Membrane channel structure of *Helicobacter pylori* vacuolating toxin: Role of multiple GXXXG motifs in cylindrical channels

Sanguk Kim, Aaron K. Chamberlain, and James U. Bowie*

Department of Chemistry and Biochemistry, University of California Los Angeles–Department of Energy Center for Genomics and Proteomics, Molecular Biology Institute, Boyer Hall, University of California, Los Angeles, CA 90095-1570

Edited by William F. DeGrado, University of Pennsylvania School of Medicine, Philadelphia, PA, and approved March 1, 2004 (received for review December 29, 2003)

***Helicobacter pylori* is a human pathogen responsible for severe gastric diseases such as peptic ulcers, gastric adenocarcinoma, and gastric lymphoma. Vacuolating toxin (VacA) is crucial in facilitating the colonization of the gastric lining by inducing cell apoptosis and immune suppression. VacA inserts into membranes and forms a hexameric, anion-selective pore. Here we present a structural model of the VacA pore that strongly resembles the structure of an unrelated anion-selective channel, MscS. In our model, Gly residues in GXXXG motifs pack against small Ala or Val side chains to generate the pore. Our model suggests that the same design of two anion-selective channels was achieved by two different evolutionary paths and provides insight into the mechanism of VacA function.**

MscS | membrane protein | structure prediction | anion channel

One of the most successful human pathogens is *Helicobacter pylori*, chronically infecting >50% of the human population, and it is responsible for severe gastric diseases such as peptic ulcers, gastric adenocarcinoma, and gastric lymphoma (1–4). Gastric cancer is the second leading cause of total cancer deaths in the United States (Harvard Center for Cancer Prevention, www.hsph.harvard.edu/cancer).

Vacuolating toxin (VacA) is a major virulence factor secreted by *H. pylori* and is a key component in the genesis of gastric diseases (5–8). The most extensively studied activity of VacA is cellular vacuolation in mammalian cells (6, 9, 10). Although the precise mechanism of VacA-induced vacuole formation is not completely understood, it involves the binding and internalization of the toxin. It has been proposed that vacuolation is a consequence of anion-selective channel formation in late endosomal compartments (11–14). In addition to its vacuole formation activity, VacA causes a series of cellular events, including depolarization of the membrane potential (13, 15), cell apoptosis (16–19), interference with epithelial cell attachment (20), and inhibition of T lymphocyte activation (21).

The secreted form of VacA is a water-soluble, 88-kDa protein that assembles on membranes to form a hexameric anion-selective pore (12, 13). VacA has a hydrophobic amino-terminal segment, VacA-transmembrane (TM) (residues 1–32), that is essential for pore formation and is thought to traverse the membrane as an α-helix. VacA-TM contains six glycine residues and three GXXXG sequence motifs, which are commonly found to mediate TM helix oligomerization (22). Cover and coworkers (23) showed that Gly-14 and Gly-18 are essential for membrane channel formation and VacA cytotoxicity. Thus, the pore assembly likely involves the association of VacA-TM helices via packing of the GXXXG elements. Structural and functional studies of VacA have been difficult, however, because of the lack of efficient expression systems and the toxicity of *H. pylori*, a type I carcinogen (24).

Here we built a structural model of the hexameric VacA channel by using a prediction method we have developed for modeling transmembrane helical bundles (25). In our model, the packing of Gly residues in the GXXXG motifs against small Ala or Val side chains generates a pore structure. The structure explains the mutagenesis results and is supported by the fact that a remarkably similar design has already been seen in nature. The pore structure bears a striking resemblance to the structure of another anion-selective channel, the mechanosensitive channel of small conductance, MscS (26). VacA and MscS control osmotic imbalances across the membrane, even though the proteins have no sequence or structural similarities outside of their pore motifs. Thus, nature appears to have developed a similar design for the two channels, by different evolutionary pathways. Our model explains how VacA-TM helices can create an anion-selective channel and reveals its structural similarities to MscS.

