa PA is then able to oligomerize into a heptameric, receptor-bound prepore (10). Oligomerization of the 63-kDa PA also creates binding sites for EF and/or LF (11). The entire toxin-receptor complex is internalized by receptor-mediated endocytosis and trafficked to a low-pH endosome (12, 13). The increasing acidity of the endosome induces a conformational rearrangement of the prepore assembly that allows formation of a pore across the endosomal membrane (14). Pore formation is coincident with

translocation of fully or partially unfolded EF and/or LF into the cytosol, where catalysis can occur.

Two PA receptors have been identified: anthrax toxin receptor/tumor endothelial marker 8 (ATR/TEM8) and capillary morphogenesis protein 2 (CMG2) (15, 16). Both appear to be expressed in a wide variety of tissues, and both are predicted to exist in multiple isoforms due to alternative splicing (15, 16). Common features of ATR/TEM8 and CMG2 include a signal peptide, an extracellular von Willebrand factor A (VWA) domain, and a single-pass transmembrane region. ATR/TEM8 has been implicated in the neovascularization of human colon tumors and developing mouse embryos (17, 18). The *CMG2* gene was identified as being highly expressed in endothelial cells during capillary morphogenesis in three-dimensional collagen matrices (19). The VWA domain of CMG2 binds selectively to collagen type IV and laminin, suggesting a role in basement membrane matrix adhesion (19). Whereas the physiological role of CMG2 is unclear, mutations within the *CMG2* gene have been shown to cause two allelic disorders, juvenile hyaline fibromatosis (JHF) and infantile systemic hyalinosis (ISH) (20, 21). Furthermore, fibroblasts isolated from patients with these diseases exhibit defects in adhesion to laminin (20).

VWA domains are often protein–protein interaction sites in cell adhesion proteins, such as integrins and extracellular matrix proteins (22). Approximately 46% of all VWA domains have a perfectly conserved metal ion-dependent adhesion site (MIDAS) motif (DXSXS...T...D, where X is any amino acid), which is often involved in ligand binding (22, 23). The best structural characterization of MIDAS-mediated protein–protein interactions has stemmed from the study of α -integrin VWA, or inserted (I), domains (24). Crystal structures and binding studies have shown that recombinant I domains exist in two distinct conformations; open and closed, with high and low affinities for ligand, respectively (24). The ability to interconvert between affinity states is believed to be critical in integrin signal transmission across the cellular membrane. In the absence of a ligand, ligand-mimetic, or engineered disulfide crosslink, I domains tend to crystallize in their low-affinity closed conformations. The conversion of I domains to an open high-affinity state requires rearrangement of the MIDAS and is linked to structural changes throughout the protein (25–27).

ATR/TEM8 and CMG2 proteins share 40% overall amino acid identity, with 60% identity within their VWA domains.

Abbreviations: PA, protective antigen; ATR/TEM8, anthrax toxin receptor/tumor endothelial marker 8; CMG2, capillary morphogenesis protein 2; S38, CMG2 residue S38-C218; R40, CMG2 residue R40-S217; VWA, von Willebrand factor A; MIDAS, metal ion-dependent adhesion site; EF, edema factor; LF, lethal factor; JHF, juvenile hyaline fibromatosis; ISH, infantile systemic hyalinosis; I, inserted.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 1SHT and 1SHU).

[†]D.B.L. and D.J.W. contributed equally to this work.

[§]To whom correspondence should be addressed. E-mail: icollier@hms.harvard.edu.

© 2004 by The National Academy of Sciences of the USA

Table 1. Data collection

	R40 Pt λ_1	R40 Pt λ_2	R40 Pt λ_3	R40 Native	S38 Native
Wavelength, λ	1.0723	1.0332	1.0720	0.9186	0.9186
Resolution, $\text{\AA}$	50–1.9	50–1.9	50–1.9	50–1.8	28.75–1.5
Reflections					
Total	98,182	98,237	97,959	144,891	214,338
Unique	13,510	13,482	13,512	15,729	30,193
Completeness, %	99.4 (99.6)	99.4 (99.9)	99.5 (100)	99.9 (100)	98.2 (90.5)
$I/\sigma(I)$	30.4 (13.7)	28.3 (13.1)	32.4 (9.4)	34.3 (4.6)	33.3 (7.7)
R_{sym} %	4.3 (14.7)	4.6 (16.3)	4.3 (22.5)	6.5 (42.6)	4.5 (17.4)

