Comparative Structure–Function Analysis of Cytolethal Distending Toxins

Xin Hu, Dragana Nesic, and C. Erec Stebbins*
Laboratory of Structural Microbiology, The Rockefeller University, New York, New York

ABSTRACT **Cytolethal distending toxins (CDTs) constitute a family of bacterial proteins that enter eukaryotic cells with genotoxic activity leading to cell cycle arrest and apoptosis. CDTs are widespread, having been found in a variety of Gram-negative pathogens with a broad tissue tropism. The recently determined crystal structure of the *Haemophilus ducreyi* CDT provides a powerful starting point for analysis of the structure and function in this toxin family. In this study, we apply comparative modeling and structural analysis to extend the experimental structural information to multiple CDT toxins from a diverse species. Analysis of structurally and functionally important residues in the active subunit, CdtB, and putative cell delivery elements, CdtA and CdtC, begins to establish the fundamental, mechanistic elements of this unique holotoxin. The results reveal that key structural features with important functional consequences are highly conserved across different CDTs, providing a blueprint for directed examination of functional hypotheses in a variety of pathogenic contexts. Proteins 2006;62:421–434.**

Key words: cytolethal distending toxin; genotoxin; nuclease; protein–protein interaction; homology modeling

INTRODUCTION

The cytolethal distending toxins (CDTs) are a recently characterized family of bacterial protein toxins and have been identified in many important pathogens, such as *Campylobacter* spp. (*jejuni* and *fetus*), *Helicobacter* spp. (*cinaedi*, *hepaticus*, and *pullorum*), *Shigella dysenteriae*, *Actinobacillus actinomycetemcomitans*, *Haemophilus ducreyi*, *Salmonella typhi* and *paratyphi* A, and certain strains of *Escherichia coli*.[1–4] These Gram-negative pathogenic bacteria have been implicated in a wide range of human diseases. For example, *C. jejuni* is the leading cause of food-borne infectious illness in industrialized nations.[5] *S. typhi* is the causative agent of typhoid fever that claims the lives of millions people each year in developing countries.[6] *H. ducreyi* is associated with the genital ulcer disease chancroid, a sexually transmitted disease potentially involved in HIV transmission,[7] whereas *A. actinomycetemcomitans* is primarily an oral pathogen also associated with a variety of diseases, such as endocarditis, meningitis, osteomyelitis, abscesses, endophthalmitis, and periodontal disease.[8,9] Although the contribution of CDT to the pathogenesis of these diseases has not been clearly determined, many studies have shown that CDT as a potential virulence factor may play an important role in disease progression and immunosuppression.[10–12]

Recent studies have provided important insights into the mechanism of action of this unique toxin family. The toxic activity is associated with the products of an operon that encodes three proteins: CdtA, CdtB, and CdtC.[13,14] All three of these protein subunits are required to constitute the fully active holotoxin by forming a tripartite complex.[15] CdtA and CdtC are hypothesized to be necessary for the delivery of CdtB,[16–18] whereas CdtB exhibits weak nuclease activity and active site sequence similarity to the deoxyribonuclease I (DNase I) family.[18,19] CDT therefore can be viewed as an AB toxin, with CdtB taking the role of the active "A" subunit, while the CdtA–CdtC subcomplex is the "B" subunit or binding element.[15,20] It is proposed that the CDT holotoxin is internalized into target cells by means of receptor-mediated endocytosis.[21] One exception is the CDT from *S. typhi* (StCDT), which notably encodes only the CdtB subunit without the expected CdtA and CdtC. StCdtB is delivered into the host cell following the internalization of the pathogen that is achieved through its type III secretion system.[22] Similar to many other toxins, such as ricin or Shiga toxins, presumably CdtB is transported from the endoplasmic reticulum (ER) to the cytosol by retrotranslocation, and finally to the nucleus, where it activates DNA damage checkpoints.[23–25] The cellular response to CDT intoxication leads to a robust DNA damage response, cell cycle arrest, nuclear enlargement, and cellular distension.[19,26] Intoxicated cells of lymphatic origin undergo apoptotic cell death.[11,27]

The crystal structure of CDT from *H. ducreyi* has been recently determined, revealing a wealth of knowledge on the molecular basis of the assembly and function of the holotoxin.[20] For example, the N-terminus of CdtC extends to the active site of CdtB and sterically occludes the active site, suggesting a possible autoinhibition for the holotoxin. CdtA and CdtC together contribute toward the formation

Grant sponsor: National Institute of Allergy and Infectious Disease; Grant number: 1U19AI056510. Grant sponsor: The Rockefeller University.

*Correspondence to: C. Erec Stebbins, Laboratory of Structural Microbiology, The Rockefeller University, New York, NY 10021. E-mail: stebbins@rockefeller.edu

Received 19 May 2005; Revised 6 July 2005; Accepted 9 August 2005

Published online 22 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20767

of a large aromatic patch and an adjacent deep groove on the protein surface, which might play an important role in cell surface binding and internalization of the holotoxin. Since CDT exists in a variety of bacterial species, it remains a question of what general features observed in the *H. ducreyi* holotoxin are conserved and applicable to other bacteria. As the three subunits of the CDT holotoxin are required to form an active tripartite complex, it is of great interest to examine the structural properties of the protein–protein interactions between CdtA and CdtB, CdtA and CdtC, and CdtB and CdtC from different CDTs to gain insight on the formation and stability of the ternary structure. An overall molecular understanding of the importance and contribution of these different elements in the holotoxin should shed light on the function and cellular effects of CDT intoxication.

Herein we apply comparative modeling and structural analysis to explore the common characteristics and structural divergences of CDTs. Sequence analysis reveals that there is significant sequential identity and consensus sequence conservation within the family, as well as with the template structure of *H. ducreyi* CDT (HdCDT), making it possible for a meaningful comparative modeling approach to three-dimensional (3D) structural prediction. The modeled structures have been validated to maintain a well-conserved global fold with satisfied stereochemical quality. Detailed structural analysis of the atomic interactions between the subunit interface of CdtA and CdtB, CdtA and CdtC, and CdtB and CdtC was conducted. Comparison of the structural features and surface properties of different CDTs, especially among the divergent CdtA and CdtC subunits, reveals much of the information on the assembly, stability, and function of the holotoxin that may encourage further site-directed mutagenesis experiments and other biological studies.

MATERIALS AND METHODS

Sequence Alignment and Homology Modeling

The amino acid sequences of CDTs were retrived from SWISS-PROT database.[28] The CDT holotoxin in bacterial species *C. jejuni*, *H. hepaticus*, *E. coli* (I, II, III), *A. actinomycetemcomitans*, *H. ducreyi* (in which all the three subunits' sequences are available), as well as *S. typhi* were selected in this study. Coordinates of the crystal structure of CDT from *H. ducreyi* are those that were deposited under entry code 1SR4 from the Protein Data Bank (PDB).[29] Multiple sequence alignment was performed by means of the CLUSTAL W package applying the default parameters.[30] The alignment was manually adjusted during the process of model optimization in an iterative approach.

The homology models of CDTs were constructed using the restraint-based comparative modeling program MODELLER 7v7.[31,32] A set of 20 models from random generation of the starting structure were calculated for each CDT sequence, and the representative model was selected with the best (lowest) value of MODELLER objective function. The 3D fold of the generated models was verified by the program PROSA II,[33] and problematic regions were spotted from the analysis of PROSA II energy profile. The resulting models were further optimized with an iterative approach until no significant improvement was obtained. The stereochemical quality and protein structure of the final models were validated by the PROCHECK[34] and PROSA II.[33] The root-mean-square deviation (RMSD) between the modeled structures and template was calculated by means of an iterative procedure with Swiss PdbViewer v3.7.[35] The solvent accessibility and surface properties analyses were performed with WHAT IF[36] and GRASP.[37]

Modeling of CdtB/DNA Complex

To model the CdtB binding to DNA, the double-stranded DNA was positioned to the CdtB active site by superimposing with the DNase I/DNA complex (PDB code: 1DNK[38]). The docked CdtB/DNA complex was then subjected to a full energy minimization to release any steric clash. The energy minimizations were carried out employing modeling package SYBYL 7.0.[39] MMFF94s force field was applied with 0.05 kcal/Å convergence and 5000 steps using the Powell method in SYBYL.

Molecular Dynamics (MD) Simulations

MD simulations were carried out on the CDT models using the SANDER module of the AMBER 8.0 package.[40] All systems were solvated by a cubic box of transferable intermolecular potential (TIP3P) water molecules. The solvated protein systems were subjected to a thorough minimization prior to MD simulations by first minimizing the water molecules while holding solute frozen (1000 steps using the steepest descent algorithm), followed by minimization of the whole system, with 5000 steps of conjugate gradient to remove close contacts and to relax the system. Bond lengths involving bonds to hydrogens were constrained with SHAKE[41] and the timestep for all MD simulations was set to 2 fs. A nonbonded cutoff of 10 Å was used, and the nonbonded pair list was updated every 25 timesteps. Periodic boundary conditions were applied to simulate a continuous system. The particle mesh Ewald (PME) method[42] was employed to calculate the long-range electrostatic interactions. All the MD runs were set up using the same protocol: The simulated system was first subjected to a gradual temperature increase from 0 K to 300 K over 20 ps, and then equilibrated for 100 ps at 300 K, followed by the production run of 1-ns length in total. Constant temperature and pressure (300 K/1 atm) were maintained using the Berendsen coupling algorithm[43] with time constant for heat bath coupling of 0.2 ps. The resulting trajectories were analyzed using the PTRAJ module of AMBER 8.0 package. The RMSDs of the backbone were calculated from the trajectories at 1-ps intervals, with the initial structure as the reference.