Methods

Modeling the VacA-TM and MscS TM3 Helix Dimeric Structures. The TM domain sequence of VacA from *H. pylori* and the TM3 sequence of MscS from *Escherichia coli* (see Fig. 2) were built into uniform α-helices having backbone torsion angles of $\phi = -65°$ and $\psi = -40°$ (27), using the INSIGHTII biopolymer package (Accelrys, San Diego). We used the backbone-dependent rotamer library program SCWRL (28) to choose the side chain rotamers. The structure prediction procedures and parameters used in the Monte Carlo (MC) searches are described (25). The dimer optimization started with 400 pairs of helices placed in random orientations with respect to each other. The packing interaction of each helix pair was then optimized by using an MC minimization method. The internal backbone and side chain positions were kept fixed during the minimizations, but the relative positions of the helices were given all six degrees of freedom. The energy function used is a softened van der Waals potential without any electrostatic component. The simulations were stopped after a maximum of 100,000 MC steps or, if 15,000 MC steps occurred, without moving to a lower energy. All six orientation parameters were changed during a step. A step of the same magnitude and direction was repeated if the previous step resulted in lower energy. The step size in each parameter was randomly selected. The temperature began at 500 K and decreased linearly to 0.1 K over the first 50,000 steps. In this manner, we generated 400 independently optimized dimeric structures.

The homodimeric structures were filtered to remove the

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: VacA, vacuolating toxin; TM, transmembrane; MC, Monte Carlo; rmsd, rms deviation.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1SEW).

*To whom correspondence should be addressed. E-mail: bowie@mbi.ucla.edu.

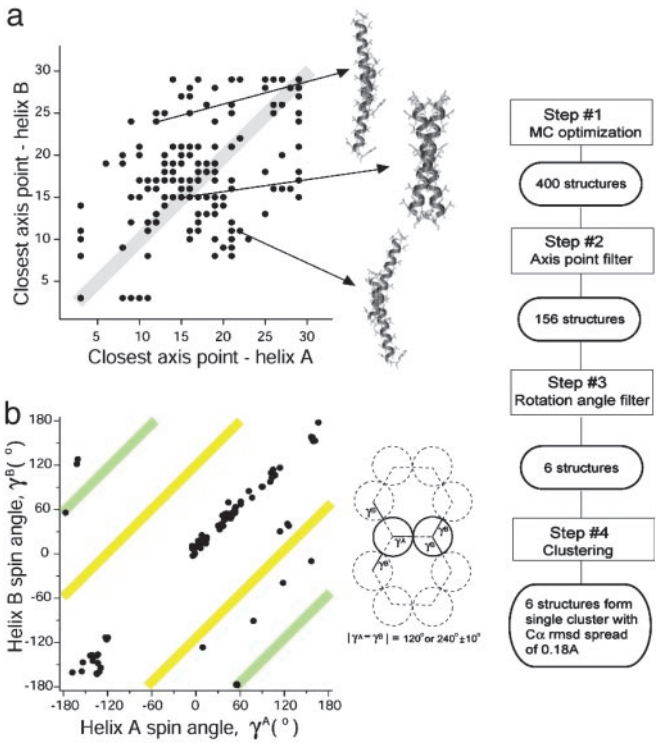

Fig. 1. Summary of the VacA-TM modeling. (*a*) The axis point filter removed the structures incompatible with lipid bilayer confinement. The diagonal strip indicates the retained structures, in which the axis points of closest approach are at a similar distance from the N terminus of each helix. (*b*) The rotational angle filter retained the structures consistent with hexameric symmetry. The shaded diagonal regions represent 120° (yellow strip) and 240° (green strip) spin angles.

structures incompatible with the bilayer constraints and known symmetry (Fig. 1). We represent the helix axis as a set of axis points, one per C^{α} atom, which are numbered from the N terminus. In each helix dimer, we found the axis point in helix A and axis point in helix B that are closest to each other. We discarded dimers in which the closest axis points were numbered more than two away from each other, because these structures have one helix significantly shifted out of the membrane. For the hexameric VacA simulations, we selected the helix pairs containing a difference in their angles of rotation about the two helical axes of 120° ± 10° or 240° ± 10° (only structures near 240° were obtained). For the heptameric MscS simulations, the rotation angles of 129° ± 10° or 231° ± 10° were used. We clustered the remaining structures by C^{α} rms distances by using NMRCLUSTER (29). The median model from the largest cluster is our final dimer structure. From the 400 MC optimized structures of the VacA-TM domain, six were consistent with the hexameric symmetry. These structures formed a single cluster and are essentially identical having a C^{α} rms deviation (rmsd) spread of 0.18 Å. The 400 MC optimizations of the MscS TM3 domains resulted in two structures consistent with heptameric symmetry. These dimers had a C^{α} rmsd of only 0.27 Å to each other and 1.8 Å to the TM3 pair of MscS (1MXM) (26). The rmsd to the TM3 pair including side chain atoms was 2.1 Å. So our method successfully predicted the packing of MscS TM3 without the use of any experimental constraints beyond the known symmetry.