Numbers in parentheses refer to the values for the outer shell.

including a perfectly conserved MIDAS motif (16). Soluble, recombinant VWA domains from ATR/TEM8 and CMG2 have been shown to bind PA directly in a divalent cation-dependent manner, and such soluble VWA domains can inhibit anthrax intoxication of cultured cells (15, 16). Preservation of the MIDAS motif is required for ATR/TEM8 binding to PA, an observation that suggests the ATR/TEM8–PA interaction will be similar to those observed between α -integrin I domains and their MIDAS-dependent ligands (28). On the other hand, key differences between these interactions are possible given that the toxin receptor VWA domains share only 15–27% amino acid identity with the α -integrin I domains. The α -integrin I domains bind their physiological ligands with weak affinity. For example, α M integrin I domain binds fibrinogen at a K_d of $\approx 3 \mu\text{M}$ (29). Whereas the affinity of CMG2 for its natural ligands is still unknown, the interaction between PA and CMG2 is extremely tight ($K_d \approx 200 \text{ pM}$; ref. 51). We have initiated structural studies of the CMG2 VWA domain in an effort to understand the molecular basis for this high affinity and to probe whether it undergoes a conformational switch similar to that seen in the I domains of α -integrins.

In this report, we have determined atomic resolution structures of the CMG2 VWA domain with and without a naturally occurring intrachain disulfide bond. Both structures contain a chelated Mg^{2+} ion in the MIDAS with bound pseudoligands contributed by either a Glu residue from a neighboring monomer or an acetate ion from the crystallization buffer. The structures most closely resemble the open high-affinity conformation of the α M integrin I domain (23). These results provide a framework for analyzing the high affinity of CMG2 for PA and the role of receptor in anthrax toxin action. The structures also serve as a basis for understanding the interaction of CMG2 with its physiological ligands and the effect of CMG2 VWA domain point mutations linked to JHF and ISH.

Materials and Methods

Expression and Purification of the CMG2 VWA Domain. Two different genes corresponding to the VWA domain of CMG2 were cloned into a pGEX-4T-1 vector (Amersham Pharmacia) and were expressed as GST fusion proteins in *Escherichia coli*. The proteins represent CMG2 residues S38–C218 (S38) and R40–S217 (R40), respectively. The S38 and R40 proteins were purified by affinity chromatography using GST Sepharose (Amersham Pharmacia) and were then cleaved with thrombin to generate pure CMG2 VWA domains. The proteins were further purified by size exclusion chromatography, were concentrated to 30 mg/ml, and were screened for crystallization by hanging drop vapor diffusion.

Crystallization. Five microliters of S38 was mixed with 5 ml of a reservoir solution containing 15% polyethylene glycol 3350, 150 mM MgCl_2 , and 100 mM [bis(2-hydroxyethyl)amino]tris(hydroxymethyl)methane, pH 6.5, and was allowed to equilibrate at

room temperature. Crystals grew in space group $P4_12_12$ with unit cell dimensions $a = b = 57.4 \text{ \AA}$ and $c = 111.5 \text{ \AA}$ and one molecule per asymmetric unit. Crystals of R40 grew by mixing 2 ml of a 5 mM CaCl_2 , 30 mg/ml solution of protein with a 2-ml reservoir solution of 20% polyethylene glycol 8K, 200 mM Mg acetate, and 100 mM of either Na cacodylate or [bis(2-hydroxyethyl)amino]tris(hydroxymethyl)methane, pH 6.5–7. After mixing and equilibration at 4°C , primitive cubic crystals formed within 4 days. The cell dimensions for the $P2_13$ space group were $a = b = c = 79.7 \text{ \AA}$.