RESULTS AND DISCUSSION

Homologs in the CDT Family

An examination of the CDT family shows that the CDT subunits exhibit varied levels of homology. CdtB is the most highly conserved, with 45% sequence identity even

TABLE I. Percentage of Amino Acid Identities and Similarities (in Parentheses) in Different Bacterial CDTs Shared with *Haemophilus ducreyi* CDT (HdCDT)

Species	CDT designation	CdtA	CdtB	CdtC
Haemophilus ducreyi	HdCDT	—	—	—
Actinobacillus actinomycetemcomitans	AaCDT	93 (95)	97 (98)	95 (97)
Helicobacter hepaticus	HhCDT	36 (55)	51 (69)	18 (32)
Campylobacter jejuni	CjCDT	30 (55)	50 (69)	19 (36)
Escherichia coli (6468/62)	EcCDT I	38 (59)	49 (67)	22 (45)
Escherichia coli (9142-88)	EcCDT II	26 (51)	48 (65)	19 (38)
Escherichia coli (S5)	EcCDT III	26 (51)	49 (65)	19 (41)
Salmonella typhi	StCdtB	—	50 (67)	—

between the most distantly related CDTs. The CdtA and CdtC subunits are quite divergent, and the pairwise identity ranges from 19% to 95%. Phylogenic analysis reveals that some bacterial CDTs are of evolutionary linkage, for instance, *C. jejuni* CDT (CjCDT) shares more than 90% homology with *H. hepaticus* CDT (HhCDT). *E. coli* CDT II (EcCDT-II) is similar to *E. coli* CDT III (EcCDT-III), whereas *E. coli* CDT I (EcCDT-I) and *S. dysenteriae* CDT (SdCDT) are more closely related (> 90% identity). The closest relative of HdCDT is the *A. actinomycetemcomitans* CDT (AaCDT). *S. typhi* CDT (StCdtB) is the most distant from other species. Table I shows the sequence identity and similarity of different CDTs with HdCDT, for which the experimental crystal structure is determined. The sequence identities between HdCDT and AaCDT are 93%, 97%, and 95% for the CdtA, CdtB, and CdtC subunits, respectively. With respect to other bacterial species, the CdtB sequence identity is between 48% and 51%, and only about 26–38% identity for CdtA and 19–22% for CdtC subunits.

CdtB also exhibits weak sequence similarity to proteins of the DNase I family,[18] and shares approximately 12% identity between HdCDT and DNase I based on structural alignment. The key residues in common with DNase I that are critical for DNA catalytic activity are strictly conserved in all CdtB homologs. These include the two catalytic histidines (His160, His274, Hd numbering), the metal coordinating resides (Glu66, Asp199, and Asp273), and DNA contacting residues (Arg117, Arg144, and Asn201). Figure 1(a) shows the multiple sequence alignment of CdtB subunit from *A. actinomycetemcomitans*, *C. jejuni*, *H. hepaticus*, and *E. coli* with the template of *H. ducreyi*. Structurally conserved regions lie in the active site and β-strands of secondary structural elements according to the template of HdCDT structure. Variable regions mainly occur within loops, for instance, the two "serine-rich" motifs (residues 136–139 and 182–186) observed in HdCdtB and AaCdtB that are lacking in other bacterial species.

The sequence alignment of the CdtA and CdtC subunits from different bacteria shows that the secondary structural elements are generally conserved despite the low overall sequence identity [Fig. 1(b and c)]. A good conservation is found for the large aromatic patch in CdtA, which consists of five to eight aromatic residues. Other structurally conserved regions in CdtA and CdtC include the presence of two intrastrand disulfide bonds and a number of hydrophobic residue motifs. CdtA and CdtC also exhibit structural similarity with each other (~40%) and both are homologous to the B-chain repeats of the plant toxin ricin (with 16% sequence identity between ricin and CdtA).[20] Structure-based sequence alignment shows that the β-trefoil fold is well maintained in these lectin-type protein domains. CdtC differs from CdtA mainly with a large deletion between residues 53–61 (equivalent to β4/β5 in CdtA), as well as a loop insertion between β2 and β3 (residues 71–85). The region with the best degree of similarity between CdtA and CdtC lies in the β10/β11-sheet in CdtA corresponding to the β6/β7-sheet in CdtC. Interestingly, the β6/β7-sheet is also the highest conserved region within the CdtC family.

Comparative Modeling of CDTs

The sequence analysis reveals that there is enough consensus amino acid conservation within different CDTs for a meaningful comparative modeling approach to 3D structural prediction. The reliability of comparative modeling depends heavily on the degree of homology to the template structure, as well as an accurate alignment of modeled sequences.[45,46] Given the high similarity of CdtB subunit to the template structure of HdCdtB across the CDT family, the CdtB models can be the most reliably constructed. The overall sequence identities of CdtA as well as CdtC homologs are generally less than 30%, posing a challenge to comparative modeling.[47] Since the alignment errors increase rapidly below 30% sequence identity and are the most substantial origin of errors in comparative models, an elaborate and convincing alignment would strongly improve the modeling accuracy.[46,47] In CdtA and CdtC, a good conservation of regions with local sequence similarity and functionally important motifs (e.g., the aromatic patch and the intrastrand disufide bonds) was found to help greatly in guiding improvements to the initial sequence alignment that was obtained using CLUSTAL W.[30] The canonical trefoil fold with hydrophobic residue motifs is known to be conserved in these lectin-type proteins; therefore, no gaps were allowed within them. In addition, we refined the sequence alignment by an iterative approach during the modeling process, which has proved an effective way to optimize the comparative models.[48]

All the CDT models were constructed using the restraint-based comparative modeling program MODELLER v7v.[31,32] The initial models were selected according to the

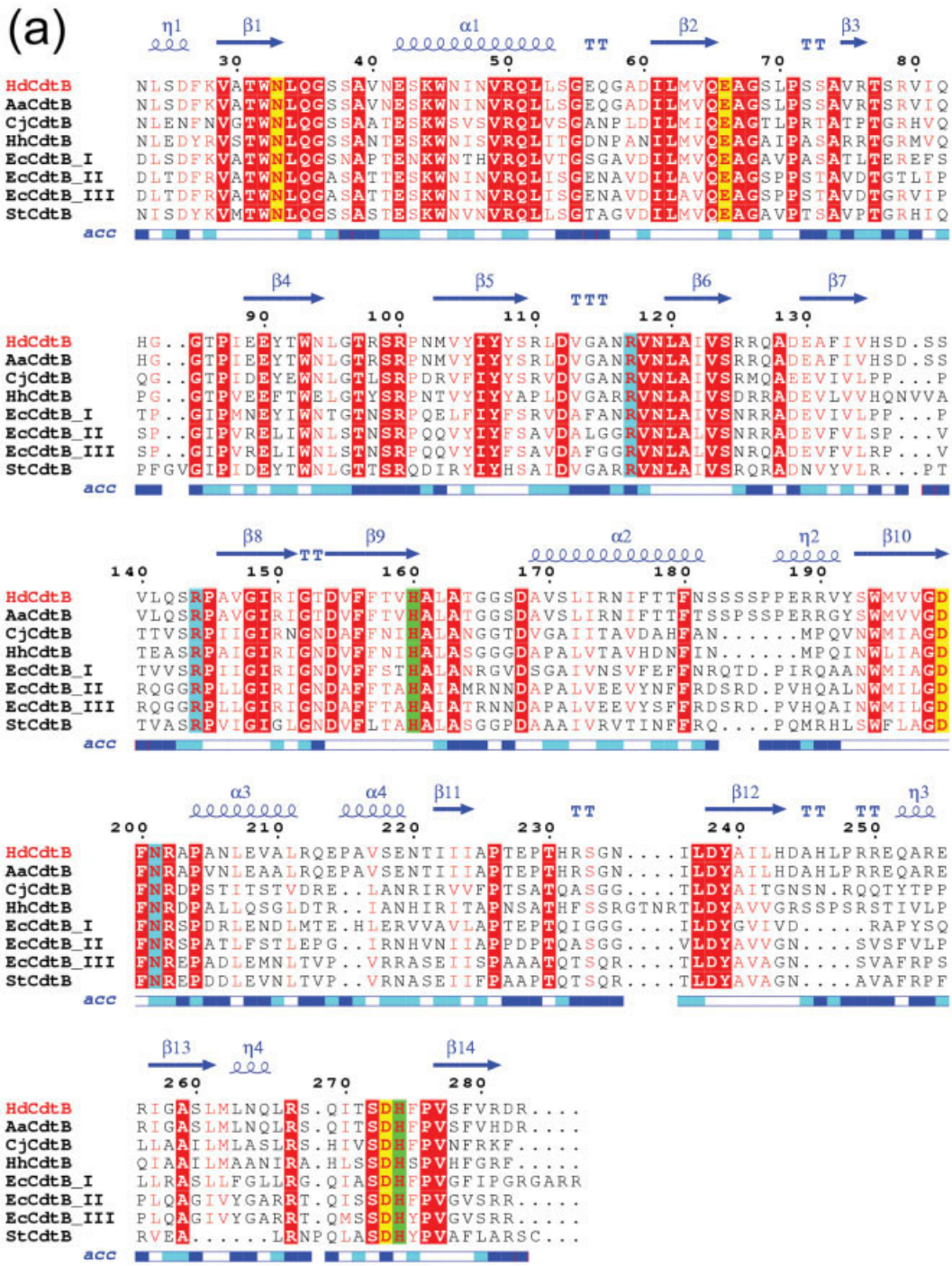

Fig. 1. Sequence alignment of (**a**) CdtB, (**b**) CdtA, and (**c**) CdtC in the CDT family. Strict sequence conservation is shown in white letters with a red background, and strong sequence conservation in red letters. The solvent accessibility of each residue in the HdCDT structure is indicated in the bar at the base of the sequences, with white representing buried residues, dark blue representing solvent-accessible residues, and light blue representing and intermediate value. The disulfides are marked with pink triangles below the cysteine residues. (a) The two catalytic histidine residues are green; the metal coordinating residues are yellow; and the DNA contacting residues are cyan. (b) The aromatic patch residues are colored in green. (c) The groove residues are yellow. The most conserved region (β6/β7 sheet) is framed in a blue box. The figure was generated using ESPript.[44]