Building the VacA Hexameric Structure. Our use of a softened van der Waals potential created steric overlaps, which were relieved by energy minimization with the crystallography and NMR system (CNS V.1.1) (30). The rmsd between the starting dimer and minimized dimer structure was 2.2 Å on C^{α} positions. We then duplicated the structure of the helix pairs by overlaying one helix of the dimer with the other four times. This duplication created a 6-fold rotational symmetry axis, which represents the center of the pore. We calculated the positions of the pore axis by averaging the positions of the equivalent C^{α} atoms in the six subunits. Optimization of the hexamers was performed *in vacuo* by using CNS. The interhelical C^{α}–C^{α} distances <7.5 Å between the modeled helix pairs were used as restraints for the hexamer structure with additional restraints to maintain the helical backbone. We energy minimized the structures with three rounds of 200 steps by using a dielectric constant of 1.0 and the nonbond cutoff of 13 Å. The rmsd between the starting and final hexamer models was 3.3 Å on C^{α} positions.

Results and Discussion

Model of the VacA Pore. We modeled the hexameric assembly of VacA-TM by using our recently developed algorithm that has proven effective for predicting homooligomeric TM helix structures (25). This method successfully built the TM helix dimer structure of glycophorin A, the tetramer structure of the influenza virus M2 proton channel, and the pentamer structure of phospholamban (25). As discussed below, the algorithm also accurately models the pore structure of the MscS channel. Thus,

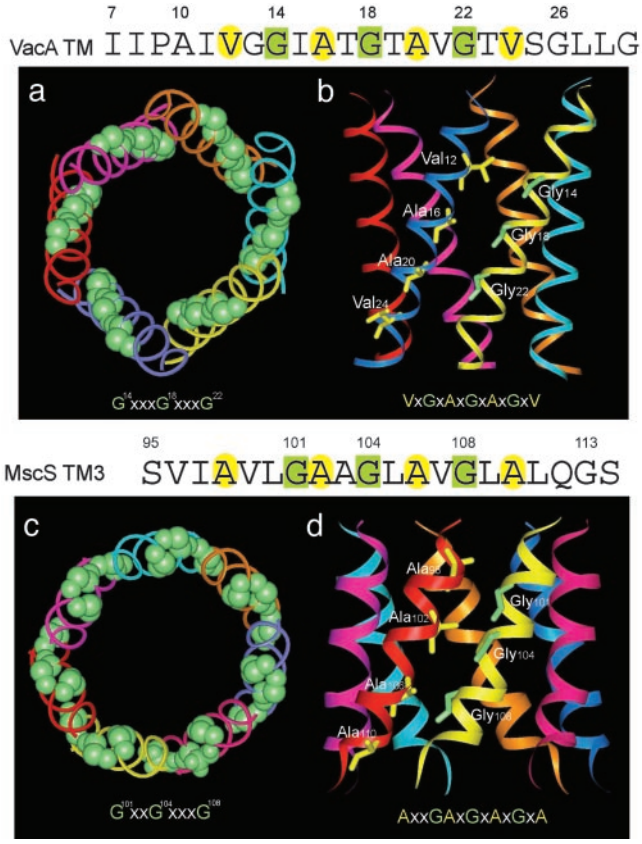

Fig. 2. The VacA-TM channel model and the structure of the MscS channel (TM3). In both structures, the helix–helix packing interactions are mediated by a series of glycines (yellow) to alanine/valine (green) contacts. (*a*) Top view of VacA-TM channel. Gly-14, Gly-18, and Gly-22 residues are shown in a space-filling representation. (*b*) Side view of the VacA-TM channel. The VacA-TM helix–helix packing interactions are mediated by the sequence V^{12}xGxAxGxAxGxV^{24}. (*c*) Top view of the MscS TM3 structure. Gly-101, Gly-104, and Gly-108 residues are shown in a space-filling representation. (*d*) Side view of the MscS TM3 channel structure. The MscS TM3 helix–helix packing interactions are mediated by the sequence A^{98}xxGAxGxAxGxA^{110}.