Structure Determination. The crystals were harvested and sequentially soaked for ≈ 1 min each in reservoir solutions containing 5%, 10%, and 15% glycerol (S38) or polyethylene glycol 400 (R40). The crystals were plunged into liquid nitrogen and were transported frozen to beamline F1 at the Cornell High Energy Synchrotron Source and beamline 8.2.1 at the Advanced Light Source. All diffraction images were processed and scaled with the program HKL2000 (30). Efforts to solve the S38 and R40 data sets by molecular replacement using integrin I domain and VWA domain search models were unsuccessful, so heavy-atom derivatization was pursued. R40 crystals were soaked with 10 mM K_2PtCl_4 for 49 h before cryocooling. The presence of two bound platinum ions was confirmed by difference Patterson methods. Phases were obtained by using a three-wavelength multiple anomalous diffraction experiment at beamline 8.2.1 at the Advanced Light Source and were calculated by using the program SOLVE (31) (figure of merit = 0.55 for all reflections). The program RESOLVE (32) was used to calculate an initial map and perform automated model building. The model was completed manually by using the program O (33) and was refined with the program CNS (34). Rigid body refinement, followed by simulated annealing with torsion angle dynamics, and finally by individual B factor refinement resulted in R_{cryst} and R_{free} values of 24.4% and 26.1%, respectively. This model was then refined against the native R40 data. In addition to the Mg^{2+} ion bound at the MIDAS site, Fourier $F_o - F_c$ maps revealed an acetate ion bound to the Mg^{2+} . With a few rounds of manual rebuilding along with the addition of 198 water molecules by using the program ARP/WARP (35), the model was refined in REFMAC (36) to final R_{cryst} and R_{free} values of 15.0% and 18.6%, respectively. The R40 model was used to identify the molecular replacement solution for the S38 data by using the program MOLREP (37). Several cycles of rebuilding and refinement were performed. After the addition of 154 water molecules using the program ARP/WARP and individual anisotropic B factor refinement in REFMAC, the R_{cryst} and R_{free} values were 18.7% and 21.5%, respectively. The stereochemical quality of each model was validated by using the program PROCHECK (38). Statistics for data collection and refinement are presented in Tables 1 and 2.

Results and Discussion

The crystal structure of the CMG2 VWA domain containing residues 40–217 (R40) was solved by phasing with multiple

Table 2. Refinement statistics

	R40	S38
Refined atoms	1,570	1,551
Protein	1,367	1,396
Water	198	154
Ligand	5	1
R_{cryst}^* %	15.0	18.7
$R_{\text{free}}^{\dagger}$ %	18.6	21.5
Average protein B factors, $\text{\AA}^2$	18.8	16.9
rms deviation		
Bond length, $\text{\AA}$	0.018	0.014
Bond angle, $^{\circ}$	1.650	1.488
Bonded B factor, $\text{\AA}^2$		
Main chain	1.143	1.378
Side chain	3.192	2.543
Ramachandran plot, %		
Favored	89.7	91.2
Allowed	8.4	8.2
Generous	1.3	0.6
Disallowed	0.6	0

* $R_{\text{cryst}} = \sum_{hkl} \|F_o - F_c\| / \sum_{hkl} \|F_o\|$.

$\dagger R_{\text{free}} = R_{\text{cryst}}$ using 5% of F_o sequestered before refinement.

anomalous diffraction data from a K_2PtCl_4 -soaked crystal (Table 1). An initial model generated from this experiment was refined against a native data set to generate a 1.8- $\text{\AA}$ structure (Table 2). The R40 structure was used to solve the structure of a second crystal form containing CMG2 residues 38–218 (S38). This S38 model, which contains a disulfide between residues Cys-39 and Cys-218, was refined to 1.5 $\text{\AA}$ (Table 2). Both CMG2 VWA domain structures consist of a six-stranded β -sheet core surrounded by six α -helices (Fig. 1). The positions of secondary structural elements in the CMG2 sequence are shown in Fig. 2. The two CMG2 VWA domain structures overlap with an rms deviation of 0.7 $\text{\AA}$ for 173 α -carbon atoms. The differences lie in the $\alpha 1$ - $\beta 2$ (residues 72–78) and $\beta 6$ - $\alpha 6$ (residues 200–204) loops, which are involved in crystal packing interactions in the R40 structure. The orientation of $\alpha 6$ also varies between the two structures (Fig. 1). The position of this C-terminal helix has been shown to be highly variable among I domains (25) and, in this case, is likely dictated by the presence (S38) or absence (R40) of the disulfide bond and the differences in the preceding $\beta 6$ - $\alpha 6$ loop.