MODELLER objective function score and accessed using the programs PROCHECK[34] and PROSA II.[33] Examination of the PROSA II energy profile showed that the conserved canonical fold in CdtA and CdtC models was comparable to the template; potential errors mainly occurred in loop regions resulting from uncertain gaps in the sequence alignment scheme. These regions, for example, the β2/β3 and β5/β6 loops in CdtC of *C. jejuni* and *H. hepaticus*, were manually adjusted, and the resulting models were evaluated. In general, all the CDT models

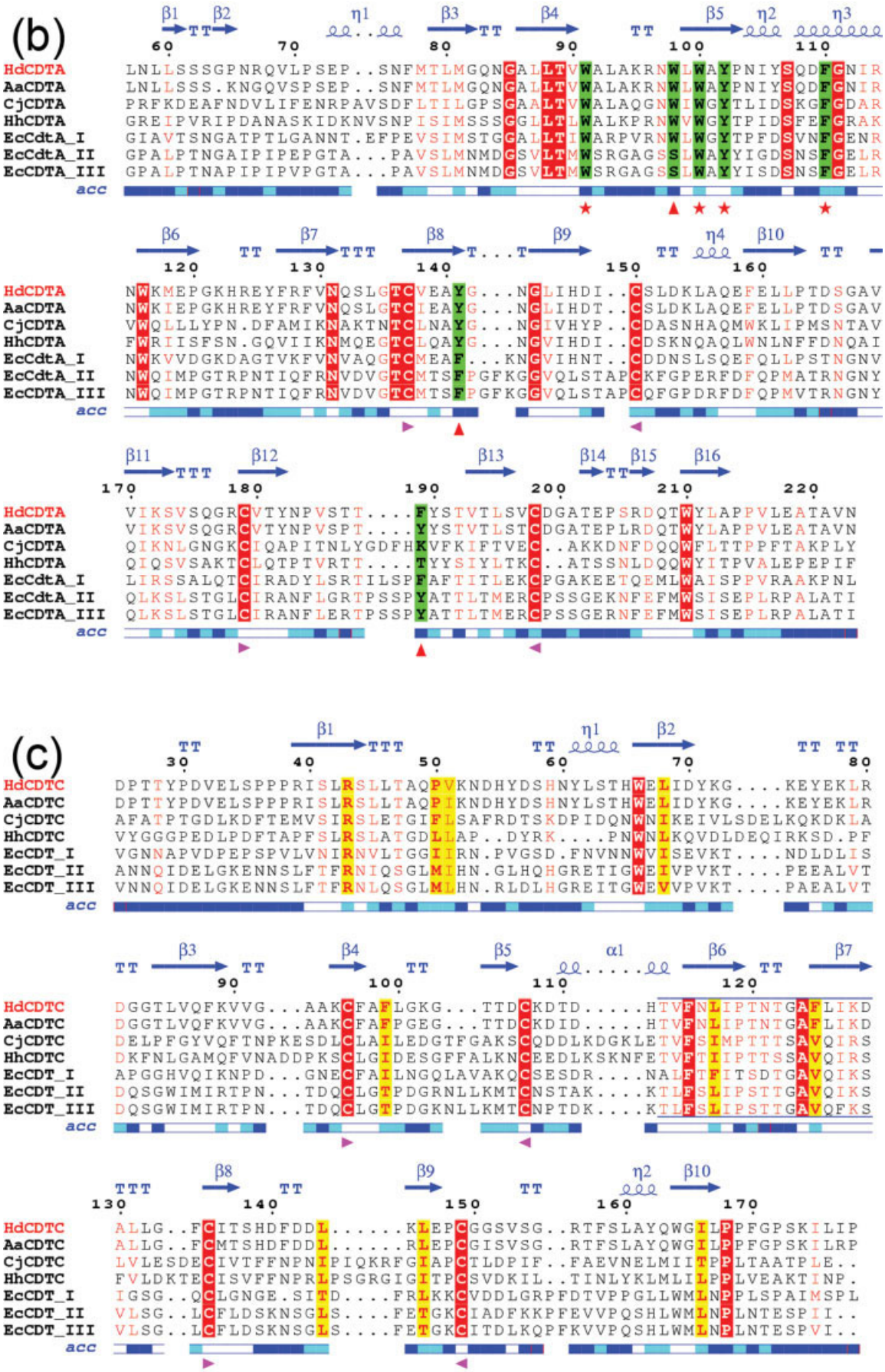

Figure 1. (Continued.)

TABLE II. PROCHECK Parameters (Percentage of Residues in the Core, Allowed, Generously Allowed, and Disallowed Regions of the Ramachandran Plot) and PROSA II *Z*-Score Calculated for the CDT Structures

CDT	Core (%)	Allowed (%)	Generous (%)	Disallowed (%)	*Z*-score CdtA/CdtB/CdtC
HdCDT (1SR4)	89.4	10.6	0	0	−5.61/−8.69/−6.10
AaCDT	92.0	7.6	0.4	0	−5.85/−8.67/−5.94
CjCDT	84.7	113.0	1.5	0.8	−5.29/−7.06/−5.50
HhCDT	85.8	12.5	1.2	0.6	−5.55/−6.87/−5.15
EcCDT I	86.8	11.1	1.4	0.6	−5.09/−7.12/−5.66
EcCDT II	86.5	11.5	1.4	0.7	−5.40/−7.82/−5.43
EcCDT III	86.4	11.4	1.4	0.8	−5.38/−7.79/−5.30
StCdtB	86.8	12.3	0.5	0.5	/−7.52/

were built with an iterative "alignment–modeling–evaluation" cycle until no significant improvement was obtained. The overall quality of stereochemistry and 3D fold of the final models were validated; the results are shown in Table II. Both PROSA II and PROCHECK confirm that reasonable models were obtained. The PROCHECK parameters are quite satisfactory, with most models featuring 85–92% of residues in the most favored regions of the Ramachandran plot (more than 99% in allowed region overall), and less than 1% of residues in disallowed region. *Z*-scores computed by PROSA II are also good, comparable to that of the template HdCDT (Table II).

MD Simulations of the CDT Tripartite Complex

MD simulations were conducted to investigate the stability and dynamic properties of the modeled CDT structures. The protocol was validated first by the MD simulations of the crystal structure of CDT from *H. ducreyi*. The RMSDs calculated on the backbone atoms of HdCDT protein reached a plateau with an average of 1.40 Å as compared to the initial crystal structure [Fig. 2(a)], indicating that the system has evolved to an equilibrium state during the 1-ns simulations. In Figure 2(b–d) are plotted the RMSDs of CDT proteins from *A. actinomycetemcomitans*, *C. jejuni*, and *E. coli* I with their respective starting structures. All trajectories appear to be equilibrated over the timecourse of 1-ns simulations. The average RMSDs are below 2.0 Å in general for all the modeled CDT structures. An analysis of the average MD structures with respect to the initial structures of CDTs reveals that the gross features of the secondary structure and 3D folding are well maintained. The stability of these CDT models allows a reliable structural comparison.

All CDT structures possess a ternary topology characterized with three subunits that interact intimately by forming three interdependent globular protein–protein interfaces (Fig. 3). The total solvent-accessible surface area for the entire holotoxin is up to 10,000 Å^2, while more than 3000 Å^2 surface area are buried in the three subunit interfaces, with each more than 1000 Å^2 of interface area (Table III). Besides the extensive protein–protein interfaces to cement the assembly of the ternary structure, the CDT complex is also stabilized by several extended, nonglobular interactions involving N- and C-terminal "tails" of CdtA and CdtC that interact with other elements of the holotoxin. These polypeptide–protein interactions are quite stable throughout the MD simulations. One exception is the long peptide of CdtA N-terminal tail that is extended to the solvent without any contacts with other elements, and therefore exhibits high mobility in the modeled CDT structures. Indeed, the RMSD plots from the simulations showed considerably increased stability, while the flexible tail of CdtA N-terminus was not included.