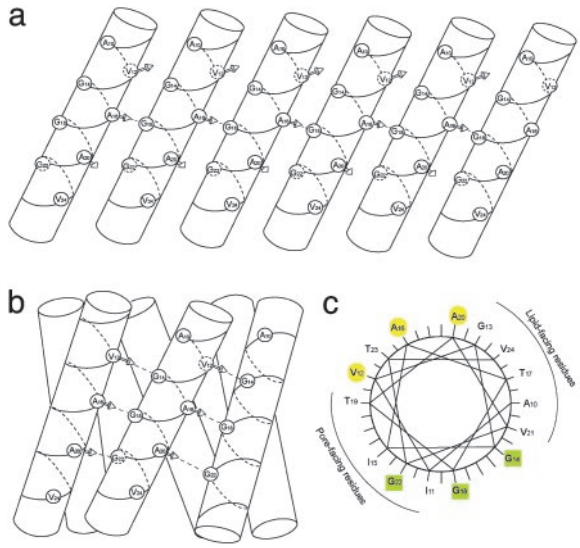

Fig. 3. Helix–helix packing interaction through multiple GXXXG motifs. (*a*) A series of rod representations of VacA-TM channel, unrolled into a planar representation. The small amino acid residues in the hexameric helix–helix contacts are shown as balls. The arrows indicate the side chain atom directions. The van der Waals contacts in between helices are shown as dashed lines. (*b*) The tightly packed α-helices forms a cylindrical hexameric channel structure. In a twisted helical bundle structure, the series of glycines can pack against alanine or valine residues in adjacent helices. (*c*) A helical wheel representation shows the pore-lining, helix–helix packing, and lipid-facing residues in the VacA-TM channel model.

as long as the oligomeric form is known, the method seems applicable to a wide variety of helix packing arrangements involving both large and small interfacial residues, although any model must be independently validated experimentally.

The modeling procedure and results are summarized in Fig. 1. Briefly, the algorithm involves the following steps. (*i*) Starting with VacA-TM helix pairs in random positions, a collection of 400 well packed helix pairs were generated by MC minimization. (*ii*) Structures lacking the appropriate symmetry were eliminated. (*iii*) The remaining structures were clustered, and a representative of the largest cluster was selected. (*iv*) The hexameric channel was built from the best helix pair packing arrangement. The selected helix pair structure was replicated around a pore axis, and steric overlaps were relieved by energy minimization.

Our model of the homo-hexameric VacA-TM channel is shown in Fig. 2*a*. The channel structure reveals a right-handed helix-crossing angle of $-25°$, which is similar to that observed in other channel structures (31). The packing arrangement is illustrated schematically in Fig. 3 and consists of a series of glycine–alanine and glycine–valine contacts between neighboring helices. Regular packing of the helices around the pore is achieved by contacts made between Val-12–Gly-14, Ala-16–Gly-18, and Ala-20–Gly-22.

Of the six Gly residues in VacA-TM, mutagenesis results indicate that Gly-14, Gly-18, and, to a lesser extent, Gly-22 are the most important residues for VacA-TM oligomerization for channel-forming activity and for cell-vacuole formation (23). As illustrated in Fig. 3*b*, Gly-14 and Gly-18 are the most closely packed positions in the structure, and Gly-22 is also in the packing interface. Thus, the model is completely consistent with the mutagenesis results.