As expected, the protein adopts a dinucleotide binding, or Rossmann, fold with homology to VWA domains and integrin I domains. Least-squares alignments to structures within these families reveal a strong conservation of the central β -sheet, whereas the register and lengths of the α -helices vary. Whereas the rms deviations for these alignments are similar (≈ 1.5 – 2.0 $\text{\AA}$ for all α -carbon atoms), the closest structural homolog to the CMG2 VWA domain is the open conformation of the αM integrin I domain (PDB ID code 1IDO; ref. 23). These molecules share 24% amino acid identity (Fig. 2), a structural rms deviation of 1.1 $\text{\AA}$ for 113 α -carbon atoms (1.5 $\text{\AA}$ overall, Fig. 3*a*), and an identical arrangement of metal coordinating residues at the MIDAS (Fig. 3*f* and *g*). In both CMG2 structures, the side chains of Ser-52, Ser-54, and Thr-118 directly coordinate the Mg^{2+} ion, whereas the side chains of Asp-50 and Asp-148 are each hydrogen-bonded to one of two Mg^{2+} -coordinated water molecules (Figs. 1 and 3*g*). The sixth MIDAS coordination site is occupied by different pseudoligands in each of the two crystal forms. In the S38 structure, a glutamate residue, Glu-194, from a neighboring CMG2 monomer is positioned such that one of its ϵ -oxygen atoms is 2.0 $\text{\AA}$ from the Mg^{2+} ion (Figs. 1 and 3*g*). In the R40 structure, an acetate ion from the crystallization buffer

Fig. 1. Structure of the CMG2 VWA domain. A ribbon diagram of the S38 structure indicates secondary structure elements. Highlighted amino acid residues include the N- and C-terminal cysteines (C39 and C218, respectively) that form a disulfide bond (the sulfur atoms are depicted in yellow) and the conserved amino acids of the MIDAS motif. The Mg^{2+} ion is shown as a large blue sphere with two bound water molecules depicted as beige spheres. The small red spheres correspond to oxygen atoms within the MIDAS amino acids. The E194 residue from a neighboring CMG2 molecule (only E194 is shown in pink) contributes the sixth coordinating residue at the MIDAS metal. The structures of S38 and R40 superimpose with an rms deviation of 0.7 $\text{\AA}$. They differ primarily in the orientation of the $\alpha 6$ C-terminal helix. This helix and its preceding loop in the R40 structure are depicted in blue. This image and Fig. 3 were generated by the program MOLSCRIPT (47) and rendered in RASTER3D (48).

is located in the same location as the S38 structure Glu-194 carboxylate group and directly coordinates the Mg^{2+} ion. The αM integrin I domain also crystallized with a ligand-mimetic, a Glu from a neighboring molecule, to reveal its open conformation (Fig. 3*f* and ref. 23). The fact that the CMG2 VWA domain structures contain an identical coordination at the MIDAS indicates that they also represent open conformations (Fig. 3*f* and *g*).

Numerous structures of open and closed α -integrin I domains have indicated conserved features that link changes in the MIDAS to structural changes throughout the domain (25–27). In the I domain closed conformation, the metal ion of the MIDAS is coordinated by three waters, two MIDAS motif serines, and the remote MIDAS motif Asp residue (DXSXS...T...D; Fig. 3*e*). In the conversion to the open high-affinity state, a conserved Ile from the C-terminal helix (Ile-316 of αM) is removed from its hydrophobic pocket (pocket II, Fig. 3*a* and *b*), and the C-terminal helix moves away from the MIDAS face by 10 $\text{\AA}$. This movement is linked to a shift in the loop preceding this helix, $\beta 6$ - $\alpha 7$ (Fig. 3*a*), and removal of a conserved phenylalanine (Phe-302) from a nearby hydrophobic pocket (pocket I, Fig. 3*a*–*c*). Changes in this pocket are coupled to movement of a second phenylalanine residue (Phe-275) into a solvent-exposed region and a shift and backbone flip for a conserved Gly (Gly-243; Fig. 3*b* and *c*).