Structure and Dynamics of CdtB Nuclease

The high homology among CdtB subunits allows for reliable structural prediction. The RMSDs of CdtB modeled structures are less than 0.6 Å from the experimental HdCdtB structure. CdtB adopts a canonical, four-layered fold consisting of a central 12-stranded β-sandwich packed between outer α-helices and loops on each side of the sandwich. The DNA binding site lies in the middle cleft of the β-sandwich packing layer, characterized by a number of residues for substrate DNA contacts and metal ion coordination as found in DNase I. All these key residues are structurally conserved in the CDT family and superimposed perfectly with DNase I. One of the putative catalytic histidines in DNase I (His134) makes hydrogen bonds with a nearby acidic residue Glu78 to facilitate the phosphodiester cleavage. The equivalent histidine in CdtB (His160, Hd numbering) is found lacking such a bonding partner and may explain the low activity of CdtB in vitro.[20] An absolutely conserved valine is present instead of glutamate at this position. Structural differences among different CdtB subunits mainly occur in connecting loops and the length/orientation of the outer helices. CjCdtB has a four–amino acid loop insertion between strand β11 and β12, while the following loop is much shorter in *E. coli* as well as in *S. typhi*. However, those insertions and deletions are exposed to solvent surface and do not affect the core structure.

A comparison of CdtB structures with DNase I shows that there is a large loop insertion (residues 74–93, Hd numbering) proximal to the active site of CdtB. This loop insertion is also one of the most highly conserved regions in the CDT family (StCDT has two more residues). Our preliminary experiments demonstrated that mutation of some residues in the loop can abrogate the cytolethal

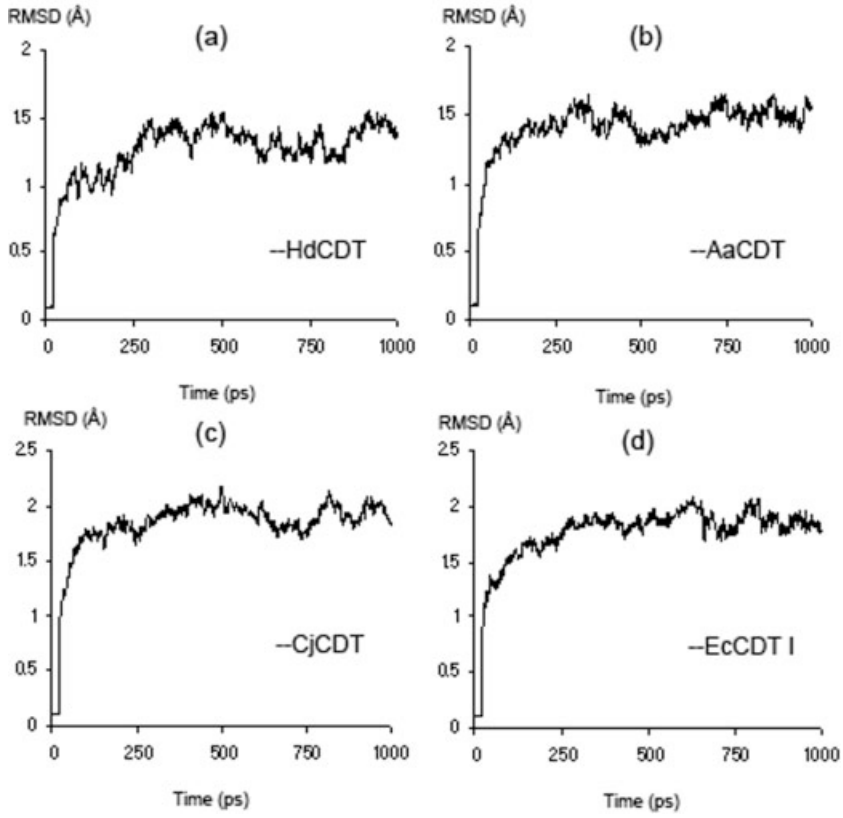

Fig. 2. RMSDs of the backbone atoms of CDTs with respect to their starting structure during 1-ns MD simulations: (**a**) *H. ducreyi*, (**b**) *A. actinomycetemcomitans*, (**c**) *C. jejuni*, and (**d**) *E. coli* I.

effects of the toxin (manuscript in preparation), raising an issue to further examine the potential role of this important region for the catalytic activity. MD simulations showed that this loop region is rather flexible and sensitive to perturbation. Since the loop is close to the active site, it might function as a "flexible activation loop" to facilitate the DNA binding, or serve as a bait able to interact with an as yet unidentified partner for DNA catalysis or nuclear import.

CdtB Autoinhibition and DNA Binding

One of the striking findings revealed by the experimental HdCDT structure is that the N-terminus of CdtC extends away from the globular β-trefoil domain and occludes the active site of the CdtB nuclease. In vitro experimental studies show that the CdtC N-terminal peptide tail appears to play an autoinhibitory role in the holotoxin of *H. ducreyi*.[20] Since the N-terminus of CdtC is the most divergent region in different CDTs and no sequence similarity can be detected in the primary sequence alignment, it remains unclear whether the N-terminal peptide binding to the CdtB active site is a general mechanism of autoinhibition in the CDT family. As shown in Figure 4(a), the nonglobular polypeptide of the N-terminus of HdCdtC makes several contacts to the active site residues (e.g., CdtB His160 with CdtC Val32) and especially to the conserved DNA binding residues Arg117 and Arg144. The residue Asp31 of HdCdtC in particular forms a salt bridge with HdCdtB Arg144, and Val32 interacts with the metal coordination residue Glu66, as well as the catalytic residue His160 of HdCdtB. Analysis of the peptide–protein interaction within other CDT homologs shows that, despite the low sequence similarity at the CdtC N-terminus, many interactions observed in HdCDT are retained in the modeled CDT complexes. In *H. hepaticus*, the residue Asp22 of HhCdtC forms a salt bridge with CdtB Arg139, and Leu23 makes van der Waals (vdW) contacts with Glu60 and His155 of HhCdtB [Fig. 4(b)]. Similar binding modes are also found in *C. jejuni* and *E. coli*. In general, our modeling results suggest that the long peptide tail of the CdtC N-terminus in different CDTs (~20 amino acids) is capable of adopting such a conformation in a complementary manner to occupy the active site of CdtB.

To gain insight on the DNA catalysis by CdtB nuclease, we also modeled DNA binding at the CdtB active site. The DNA double-strand was docked into the CdtB subunit by aligning with the DNase I/DNA complex. The CdtB/DNA complexes were then subjected to a full-energy minimization to release any steric clash. As shown in Figure 5, the DNA fit well into the active site of AaCdtB and StCdtB, in a similar manner as in DNase I, mainly forming salt bridge and hydrogen bonds with the

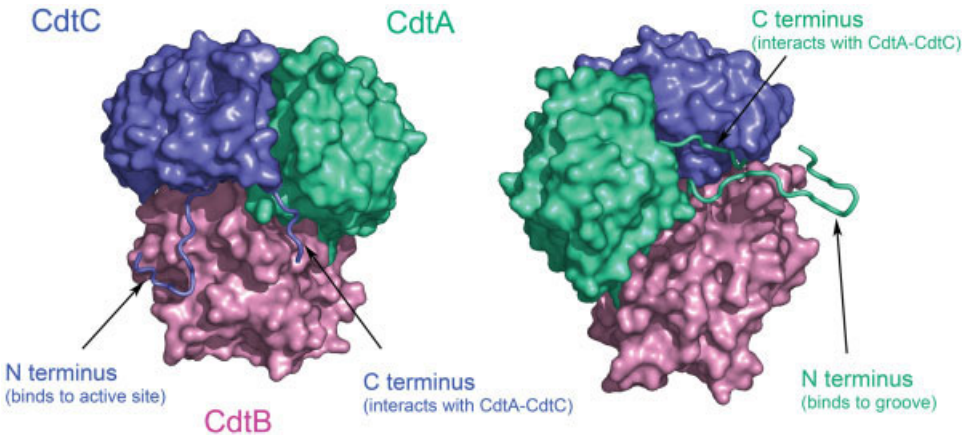

Fig. 3. Surface representations of CDT holotoxin shown in green (CdtA), pink (CdtB), and blue (CdtC), respectively. The N and C terminus of CdtA and CdtC are shown as cartoon tube.

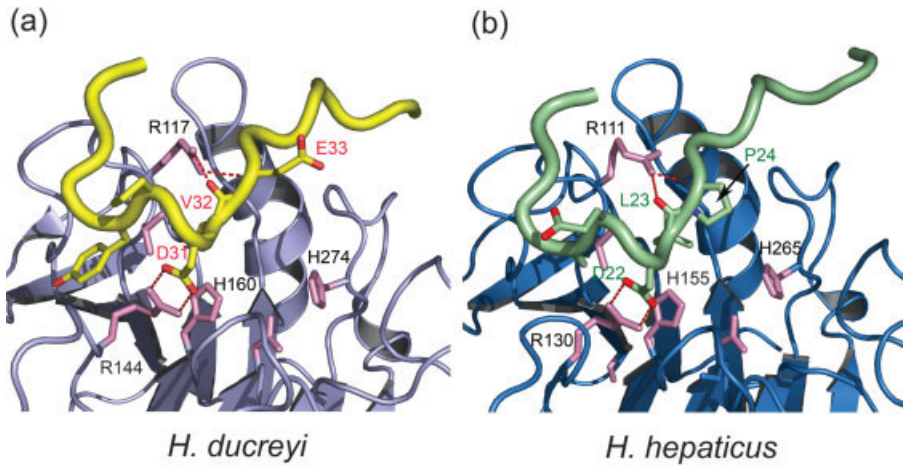

Fig. 4. The N-terminal polypeptide of CdtC binds to the CdtB active site of (**a**) *H. ducreyi* and (**b**) *H. hepaticus*. The side-chain atoms of the key residues at the CdtB active site are pink.

