J Mol Evol (2017) 84:174–186
DOI 10.1007/s00239-017-9791-y

ORIGINAL ARTICLE

Molecular Evolutionary Constraints that Determine the Avirulence State of *Clostridium botulinum* C2 Toxin

A. Prisilla[1] · R. Prathiviraj[1] · P. Chellapandi[1]

Received: 1 June 2016 / Accepted: 30 March 2017 / Published online: 5 April 2017

Abstract *Clostridium botulinum* (group-III) is an anaerobic bacterium producing C2 toxin along with botulinum neurotoxins. C2 toxin is belonged to binary toxin A family in bacterial ADP-ribosylation superfamily. A structural and functional diversity of binary toxin A family was inferred from different evolutionary constraints to determine the avirulence state of C2 toxin. Evolutionary genetic analyses revealed evidence of C2 toxin cluster evolution through horizontal gene transfer from the phage or plasmid origins, site-specific insertion by gene divergence, and homologous recombination event. It has also described that residue in conserved NAD-binding core, family-specific domain structure, and functional motifs found to predetermine its virulence state. Any mutational changes in these residues destabilized its structure–function relationship. Avirulent mutants of C2 toxin were screened and selected from a crucial site required for catalytic function of C2I and pore-forming function of C2II. We found coevolved amino acid pairs contributing an essential role in stabilization of its local structural environment. Avirulent toxins selected in this study were evaluated by detecting evolutionary constraints in stability of protein backbone structure, folding and conformational dynamic space, and antigenic peptides. We found 4 avirulent mutants of C2I and 5 mutants of C2II showing more stability in their local structural environment and backbone structure with rapid fold rate, and low conformational flexibility at mutated sites. Since, evolutionary constraints-free mutants with lack of catalytic and pore-forming function suggested as potential immunogenic candidates for treating *C. botulinum* infected poultry and veterinary animals. Single amino acid substitution in C2 toxin thus provides a major importance to understand its structure–function link, not only of a molecule but also of the pathogenesis.

Keywords Binary toxin A · ADP ribosyltransferase · Site-directed mutagenesis · Molecular dynamics · Coevolution

Introduction

Clostridium botulinum is a spore-forming anaerobic bacterium causing avian and mammalian botulism outbreaks. Physiological group-III strains of *C. botulinum* are capable of producing C2 and C3 toxins, in addition to botulinum neurotoxin serotypes C and D (Moriishi et al. 1991, 1993). C2 and C3 toxins are belonging to ADP-ribosylation superfamily or bacterial ADP-ribosyltransferases (BADPRTs), which is categorized into four types designated I–IV (Holbourn et al. 2005). Type I consists of heat-labile enterotoxin A, cholera toxin A, and pertussis toxin S. Type II includes diphtheria toxin A and exotoxin A and Type III encompasses C3-like exotoxins. Type IV contains *C. botulinum* C2 toxin, *C. perfringens* ι-toxin, *C. difficile* Cdt toxin, *C. spiroforme* toxin, and *Bacillus cereus* vegetative insecticidal protein 2 (VIP2). Type IV is commonly known as binary toxin A (BTA) that transfers an ADP ribose moiety of NAD^+ to G-actin and F-actin (Tsuge et al. 2008; Wiegers et al. 1991).

Electronic supplementary material The online version of this article (doi:10.1007/s00239-017-9791-y) contains supplementary material, which is available to authorized users.

✉ P. Chellapandi
pchellapandi@gmail.com

[1] Molecular Systems Engineering Lab, Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli, Tamil Nadu 620 024, India

C2 toxin consists of an enzymatic component C2I and a binding/translocation component C2II (Ohishi et al. 1980). C2I composes of an adapter domain and a catalytic domain related with VIP2 and iota toxin a (Ia) (Aktories and Barth 2004; Tsuge et al. 2003). In C2 toxin, a functional motif serine–threonine–serine sequence (STS), a α1-helical structure and folding of a core β-strands surrounded by α-helices are highly conserved to maintain its structure–function link similar to other BTA family members (Takada et al. 1995). Glu187 is a catalytic residue in C2I that deprotonates Arg177 in G-actin. Glu189 forms H-bond with the O′2 on nicotinamide ribose that stabilizes the oxocarbenium cation, leading to α-selective ADP-ribosylation (Barth et al. 1998; Domenighini et al. 1994; Domenighini and Rappuoli 1996; Fujii et al. 1996; Jank and Aktories 2013). C2II composes of five domains ($C2II_{20}$, D1–D4) in which D1–D3 consist of C2I binding components evolutionarily related to protective antigen (PA) of anthrax toxin and iota toxin b (Ib) (Barth et al. 2004; Leppla 1995; Schleberger et al. 2006). D4 domain structure has no similarity with D4 of PA and Ib (Blöcker et al. 2000). D4 contributes to stabilize a conformation needed for receptor binding and cytotoxicity (Blöcker et al. 2000; Varughese et al. 1999). Glu307 is the only negatively charged amino acid sharing its function of pore-forming and C2I–C2IIa complex formation, which phylogenetically close with PA (Blöcker et al. 2003a, b; Benson et al. 1998; Nagahama et al. 2003; Petosa et al. 1997).