Fig. 2. Sequence alignment of the VWA domains from CMG2, ATR/TEM8, and the α M integrin. The sequences were aligned with CLUSTALW (49) and were displayed in ESPRIT (50) along with the secondary structure assignments for CMG2 (Top) and the open conformation of the α M integrin (PDB ID code 1IDO, Bottom; ref. 23). η , 3_{10} -helix. White lettering boxed with a red background indicates residues that are conserved in all three sequences, and red lettering indicates similar residues. The numbering (Top) corresponds only to the CMG2 sequence.

The net effect is rearrangement of residues in the loops and strands surrounding the MIDAS. The loop containing the DxSxS sequence shifts closer to the threonine-containing loop so that the metal ion can bridge both the MIDAS serines and the threonine (DXSXS...T...D; Fig. 3*e* and *f*). The remote MIDAS motif Asp residue is shifted such that it is no longer able to directly coordinate the metal and instead coordinates indirectly through a water molecule (Fig. 3*e* and *f*). This shift is thought to increase the electrophilicity of the metal and increase the affinity for ligand (26).

Currently, there is no evidence that the CMG2 VWA domain can adopt a closed conformation. It is therefore interesting to note that many of the key I domain residues implicated in the structural transformation between open and closed conformations are conserved in CMG2. Specifically, Gly-149 of CMG2 aligns with Gly-243 of α M (Fig. 2) and adopts the same backbone conformation as in the open conformation (data not shown). CMG2 Phe-203 aligns with α M Phe-302 and is located in a solvent-exposed region of the β 6- α 6 loop rather than being buried within hydrophobic pocket I (Figs. 2 and 3*d*). Whereas α M Phe-275 aligns with CMG2 Glu-182 in sequence, CMG2 Phe-181 partially overlaps α M Phe-275 in their aligned structures (Figs. 2 and 3*c* and *d*). The exception to this conservation is the α M Ile-316 residue that packs into hydrophobic pocket II in the closed conformation (Fig. 3*b*). This isoleucine is conserved in all α -integrin I domains and its mutation has been shown to shift α M into a open conformation (39). However, this residue aligns to Ser-217 in CMG2 (Fig. 2). The fact that this residue is polar and that there are no other residues in this vicinity that might bind in hydrophobic pocket II suggests that the C-terminal helix position associated with closed conformations may not be stable. Furthermore, in both of the CMG2 open conformation structures observed here, hydrophobic pocket II is occupied by an isoleucine, Ile-213, from a different part of the C-terminal helix (Fig. 3*d*, R40 structure not shown). Stabilization of the C-terminal helix position through this hydrophobic pocket may increase the probability of the CMG2 VWA domain existing in an open conformation.

Researchers have been able to lock I domains into low-, intermediate-, or high-affinity conformations by engineering disulfides between the C-terminal helix and the N terminus of

the protein (25, 29, 40). This strategy works because the C-terminal helix and its preceding loop act as an allosteric switch for the structural rearrangements at the MIDAS (39). We were curious as to whether the natural disulfide bond between the N and C termini of the CMG2 VWA domain might play a role in regulating the affinity for ligand or open and closed domain conformations. Whereas our structures with (S38) and without (R40) the disulfide in CMG2 revealed slightly different orientations for the C-terminal helix, the fact that both proteins crystallized in the open conformation suggests that the conformation is independent of the disulfide bond. Further, the disulfide bond does not seem to play a significant role in PA binding, because both CMG2 S38 and R40 proteins bind PA with subnanomolar affinity (51).