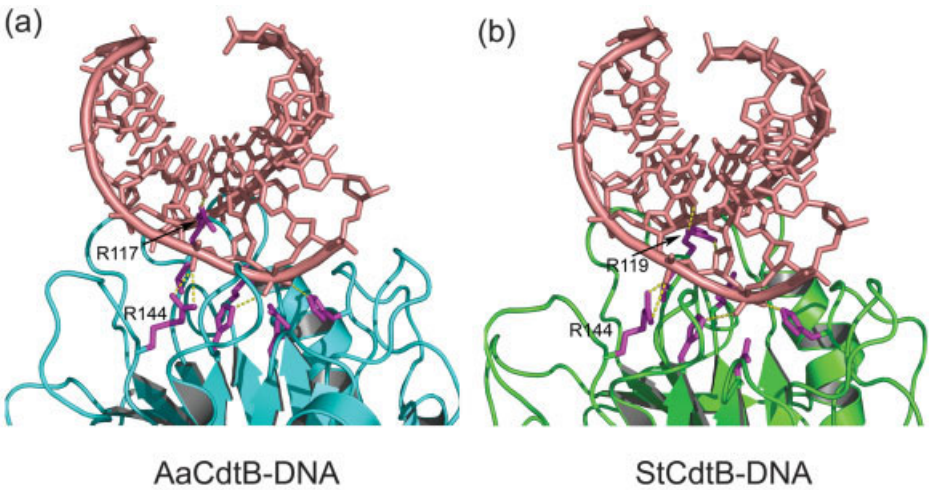

Fig. 5. The double-stranded DNA binds at the CdtB active site of (**a**) *A. actinomycetemcomitans* and (**b**) *S. typhi*. The side-chain atoms of the DNA contacting residues at the active site are magenta.

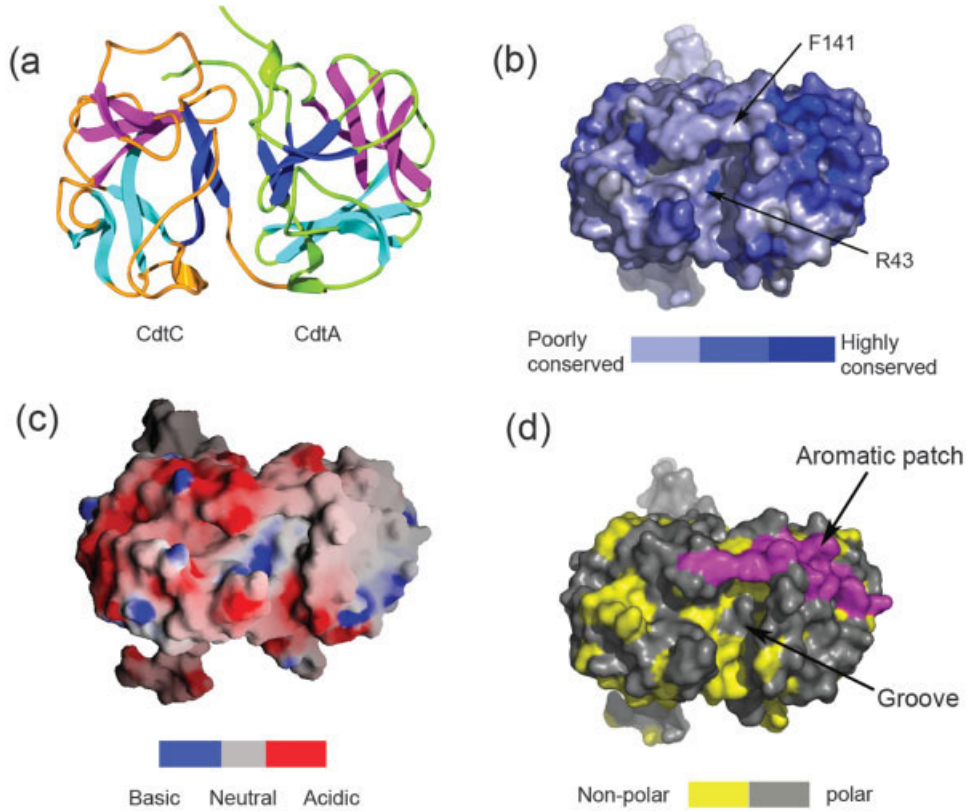

Fig. 6. Structural features of the lectin proteins CdtA and CdtC. (**a**) The symmetric trefoil fold in CdtA and CdtC subunits. The trefoil folds are in blue, magenta, and cyan. CdtC lacks two strands within one fold (blue). (**b**) Mapping of sequence conservation among CdtA and CdtC subunits in the CDT family. The figure was prepared with ProtSkin (http://www.mcgnmr.ca/protskin/) and PyMOL (http://pymol.sourceforge.net/).[51] Solid surface represented by *H. ducreyi* CDT (HdCDT). (**c**) Electrostatic surface representations of CdtA–CdtC (AaCDT). The electrostatic potential is calculated using GRASP.[37] (**d**) Surface groove and the hydrophobicity of CdtA–CdtC of the holotoxin (AaCDT). The aromatic patch is colored in magenta.

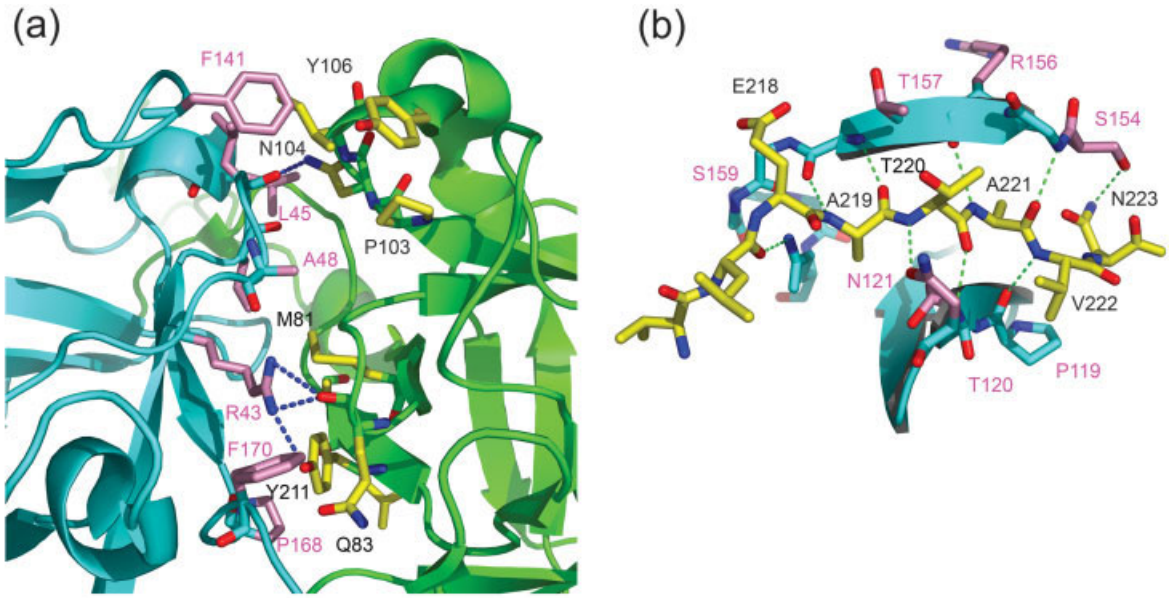

Fig. 7. (**a**) Globular protein–protein interaction interface between CdtA (green) and CdtC (cyan) in *H. ducreyi* CDT. The side-chain carbon atoms of residues involved in the interfacial interaction are colored in yellow (CdtA) and pink (CdtC). (**b**) Intermolecular β-sheet interaction of CdtA C-terminal peptide (yellow) binding into CdtC (cyan and pink) in *H. ducreyi* CDT.

TABLE III. Solvent-Accessible Surface Area and Interface Properties for CdtA, CdtB, and CdtC Subunits in CDTs

	Solvent-accessible surface area (Å^2)			Interface area (Å^2)			Interface hydrophobicity (% of nonpolar atoms)		
CDT	CdtA	CdtB	CdtC	CdtA–CdtC	CdtB–CdtC	CdtB–CdtA	CdtA–CdtC	CdtB–CdtC	CdtB–CdtA
HdCDT	9347.8	11,542.7	9238.0	1556.0	1451.1	1361.6	69.7	56.7	49.5
AaCDT	9335.0	11,465.1	9286.1	1497.7	1409.9	1297.1	69.2	58.1	54.5
CjCDT	9895.0	10,944.5	10,660.6	1453.2	1390.1	1085.2	69.7	58.5	64.1
HhCDT	9965.3	11,348.9	10,365.3	1523.0	1382.2	1242.0	70.8	60.6	56.8
EcCDT I	9811.0	11,591.9	9911.6	1561.7	1445.1	1159.1	72.7	61.9	56.7
EcCDT II	10,145.2	11,679.3	9923.4	1527.6	1531.1	1147.3	69.2	56.9	58.6
EcCDT III	10,082.6	11,645.1	9909.0	1506.5	1522.6	1145.4	68.9	56.3	59.9
StCdtB	—	11,812.0	—	—	—	—	—	—	—

DNA backbone phosphates through conserved residues Arg144 and Arg117 (Aa numbering). Except for a short loop at the N-terminus connecting the strand β1 and α-helix H1, which was reoriented slightly due to steric clash with the DNA, no significant conformation changes were observed for the CdtB nuclease.