VacA-TM Packing Is Similar to the Packing of the MscS Channel. Our proposed structure of VacA-TM is also supported by its similarity to another anion-selective channel of known structure, MscS (26). MscS is a heptameric, integral membrane protein that contains three transmembrane helices per monomer. There is no sequence similarity between MscS and VacA, and before our VacA modeling, there was no reason to suspect a structural similarity. We found, however, that the packing of the pore-lining helix of MscS (TM3) bears a striking resemblance to the packing seen in our VacA-TM model (Fig. 2). In MscS, a face of Gly residues packs against a face of small (all Ala) residues (32), which is similar to VacA, where the Gly residues pack against two Ala and one Val residue. One important difference between the structures is that the helices are in opposite orientations: the N termini of the VacA helices are in the cytoplasm, whereas in MscS, the C termini are in the cytoplasm (Fig. 4).

Because of the apparent similarity between VacA and MscS, we felt that MscS TM3 would provide an ideal test of our application of the modeling algorithm to VacA-TM. We therefore modeled MscS TM3 by using the same procedure used for VacA. The final model for MscS was found to be within 1.8 Å C^{α} rmsd of the crystal structure (the average rmsd of all of the structures before filtering was 7.0 Å) (26). This result lends further credence to our modeling procedure and to our VacA-TM model.

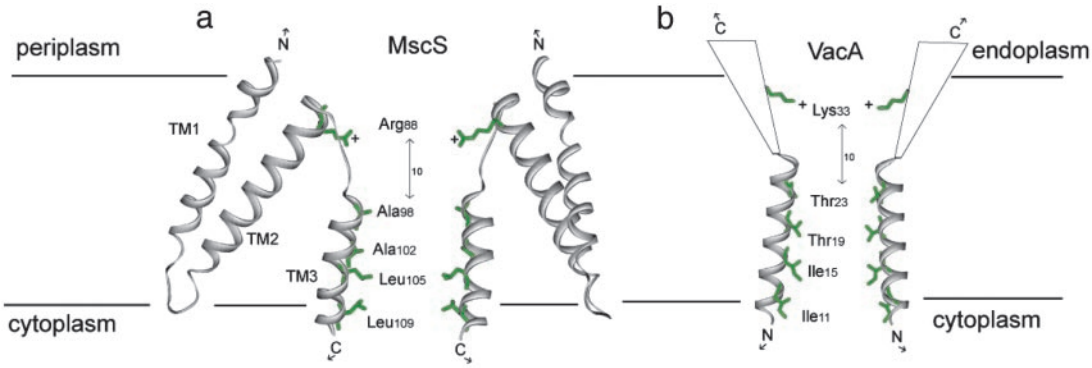

Fig. 4. Transmembrane channel models of MscS and VacA. MscS and VacA channels have similar pore-lining amino acids. (*a*) Structure of the MscS TM region (subunits A and E). The residues in the channel lining (Leu-109, Leu-105, Ala-102, and Ala-98) and an important residue in anion selectivity (Arg-88) are depicted in stick representations. The consensus amino acids in the MscS sequences have Ile, Thr, and Thr at positions 105, 102, and 98 (26, 32). (*b*) Model of VacA channel. Residues in ion permeation pathway (Ile-11, Ile-15, Thr-19, and Thr-23) and a residue potentially responsible for anion selectivity (Lys-33) are shown. The side chains of Leu-109 and Leu-105 in MscS and Ile-11 and Ile-15 in VacA create constrictions in the channels. The well conserved positively charged residues in both channels, Arg-88 of MscS and Lys-33 of VacA, are located 10 residues toward the helix N terminus and C terminus, respectively.

Pore-Lining Residues and Anion Selectivity. Besides the helix-packing residues in VacA and MscS, the pore-lining residues of the two proteins are also similar (Fig. 4). In our VacA-TM model, the key pore-lining residues are Ile-11, Ile-15, Thr-19, and Thr-23. From the cytosolic side of the channel, the apolar side chains of Ile-11 and Ile-15 line the channel and form the narrowest constriction of the pore followed by Thr-19 and Thr-23. In the MscS channel, the first two pore residues on the cytoplasmic side of MscS are also large aliphatic side chains (Leu-109 and Leu-105), which are followed by two small residues (Ala-102 and Ala-98). In the homologs of *E. coli* MscS, Leu-109 is substituted by Ile, and both Ala-102 and Ala 98 are replaced with Thr (26, 33). Thus, residues in the ion conductance pathway of *E. coli* MscS are nearly identical to those found in VacA. The diameter of the permeation pathways are ≈11Å in MscS and ≈8Å in the VacA model. Both VacA and MscS channels function as low-conductance ion channels (14, 26). The large hydrophobic residues at the cytoplasmic mouths of the channels may create a barrier limiting the rate of ion flow.