The extremely tight binding affinity between the CMG2 VWA domain and PA suggests that their interaction will differ from those observed between I domains and their physiological ligands. Precedent for this exists even within α -integrin I domains, which have been exploited by a number of microbial pathogens as a point of interaction with the cell. For instance, binding of echovirus-1 to integrin α 2 is not MIDAS metal-dependent (41, 42). In the case of the hookworm neutrophil inhibitory factor binding to integrin α M, the MIDAS metal is needed, but binding does not require the canonical open conformation (39, 43–45). Mutational analysis of ATR/TEM8 suggests that preservation of the MIDAS motif is required for PA interaction (28). Binding also depends on the Asp-683 residue of PA, suggesting that its carboxylate group completes the MIDAS metal coordination in the open conformation of ATR/TEM8 (28, 46). It is expected that CMG2 will interact similarly with PA.

The high affinity of CMG2 for PA implies that there are additional VWA domain contacts beyond the MIDAS residues. Whereas ATR/TEM8 and CMG2 share high sequence identity, mapping the ATR/TEM8 sequence onto the surface of CMG2 did not indicate an obvious conserved patch that might represent a common PA-binding motif. An analysis of surface exposed residues on the face surrounding the MIDAS reveals predominantly polar residues: Asn-57, Glu-117, Lys-150, Asp-152, Asp-180, and Glu-182. These residues are also polar in ATR/TEM8 (Fig. 2). The only hydrophobic residues within this region, Val-115 and Leu-179, correspond to Gly-115 and Lys-181 in

Fig. 3. The CMG2 VWA domain is in an open conformation. (a) The backbone structure of the CMG2 VWA domain (light green) was superimposed onto the aligned structures of the α M integrin I domains in their open (dark green, PDB ID code 1IDO) and closed (blue, PDB ID code 1JLM) conformations (23, 26). The hydrophobic pockets I and II are indicated by gray ovals and the Mg^{2+} ion for CMG2 is depicted as a blue sphere. (b) The closed conformation of α M with Phe-302 buried in hydrophobic pocket I and Ile-316 buried in hydrophobic pocket II. (c) The open conformation of α M shows a shift in the C-terminal helix from its position in *b* such that Phe-302 becomes solvent-exposed, and hydrophobic pocket II is now occupied by Leu-312. The positions of Phe-275 and Gly-243 have also shifted. (d) The structure of the CMG2 VWA domain is similar to that of *c* and is therefore an open conformation. It is hypothesized that the presence of Ile-213 bound in hydrophobic pocket II and the absence of a downstream hydrophobic residue equivalent to α M Ile-316 might help stabilize the open conformation. Residues that, when mutated, result in ISH and JHF disease (Leu-45, Gly-105, Ile-189, and Cys-218) are depicted in yellow. (e) In the closed structure of α M, the Mn^{2+} ion (blue sphere) is coordinated by three waters, two MIDAS serines, and an aspartic acid. The bond to the MIDAS threonine has been broken. (f) In the open structure of α M, the Mg^{2+} ion is coordinated directly by two serines, two waters (medium red spheres), a threonine, and a glutamate from a neighboring monomer. (g) The coordination of the MIDAS metal in the CMG2 VWA domain structure is identical to the coordination observed for the open conformation of α M shown in *f*.

ATR/TEM8, and are therefore not conserved. Further work is needed to determine the means by which these molecules bind PA and if their specificities differ.

Recently, two groups have reported genotype-phenotype analyses that link mutations in CMG2 with JHF and ISH (20, 21). Four point mutations within the VWA domain have been identified: L45P, G105D, I189T, and C218R. Leu-45 is a highly conserved core residue located within β 1 and hydrophobic pocket II (Fig. 3*d*). The unique structural properties of proline and the rarity of this residue within β -strands suggest that this mutation will be destabilizing. Glycines also represent a unique structural niche and are rarely found within β -strands and α -helices; however, Gly-105 is located in the middle of α 2 and is conserved in ATR/TEM8 (Figs. 2 and 3*d*). Reasonable rotamers of an Asp residue side chain could not be modeled into this site without generating steric clash. Ile-189 is located at the end of α 5, in a buried and hydrophobic region of the structure (Fig. 3*d*). Mutation to Thr is likely to create a destabilizing cavity within this region. The location and nature

of these mutations suggest that all will have destabilizing effects on the domain structure rather than a direct effect on ligand binding. The final mutant, C218R, disrupts the natural disulfide that links the N and C termini of the domain (Figs. 1 and 3*d*). Whereas this disulfide does not seem to be required for obtaining the open conformation or high-affinity binding to PA (51), it may play a role in binding physiological ligand. It is also possible that the C218R mutation could be destabilizing within the context of full-length CMG2.