Highly Conserved Motifs in CdtA and CdtC Subunits

CdtA and CdtC are both lectin-type structures homologous to the B-chain repeats of the toxin ricin, a receptor-binding domain for cellular uptake.[49,50] The core fold consists of a three-fold pseudosymmetric trefoil of 12 β-strands, arranged in three separate β-sheets. The β-trefoil fold in CdtA and CdtC possesses a cylindrical shape with a short length between relatively flat ends, and appears well conserved in all modeled structures (Fig. 6). The RMSDs of CdtA and CdtC with respect to the crystal structure of HdCdtA and HdCdtC are less than 1.0 Å. *C. jejuni* and *H. hepaticus* contain several loop insertions between conserved β-sheets (e.g., β4/β5 and β7/β8 in CdtC, β6/β7 in CdtA). Similar to the structures of CdtB, these insertion/deletions are mostly exposed to solvent and contribute to a relatively large solvent-accessible surface area (Table III). Inspection of the surface properties of CdtA and CdtC from different species shows that they generally resemble each other in terms of shape and charge distribution. The overall electrostatic potential of CdtA–CdtC subcomplex is quite acidic compared to CdtB, and CdtC is more acidic than CdtA. Except for the aromatic patch, other residues on the protein surface of CdtA and CdtC are divergent, with only approximately 25% of solvent-exposed surface residues being structurally conserved among the CDT family.

The aromatic patch and adjacent groove show a good conservation in all CDTs. The large aromatic patch on the top of surface is striking, not only forming a hydrophobic cluster but also contributing a large portion of the solvent-accessible surface (~40% of total solvent-accessible surface area of CdtA). Both of these aspects are significant to the potential role in cell surface binding. The groove is formed mainly along the CdtA and CdtC interface, and extends to the hydrophobic core of CdtC, which is exposed to solvent from the deletion of a β-strand within the trefoil fold. The groove is characterized by a hydrophobic stretch at the bottom, and a number of charged and polar residues flanking on the top (Fig. 6). A positively charged arginine residue (Arg43 in HdCdtC) is situated at the bottom of the groove and strictly conserved in all CDTs. Although other residues surrounding the groove are not identical cross the CDT species, they are mostly hydrophobic (such as Phe99, Leu144, and Leu146 in HdCdtC) and form a highly conserved hydrophobic surface that possesses potential for receptor binding. The electrostatic potential of the groove is relatively neutral compared to the acidic surface of CdtA and CdtC. A comparison of the groove in different CDTs shows that the size of the groove, particularly the opening at the surface, appears to be more variable. The groove in *H. ducreyi* and *A. actinomycetemcomitans* is shielded by an aromatic phenylalanine residue (Phe141 on CdtC) on the top, which also makes a connection to the large aromatic patch on CdtA. However, such a residue is absent in *E. coli* (which possesses an asparagine at this position), making the groove more open in these species.

In line with the low sequence identities in CdtA and CdtC homologs, the highly conserved structures of the aromatic patch and the adjacent groove in CdtA and CdtC have strong functional implications. It has been shown in HdCDT that mutation of the patch renders the toxin biochemically unperturbed but inactive against cells, suggesting that these residues are potentially involved in the receptor binding but not in protein assembly.[20] Other studies have also shown that CdtA and CdtC are each able to bind to the cell surface, but the binding affinities increase significantly when both components are combined.[17,52] Our modeling results indicate that both CdtA and CdtC share similar structural properties across different species with an essential element likely devoted to host cell binding, and thus may explain the ability of each protein to bind to the cell surface independently of the other. On the other hand, since CdtA and CdtC are both integral components of the groove, the lack of one or the other of these proteins would likely severely impair the function of this structural element. Further mutagenesis studies on the groove, especially the highly conserved residues such as Arg43 and Phe141 in CdtC, should provide experimental evidence and more insight on the role of CdtA and CdtC in host cell receptor binding.

Intriguingly, the experimental structure of HdCDT shows that CdtA N-terminal extension forms a crystal-

TABLE IV. Dynamics of Hydrogen Bonds in the CdtA–CdtC Interaction Interface With Occurrence (%) During Simulations

H-bonds		X-ray distance	MD-average distance	Occupies
CdtA	CdtC	(Å)	(Å)	(%)
Intermolecular β-sheet				
Leu217 O	Gln163 NH2	2.91	2.93 (0.15)	99.98
Ala219 O	Phel58 NH	2.81	2.94 (0.14)	98.98
Thr220 NH	Asn121 OD1	2.83	2.92 (0.14)	98.36
Ala221 O	Gly155 NH	2.79	3.01 (0.19)	83.62
Ala221 NH	Arg156 O	2.98	3.05 (0.16)	98.36
Val222 NH	Pro119 O	3.04	2.91 (0.13)	99.90
Asn223 NH2	Asp81 O	2.71	0.90 (0.14)	98.98
Globular interface				
Gln83 O	Arg43 NH1	2.76	2.95 (0.19)	81.27
Tyr211 O	Arg43 NH2	3.11	3.16 (0.18)	33.98
Asn104 NH2	Leu46 O	2.96	3.04 (0.19)	10.54

packing contact by interacting with the groove in a symmetry-related molecule and inserting deeply into it.[20] It is likely to be a fortuitous crystal-packing arrangement induced by high protein concentration rather than a physiological interaction, as the CDT complex shows no signs of oligomerization by gel-filtration chromatography.[20] Further studies demonstrated that the peptide corresponding to the N-terminal residues 58–69 does not bind to CDT to inhibit toxicity when applied in cellular assays (unpublished data). As a functional implication, the CdtA N-terminus in the crystals may mimic a peptide from a cell surface receptor for CDT binding. The observed N-terminal tail of CdtA from *H. ducreyi* binds in the groove by forming a salt bridge with the conserved arginine residue and making a number of polar interactions with the aromatic patch. From the point of inhibitor design, the interaction of the N-terminus within the groove provides a good initial choice for designing a peptide mimic. This may point to a possible mode of interaction with a cell surface receptor in which the β-trefoil fold is completed through a strand insertion from a host peptide.

CdtA–CdtC Interaction Interface

To probe the functionally important residues and conserved pattern involved in the interfacial interaction, we analyzed the three subunit interfaces of CdtA–CdtC, CdtA–CdtB, and CdtB–CdtC for each of the modeled CDT structures. The protein–protein interaction interface is defined by residues with 10% or more of their side-chains shielded from solvent by interaction with another subunit. The interface properties including solvent-buried surface area, surface electrostatic potential and hydrophobicity, polar, nonpolar interactions, and hydrogen bonds have been compared in detail. Some results are shown in Table III.

The CdtA-CdtC interface is centered on a relatively flat, globular protein–protein interface with extensive nonglobular polypeptide interactions formed by the C-terminal tails of each subunit. Overall, the potential of the interface is neutral and dominated by nonpolar interactions (~70%). The residues involved in the globular interaction interface in *H. ducreyi* include Met81, Gln83, Gly85, Pro103, Asn104, Ile105, Try106, Try211, Leu212, Ala213, Pro214, and Pro215 in CdtA, and Arg43, Leu45, Leu46, Ala48, Phe141, Val153, Leu160, Pro168, and Phe170 in CdtC, but only residues Met81, Gly85 (CdtA) and Arg43, and Pro168 (CdtC) are generally conserved among the CDT family (*C. jejuni* lacks methionine at this position and possesses instead a leucine residue). The side-chain of Arg43 is stretched out at the bottom of the groove and forms hydrogen bonds with CdtA Gln83 (2.85 Å with a backbone oxygen) and Tyr211 (3.16 Å with the hydroxyl group of the side-chain). CdtA Met81 makes extensive vdW contacts with residues Arg43, Leu45, and Ala48 in CdtC [Fig. 7(a)]. These interactions are generally conserved in other CDTs. In *E. coli* II and III, the corresponding arginine residue makes two hydrogen bonds with CdtA, but only one hydrogen bond is formed in *E. coli* I and *C. jejuni* due to the presence of a nonpolar residue in position Tyr211.