The close relationship between the VacA and MscS structures allows us to propose a possible mechanism of anion selectivity. In MscS, Arg-88 is thought to be important for the anion preference and is located near the mouth of the pore (Fig. 4). Likewise, the VacA sequence also contains the first positively charged residue at Lys-33. Both Arg-88 of MscS and Lys-33 of VacA are spaced 10 residues away from the last pore residue at the noncytoplasmic side of the channel helix. The spacings are equivalent despite the positive charge being C-terminal of the pore in VacA and N-terminal of the pore in MscS. We therefore suggest that Lys-33 may be playing a key role in anion selectivity in the VacA pore.

The two channel structures may also be similar on the cytoplasmic side of the pore. VacA has a Pro residue just beyond the cytoplasmic end of our Vac-TM helix model. Mutation of this residue (Pro-9) to Ala disrupts the channel formation and cell-vacuolating activity (23). In MscS, a severe kink occurs just on the cytoplasmic side of TM3. VacA-TM may be similarly kinked at Pro-9, and this structural feature may be disrupted by the mutation to Ala.

Conclusion

Besides its cellular vacuolation function, VacA localizes to the mitochondrial membrane and induces cytochrome *c* release into the cytosol during cell apoptosis (17). The mechanism of cytochrome *c* release from mitochondria by VacA is still unclear. The mitochondrial membrane is different from the eukaryotic cell membrane and is more closely related to the prokaryotic cell membrane. Considering the structural resemblance between VacA and the prokaryotic MscS, VacA could perturb the osmotic balance across the mitochondrial membrane. The osmotic shock could disrupt the integrity of the membrane causing cytochrome *c* release. Cellular vacuolation and cytochrome *c* release have independent mechanisms (16), but the two events are mainly caused by VacA channel formation.

It is unlikely that the close resemblance between MscS and our model of VacA occurs because their sequences diverged from a common ancestor. There is no sequence similarity between the two proteins; the global structure of the two proteins is different, one being largely soluble and the other being a constitutive integral membrane protein; and, most significantly, the helix orientations are reversed with respect to one another. Therefore, we propose that the structural similarities arose because these structural features are important for the channel function or regulation. The opposing helix orientations may be tolerated because the packing residues in VacA form a pseudosymmetric packing motif, VxGxAxGxAxGxV. In MscS, the small residues in the helix-packing interface could facilitate helix movement during gating by providing a smooth sliding surface (34). VacA could be gated by the same mechanism. Because both proteins are involved in altering osmotic pressure between compartments, there may be yet more similarities between MscS and VacA.

We thank Sehat Nauli, Chongwoo Kim, and Hoang Tran for helpful comments on the manuscript. This work was supported by National Institutes of Health Grant RO1 GM063919.