Our studies reveal two structures of the CMG2 VWA domain in an open conformation with carboxylate ligand-mimetics completing the coordination at the MIDAS. Further work is needed to understand the physiological role of CMG2, whether it has a high affinity for its natural ligands, and the role CMG2 defects may play in ISH, JHF, and other human diseases. Understanding the role of CMG2 in binding and internalizing the anthrax toxin will be an important part of understanding its mechanism of intoxication. There is also potential that the pM affinity of soluble CMG2 VWA domain for PA and its ability to inhibit

toxicity in cell culture (16) can be exploited to develop new therapeutics for anthrax.

We thank Bryan Krantz, Jonah Rainey, and Ben Spiller for critical review of the manuscript; Lauren Greene for assistance with protein

expression and purification; and Steve Harrison, Piotr Sliz, and the staff of the Cornell High Energy Synchrotron Source beamline F1 and Advanced Light Source beamline 8.2.1 for their assistance with x-ray crystallography. This work was supported by National Institutes of Health Grants AI022021, AI048489, and AI056013.

- Collier, R. J. & Young, J. A. (2003) *Annu. Rev. Cell Dev. Biol.* **19**, 45–70.
- Leppla, S. H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3162–3166.
- O'Brien, J., Friedlander, A., Dreier, T., Ezzell, J. & Leppla, S. (1985) *Infect. Immun.* **47**, 306–310.
- Duesbery, N. S., Webb, C. P., Leppla, S. H., Gordon, V. M., Klimpel, K. R., Copeland, T. D., Ahn, N. G., Oskarsson, M. K., Fukasawa, K., Paull, K. D., *et al.* (1998) *Science* **280**, 734–737.
- Vitale, G., Pellizzari, R., Recchi, C., Napolitani, G., Mock, M. & Montecucco, C. (1998) *Biochem. Biophys. Res. Commun.* **248**, 706–711.
- Agrawal, A., Lingappa, J., Leppla, S. H., Agrawal, S., Jabbar, A., Quinn, C. & Pulendran, B. (2003) *Nature* **424**, 329–334.
- Park, J. M., Greten, F. R., Li, Z. W. & Karin, M. (2002) *Science* **297**, 2048–2051.
- Kirby, J. E. (2004) *Infect. Immun.* **72**, 430–439.
- Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R. & Thomas, G. (1992) *J. Biol. Chem.* **267**, 16396–16402.
- Milne, J. C., Furlong, D., Hanna, P. C., Wall, J. S. & Collier, R. J. (1994) *J. Biol. Chem.* **269**, 20607–20612.
- Mogridge, J., Cunningham, K., Lacy, D. B., Mourez, M. & Collier, R. J. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 7045–7048.
- Friedlander, A. M. (1986) *J. Biol. Chem.* **261**, 7123–7126.
- Abrami, L., Liu, S., Cosson, P., Leppla, S. H. & van der Goot, F. G. (2003) *J. Cell Biol.* **160**, 321–328.
- Blaustein, R. O., Koehler, T. M., Collier, R. J. & Finkelstein, A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2209–2213.
- Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. (2001) *Nature* **414**, 225–229.
- Scobie, H. M., Rainey, G. J., Bradley, K. A. & Young, J. A. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 5170–5174.
- St. Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K. E., Montgomery, E., Lal, A., Riggins, G. J., Lengauer, C., Vogelstein, B., *et al.* (2000) *Science* **289**, 1197–1202.
- Carson-Walter, E. B., Watkins, D. N., Nanda, A., Vogelstein, B., Kinzler, K. W. & St. Croix, B. (2001) *Cancer Res.* **61**, 6649–6655.
- Bell, S. E., Mavila, A., Salazar, R., Bayless, K. J., Kanagala, S., Maxwell, S. A. & Davis, G. E. (2001) *J. Cell Sci.* **114**, 2755–2773.