Analysis of the nonglobular polypeptide–protein interaction shows that an intermolecular β-sheet interaction comprising the CdtA C-terminal tail with the β-sheet of β6/β7 in CdtC is remarkably conserved in the CDT family [Fig. 7(b)]. Extensive polar and nonpolar interactions, as well as a number of hydrogen bonds, are formed in this interfacial β-sheet region. MD simulations show that this H-bonding network is well maintained among CDTs. Table IV lists the hydrogen bonds of HdCDT with their percentages of occurrence during 2-ns MD simulations. Compared to the high stability at the terminal β-sheet region, the H-bonds involved in the globular CdtA–CdtC interface, such as Arg43 with Gln83 and Tyr211, are varied, with relatively short duration along the trajectory, indicating that the globular interface, especially the groove region, is more dynamic. As noted in the sequence analysis, β-sheet β6/β7 is the most highly conserved region in CdtC (up to 80% similarity vs the 30% overall in the CDT family), and the C-terminus of CdtA from different CDTs is generally hydrophobic. Given the divergence of sequence homology among CdtA and CdtC subunits, such a highly conserved interfacial interaction appears evolutionarily

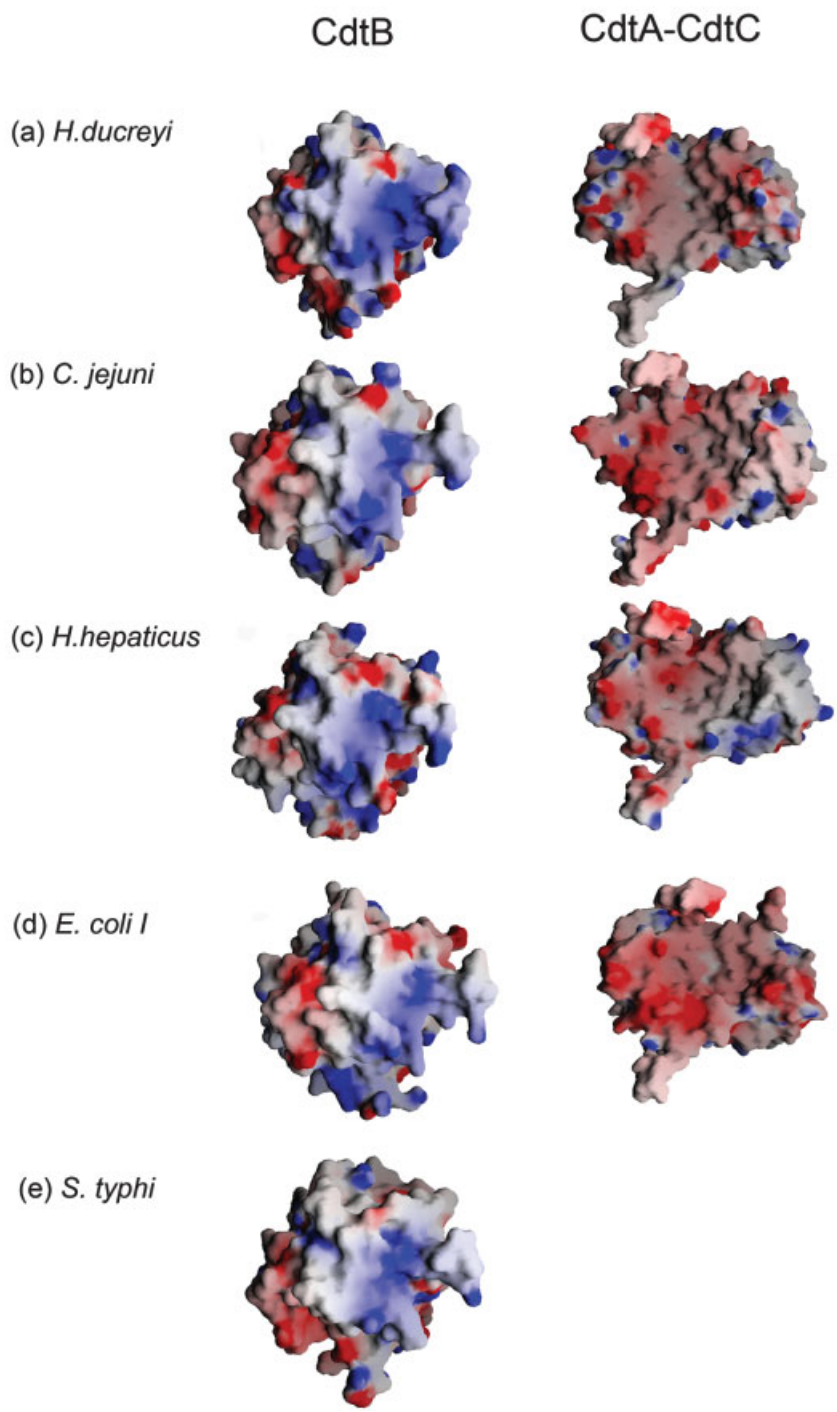

Fig. 8. Surface representations of the protein–protein interaction interface of CdtB and CdtA–CdtC subcomplex from (**a**) *H. ducreyi*, (**b**) *H. hepaticus*, (**c**) *E. coli* I, (**d**) *E. coli* III, and (**e**) *S. typhi*. The surfaces are colored by electrostatic potential (blue, positive; red, negative), calculated using GRASP.[37]

significant, suggesting that it plays an crucial role in the formation of CdtA–CdtC subcomplex.

CdtB Interaction Interface

The interface of CdtB with CdtA and CdtC subunits is the most important interactions associated with the assembly/disassembly of CDT holotoxin. A total of more than 4000 Å^2 of solvent-accessible surface area is buried in the formation of the tripartite complex. Similar to the CdtA–CdtC interface, a large number of both polar and nonpolar hydrophobic interactions involved in the CdtB interface, and a total of about 20–30 H-bonds are formed in different CDTs. The electrostatic potential of the interface for *H. ducreyi*, *H. hepaticus*, *E. coli*, and *S. typhi* are shown in

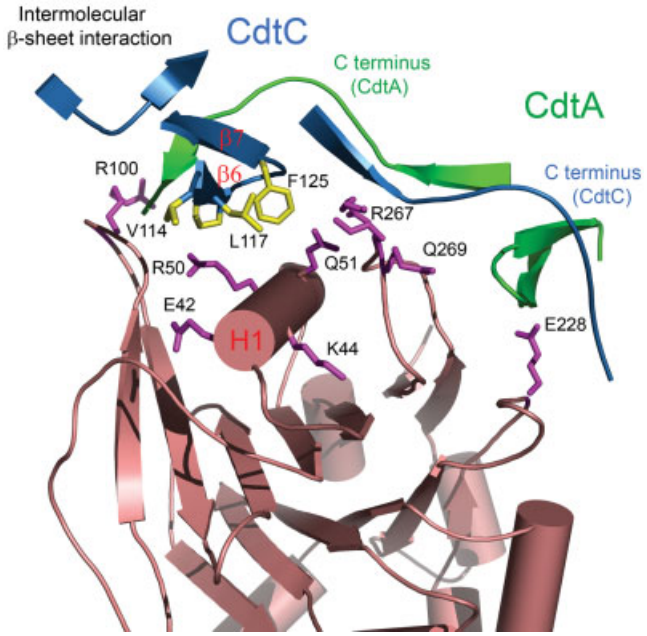

Fig. 9. Illustration of interfacial interactions between CdtB and CdtA–CdtC. Structural elements of α-helix H1 (CdtB) and β-sheet β6/β7 are indicated. The side-chain atoms of key residues involved in the interaction are magenta (CdtB) and yellow (CdtC).

Figure 8. A stretch with positively charged potential can be seen along the interface of CdtB, with several neutral bulges protruding out. Correspondingly, the interface on CdtA–CdtC is generally acidic, with a hydrophobic core in the middle. Indeed, mapping of interface residues in different CdtB enzymes shows that the positively charged patch including Lys44, Arg50, Arg100, Arg267 (Hd numbering) is highly conserved within the CDT family.

CdtB interacts with CdtA and CdtC mainly through α-helix H1 (41–54, Hd number) and a large loop (261–272) following β13 (Fig. 9). The loop extends to the interface of CdtA–CdtC, and is the major interacting element with both CdtA and CdtC. The residues Arg267 and Gln269 are generally conserved in the CDT family, forming hydrogen bonds and polar interactions with CdtA and CdtC. The major interfacial region of CdtC, the β-sheet of β6/β7, lies right above the α-helix H1 of CdtB. The stabilizing forces within this interface are dominated by nonpolar and hydrophobic interactions (conserved residues Val114, Leu117, Pro119, and Phe125 in the β-sheet), as well as a hydrogen bonding network formed by conserved residues CdtB Glu42, Lys44, Arg50, Gln51, and Arg100 with residues in CdtC. Consequently, CdtB assembles with CdtC in a highly conserved manner and appears to be more extensive and strong compared to CdtB–CdtA interaction. Of note, the β-sheet of β6/β7 is the most highly conserved region in CdtC, and also participates in the previously described intermolecular β-sheet interaction with CdtA. Therefore, the β-sheet of β6/β7 represents a key element in the assembly of the tripartite CDT complex. In agreement with mutagenesis data, a mutant CdtC subunit with deletion of β6/β7 sheet in *C. jejuni* (amino acids 115–136) was found to be unable to interact with CdtB and CdtC to reconstitute an active holotoxin.[17]

The absence of the CdtA and CdtC proteins in *S. typhi* CDT encourages an analysis of the structural properties of this enzyme in relation to other members of CdtB. Consistent with the experimental observation that StCdtB can form an active CDT holotoxin with *C. jejuni* CdtA and CdtC,[22] our structural analysis reveals that StCdtB is remarkably similar to other bacteria CdtB proteins in terms of surface shape and charge distribution on the interaction interface (Fig. 8). Comparison of the modeled complex of StCdtB with CjCdtA and CjCdtC shows that the interactions involved in the protein–protein interface, especially the conserved arginine patch participating in a H-bond network, is generally the same as those in CjCDT complex. However, structural differences are also observed between StCdtB and other CdtB subunits in the tripartite complex. An important one, the loop following β13 (residues 250–257, St numbering), which governs the interactions with both CdtA and CdtC, is relatively short in *S. typhi* due to the lack of six residues in this region compared to other bacterial CDTs. Therefore, the significance of this loop to the assembly of CDT holotoxin may be worth further experimental investigation.