1. Montecucco, C. & Rappuoli, R. (2001) *Nat. Rev. Mol. Cell Biol.* **2,** 457–466.
2. Covacci, A., Telford, J. L., Del Giudice, G., Parsonnet, J. & Rappuoli, R. (1999) *Science* **284,** 1328–1333.
3. Suerbaum, S. & Mitchetti, P. (2003) *N. Engl. J. Med.* **347,** 1175–1186.
4. Parsonnet, J. (1995) *Aliment. Pharmacol. Ther. Suppl.* **2,** 45–51.
5. Marshall, B. J. & Warren, J. R. (1984) *Lancet* **1,** 1311–1315.
6. Cover, T. L. & Blaser, M. J. (1992) *J. Biol. Chem.* **267,** 10570–10575.
7. Blaser, M. J. (1993) *Trends Microbiol.* **1,** 255–259.
8. Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A., Warnke, A., Jellum, E., Orentreich, N., Vogelman, J. & Friedman, G. (1994) *N. Engl. J. Med.* **330,** 1267–1271.
9. Catrenich, C. E. & Chestnut, M. H. (1992) *J. Med. Microbiol.* **37,** 389–395.
10. Papini, E., Bugnoli, M., de Bernard, M., Figura, N., Rappuoli, R. & Montecucco, C. (1993) *Mol. Microbiol.* **7,** 323–327.
11. Papini, E., Zoratti, M. & Cover, T. L. (2001) *Toxicon* **39,** 1757–1767.
12. Czajkowsky, D. M., Iwamoto, H., Cover, T. L. & Shao, Z. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 2001–2006.
13. Szabo, I., Brutsche, S., Tombola, F., Moschioni, M., Satin, B., Telford, J. L., Rappuoli, R., Montecucco, C., Papini, E. & Zoratti, M. (1999) *EMBO J.* **18,** 5517–5527.
14. Tombola, F., Carlesso, C., Szabo, I., de Bernard, M., Reyrat, J. M., Telford, J. L., Rappuoli, R., Montecucco, C., Papini, E. & Zoratti, M. (1999) *Biophys J.* **76,** 1401–1409.
15. Schraw, W., Li, Y., McClain, M. S., van der Goot, F. G. & Cover, T. L. (2002) *J. Biol. Chem.* **277,** 34642–34650.
16. Willhite, D. C., Cover, T. L. & Blanke, S. R. (2003) *J. Biol. Chem.* **278,** 48204–48209.
17. Galmiche, A., Rassow, J., Doye, A., Cagnol, S., Chambard, J. C., Contamin, S., de Thillot, V., Just, I., Ricci, V., Solcia, E., *et al.* (2000) *EMBO J.* **19,** 6361–6370.
18. Kuck, D., Kolmerer, B., Iking-Konert, C., Krammer, P. H., Stremmel, W. & Rudi, J. (2001) *Infect. Immun.* **69,** 5080–5087.
19. Peek, R. M., Jr., Blaser, M. J., Mays, D. J., Forsyth, M. H., Cover, T. L., Song, S. Y., Krishna, U. & Pietenpol, J. A. (1999) *Cancer Res.* **59,** 6124–6131.
20. Fujikawa, A., Shirasaka, D., Yamamoto, S., Ota, H., Yahiro, K., Fukada, M., Shintani, T., Wada, A., Aoyama, N., Hirayama, T., *et al.* (2003) *Nat. Genet.* **33,** 375–381.
21. Gebert, B., Fischer, W., Weiss, E., Hoffmann, R. & Haas, R. (2003) *Science* **301,** 1099–1102.
22. Russ, W. P. & Engelman, D. M. (2000) *J. Mol. Biol.* **296,** 911–919.
23. McClain, M. S., Iwamoto, H., Cao, P., Vinion-Dubiel, A. D., Li, Y., Szabo, G., Shao, Z. & Cover, T. L. (2003) *J. Biol. Chem.* **278,** 12101–12108.
24. McClain, M. S. & Cover, T. L. (2003) *Infect. Immun.* **71,** 2266–2271.
25. Kim, S., Chamberlain, A. K. & Bowie, J. U. (2003) *J. Mol. Biol.* **329,** 831–840.
26. Bass, R. B., Strop, P., Barclay, M. & Rees, D. C. (2002) *Science* **298,** 1582–1587.
27. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M. & Dobson, C. M. (1996) *J. Mol. Biol.* **255,** 494–506.
28. Bower, M. J., Cohen, F. E. & Dunbrack, R. L. (1997) *J. Mol. Biol.* **267,** 1268–1282.
29. Kelly, L. A., Gardner, S. P. & Sutcliffe, M. J. (1996) *Protein Eng.* **9,** 1063–1065.
30. 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.* **54,** 905–921.
31. Rees, D. C., Chang, G. & Spencer, R. H. (2000) *J. Biol. Chem.* **275,** 713–716.
32. Calladine, C. R., Pratap, V., Chandran, V., Mizuguchi, K. & Luisi, B. F. (2003) *Science* **299,** 661–662.
33. Miller, S., Bartlett, W., Chandrasekaran, S., Simpson, S., Edwards, M. & Booth, I. R. (2003) *EMBO J.* **22,** 36–46.
34. Perozo, E. & Rees, D. C. (2003) *Curr. Opin. Struct. Biol.* **13,** 432–442.