- Dowling, O., Difeo, A., Ramirez, M. C., Tukul, T., Narla, G., Bonafe, L., Kayserili, H., Yuksel-Apak, M., Paller, A. S., Norton, K., *et al.* (2003) *Am. J. Hum. Genet.* **73**, 957–966.
- Hanks, S., Adams, S., Douglas, J., Arbour, L., Atherton, D. J., Balci, S., Bode, H., Campbell, M. E., Feingold, M., Keser, G., *et al.* (2003) *Am. J. Hum. Genet.* **73**, 791–800.
- Whittaker, C. A. & Hynes, R. O. (2002) *Mol. Biol. Cell* **13**, 3369–3387.
- Lee, J. O., Rieu, P., Arnaout, M. A. & Liddington, R. (1995) *Cell* **80**, 631–638.
- Shimaoka, M., Takagi, J. & Springer, T. A. (2002) *Annu. Rev. Biophys. Biomol. Struct.* **31**, 485–516.
- Shimaoka, M., Xiao, T., Liu, J. H., Yang, Y., Dong, Y., Jun, C. D., McCormack, A., Zhang, R., Joachimiak, A., Takagi, J., *et al.* (2003) *Cell* **112**, 99–111.
- Lee, J. O., Bankston, L. A., Arnaout, M. A. & Liddington, R. C. (1995) *Structure (London)* **3**, 1333–1340.
- Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J. & Liddington, R. C. (2000) *Cell* **101**, 47–56.
- Bradley, K. A., Mogridge, J., Rainey, G. J., Batty, S. & Young, J. A. (2003) *J. Biol. Chem.* **278**, 49342–49347.
- McCleverty, C. J. & Liddington, R. C. (2003) *Biochem. J.* **372**, 121–127.
- Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276**, 307–326.
- Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55**, 849–861.
- Terwilliger, T. C. (1999) *Acta Crystallogr. D* **55**, 1863–1871.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47**, 110–119.
- Brunger, 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., *et al.* (1998) *Acta Crystallogr. D* **54**, 905–921.
- Perrakis, A., Morris, R. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463.
- Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999) *Acta Crystallogr. D* **55**, 247–255.
- Vagin, A. & Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025.
- Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993) *J. Mol. Biol.* **231**, 1049–1067.
- Xiong, J. P., Li, R., Essafi, M., Stehle, T. & Arnaout, M. A. (2000) *J. Biol. Chem.* **275**, 38762–38767.
- Shimaoka, M., Lu, C., Palframan, R. T., von Andrian, U. H., McCormack, A., Takagi, J. & Springer, T. A. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 6009–6014.
- King, S. L., Kamata, T., Cunningham, J. A., Emsley, J., Liddington, R. C., Takada, Y. & Bergelson, J. M. (1997) *J. Biol. Chem.* **272**, 28518–28522.
- Xing, L., Huhtala, M., Pietiainen, V., Kapyla, J., Vuorinen, K., Marjomaki, V., Heino, J., Johnson, M. S., Hyypia, T. & Cheng, R. H. (2003) *J. Biol. Chem.* **279**, 11632–11638.
- Li, R., Rieu, P., Griffith, D. L., Scott, D. & Arnaout, M. A. (1998) *J. Cell Biol.* **143**, 1523–1534.
- Rieu, P., Ueda, T., Haruta, I., Sharma, C. P. & Arnaout, M. A. (1994) *J. Cell Biol.* **127**, 2081–2091.
- Muchowski, P. J., Zhang, L., Chang, E. R., Soule, H. R., Plow, E. F. & Moyle, M. (1994) *J. Biol. Chem.* **269**, 26419–26423.
- Rosovitz, M. J., Schuck, P., Varughese, M., Chopra, A. P., Mehra, V., Singh, Y., McGinnis, L. M. & Leppla, S. H. (2003) *J. Biol. Chem.* **278**, 30936–30944.
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
- Merritt, E. A. & Bacon, D. J. (1997) *Methods Enzymol.* **277**, 505–524.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680.
- Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999) *Bioinformatics* **15**, 305–308.
- Wigelsworth, D. J., Krantz, B. A., Christensen, K. A., Lacy, D. B., Juris, S. J. & Collier, R. J. (2004) *J. Biol. Chem.*, in press.