CONCLUSIONS

We have constructed the 3D structures of CDTs from several bacterial pathogens based on the homology template of *H. ducreyi* and performed a comparative analysis to probe the structural characteristics and functions of the CDT family. Our modeling studies have shown that a number of structural features are remarkably conserved in the diverse bacterial CDTs that imply important functional consequences. In particular, the high conservation of key residues for DNA binding and catalysis indicate that the active subunit of the CdtB functions in highly similar fashion in all CDTs. The N-terminus of CdtC binding to the CdtB active site may represent a common autoinhibitory mechanism for the CDT holotoxin. The close interplay of CdtA and CdtC in the formation of the groove and aromatic patch, and the similarity in their positioning with the two lectin repeats in the ricin B-chain, suggest that these two components of CDT work together for a related function in cell surface binding and internalization of the holotoxin. In addition, the CDT ternary complex possesses a number of globular protein–protein interfaces and terminal polypeptide–protein interactions. A highly conserved intermolecular β-sheet interaction is found to play an important role in the formation of CdtA and CdtC. Despite the generally low level of sequence similarity in the lectin subunits, the β-sheet β6/β7 in CdtC is the most conserved element from sequence and structure–function considerations. Taken together, these studies shed light on the assembly and function of this important bacterial genotoxin family, and furthermore, may provide targets for therapeutic design.

REFERENCES

1. Pickett CL, Whitehouse CA. The cytolethal distending toxin family. Trends Microbiol 1999;7:292–297.

2. De Rycke J, Oswald E. Cytolethal distending toxin (CDT): a bacterial weapon to control host cell proliferation? FEMS Microbiol Lett 2001;203:141–148.
3. Lara-Tejero M, Galan JE. Cytolethal distending toxin: limited damage as a strategy to modulate cellular functions. Trends Microbiol 2002;10:147–152.
4. Thelestam M, Frisan T. Cytolethal distending toxins. Rev Physiol Biochem Pharmacol 2004;152:111–133.
5. Friedman CR, Neimann J, Wegener HC, Tauxe RV. Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialized nations. 2nd ed. Washington, DC: American Society for Microbiology; 2000.
6. Pang T, Bhutta ZA, Finlay BB, Altwegg M. Typhoid-fever and other salmonellosis—a continuing challenge. Trends Microbiol 1995;3:253–255.
7. Trees DL, Morse SA. Chancroid and *Haemophilus ducreyi*: an update. Clin Microbiol Rev 1995;8:357–375.
8. Wilson M, Henderson B. Virulence factors of *Actinobacillus actinomycetemcomitans* relevant to the pathogenesis of inflammatory periodontal diseases. FEMS Microbiol Rev 1995;17:365–379.
9. Sullivan P, Clark WL, Kaiser PK. Bilateral endogenous endophthalmitis caused by HACEK microorganism. Am J Ophthalmol 2002;133:144–145.
10. Purdy D, Buswell CM, Hodgson AE, McAlpine K, Henderson I, Leach SA. Characterisation of cytolethal distending toxin (CDT) mutants of *Campylobacter jejuni*. J Med Microbiol 2000;49:473–479.
11. Shenker BJ, McKay T, Datar S, Miller M, Chowhan R, Demuth D. *Actinobacillus actinomycetemcomitans* immunosuppressive protein is a member of the family of cytolethal distending toxins capable of causing a G(2) arrest in human T cells. J Immunol 1999;162:4773–4780.
12. Young VB, Knox KA, Pratt JS, Cortez JS, Mansfield LS, Rogers AB, Fox JG, Schauer DB. In vitro and in vivo characterization of *Helicobacter hepaticus* cytolethal distending toxin mutants. Infect Immun 2004;72:2521–2527.
13. Pickett CL, Cottle DL, Pesci EC, Bikah G. Cloning, sequencing, and expression of the *Escherichia coli* cytolethal distending toxin genes. Infect Immun 1994;62:1046–1051.
14. Scott DA, Kaper JB. Cloning and sequencing of the genes encoding *Escherichia coli* cytolethal distending toxin. Infect Immun 1994;62: 244–251.
15. Lara-Tejero M, Galan JE. CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. Infect Immun 2001;69:4358–4365.
16. Deng K, Hansen EJ. A CdtA–CdtC complex can block killing of HeLa cells by *Haemophilus ducreyi* cytolethal distending toxin. Infect Immun 2003;71:6633–6640.
17. Lee RB, Hassane DC, Cottle DL, Pickett CL. Interactions of *Campylobacter jejuni* cytolethal distending toxin subunits CdtA and CdtC with HeLa cells. Infect Immun 2003;71:4883–4890.
18. Lara-Tejero M, Galan JE. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. Science 2000;290: 354–357.
19. Elwell CA, Dreyfus LA. DNase I homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest. Mol Microbiol 2000;37:952–963.
20. Nesic D, Hsu Y, Stebbins CE. Assembly and function of a bacterial genotoxin. Nature 2004;429:429–433.
21. Cortes-Bratti X, Chaves-Olarte E, Lagergard T, Thelestam M. Cellular internalization of cytolethal distending toxin from *Haemophilus ducreyi*. Infect Immun 2000;68:6903–6911.
22. Haghjoo E, Galan JE. *Salmonella typhi* encodes a functional cytolethal distending toxin that is delivered into host cells by a bacterial-internalization pathway. Proc Natl Acad Sci USA 2004; 101:4614–4619.
23. Frisan T, Cortes-Bratti X, Thelestam M. Cytolethal distending toxins and activation of DNA damage-dependent checkpoint responses. Int J Med Microbiol 2002;291:495–499.
24. Alby F, Mazars R, de Rycke J, Guillou E, Baldin V, Darbon JM, Ducommun B. Study of the cytolethal distending toxin (CDT)-activated cell cycle checkpoint: involvement of the CHK2 kinase. FEBS Lett 2001;491:261–265.
25. Cortes-Bratti X, Frisan T, Thelestam M. The cytolethal distending toxins induce DNA damage and cell cycle arrest. Toxicon 2001;39: 1729–1736.
26. Peres SY, Marches O, Daigle F, Nougayrede JP, Herault F, Tasca C, De Rycke J, Oswald E. A new cytolethal distending toxin (CDT) from *Escherichia coli* producing CNF2 blocks HeLa cell division in G2/M phase. Mol Microbiol 1997;24:1095–1107.
27. Ohara M, Hayashi T, Kusunoki Y, Miyauchi M, Takata T, Sugai M. Caspase-2 and caspase-7 are involved in cytolethal distending toxin-induced apoptosis in Jurkat and MOLT-4 T-cell lines. Infect Immun 2004;72:871–879.
28. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 2000;28:45–48.
29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
30. Thompson JD, Higgins DG, Gibson TJ. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
31. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
32. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M. Evaluation of comparative protein modeling by MODELLER. Proteins 1995;23:318–326.
33. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins 1993;17:355–362.
34. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck—a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
35. Guex N, Diemand A, Peitsch MC. Protein modelling for all. Trends Biochem Sci 1999;24:364–367.
36. Vriend G. What If—a molecular modeling and drug design program. J Mol Graph 1990;8:52–55.
37. Nicholls A, Sharp KA, Honig B. Protein folding and association—insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
38. Weston SA, Lahm A, Suck D. X-ray structure of the DNase I-d(GGTATACC)2 complex at 2.3 Å resolution. J Mol Biol 1992;226: 1237–1256.
39. SYBYL Molecular Modelling Software, v7.0. St. Louis, MO: Tripos Associates; 2004.
40. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WS, and Kollman PA. AMBER 8, San Francisco: University of California; 2004.
41. Ryckaert JP, Ciccotti G, Berendsen JC. Numerical integration of the Cartesian equation of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
42. Darden T, York D, Pedersen L. Particle mesh Ewald—an N.Log(N) method for Ewald sums in large systems. J Chem Phys 1993;98: 10089–10092.
43. Berendsen HCJ, Postma JPM, van Gunsteren WF, DiNole A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
44. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999; 15:305–308.
45. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Bioph Biom 2000;29:291–325.
46. Baker D, Sali A. Protein structure prediction and structural genomics. Science 2001;294:93–96.
47. Sanchez R, Sali A. Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. Proc Natl Acad Sci USA 1998; 95:13597–13602.
48. John B, Sali A. Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucleic Acids Res 2003;31:3982–3992.
49. Montfort W, Villafranca JE, Monzingo AF, Ernst SR, Katzin B, Rutenber E, Xuong NH, Hamlin R, Robertus JD. The 3-dimensional structure of ricin at 2.8-Å. J Biol Chem 1987;262:5398–5403.
50. Hazes B. The (QxW)(3) domain: a flexible lectin scaffold. Protein Sci 1996;5:1490–1501.
51. Delano WL. The PyMOL molecular graphics system. San Carlos, CA: Delano Scientific; 2002.
52. Shenker BJ, Besack D, McKay T, Pankoski L, Zekavat A, Demuth DR. Induction of cell cycle arrest in lymphocytes by *Actinobacillus actinomycetemcomitans* cytolethal distending toxin requires three subunits for maximum activity. J Immunol 2005;174:2228–2234.