After proteolytic cleavage, a residual fragment C2IIa released from C2II binds with C2I and then recognizes on the cell surface receptors (Aktories and Barth 2004; Blöcker et al. 2000; Eckhardt et al. 2000). The uptake of C2 toxin is mediated by dynamin-dependent pathways and endocytosed by processes dependent on clathrin and RhoA (Pust et al. 2010). C2IIa heptamer is inserted into the membrane and forms pores in a late endosome under acidic pH and then bound C2I is translocated into the cytosol (Blöcker et al. 2003a, b). C2I catalyzes the ADP-ribosylation at the residue Arg177 of G-actin and exhibits to reduce the ability of actin to undergo polymerization, leading to disruption of the cytoskeleton architecture (Sterthoff et al. 2010; Tsuge et al. 2008).

Site-mutagenesis studies evidenced its substrate specificity, recognition of NAD-binding core, and channel/pore-forming property (Barth et al. 1998; Blöcker et al. 2003a, b; Fujii et al. 1996; Han and Tainer 2002; Lang et al. 2008; Schleberger et al. 2006). Conserved residues contribute to the overall stability of a protein structure and function (Kowarsch et al. 2010; Vitkup et al. 2003). Evolutionary constraints on a disease-related gene are relaxed allowing it to accumulate mutation after gene duplication. Physical and evolutionary restraints also signify the associations between patterns of amino acid replacement and protein structure (Han et al. 1999). Such constrictions reflect on the substitution rate of residues in different secondary structural environments and of different solvent accessibilities (Tseng and Liang 2006). Thus, evolutionary forces determine its structure–function integrity, avirulence state and disease-susceptibility (Chellapandi et al. 2013; Chellapandi 2014; Prathiviraj et al. 2015; Xia and Levitt 2002).

Enzymatic inactive mutants from C2 toxin (Fahrer et al. 2010a, b) and gene-mediated adjuvants from *Escherchia coli* heat-labile toxin (LTB) (Cunha et al. 2014; Gil et al. 2013; Krüger et al. 2013) and non-catalytic cholera toxin A (Wan et al. 2014) were recently developed for therapeutic applications. No vaccine has been designed or developed for C2 toxin-mediated diseases in poultry and veterinary animals to date. Thus, the present study was aimed to reveal how evolutionary forces enforcing on BTA family are useful measures to determine the functional diversity and avirulence states of C2 toxin. Finding structural constraints in C2 toxin would further improve the accuracy and reliability of our screening and selection of its avirulent toxins with increased immunogenic nature.

Materials and Methods

Analysis of Phylogenetic Diversity

Crystallographic structural information for C2I (PDB ID: 2J3V; UniProt: 069275) and C2II (PDB ID: 2J42; UniProt: 086171) were obtained from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) and corresponding protein sequences retrieved from the UniProt database (http://www.uniprot.org/). The similarity hits for these template sequences were searched from NCBI-non-redundant protein database by BlastP program (Altschul et al. 1997). Multiple sequence alignment was conducted with ClustalX 2.0.11 software (Thompson et al. 1997). Estimates of phylogeny for sequence alignments were obtained using the neighbor joining algorithm in the MEGA 5.05 software (Tamura et al. 2011) with 1000 bootstrapped alignments. The constructed trees were visualized and edited with MEGA.

Analysis of Functional Divergence

The type I (θ_I) and type II (θ_{II}) functional divergence coefficients along with the rate for gamma distributions (α) were examined across the BTA family by DIVERGE 1.04 (Gu and Vander Velden 2002). Phylogenetic functional divergence and its parameters of this family were analyzed by SplitsTree 4.0 software using BioNJ algorithm (Huson and Bryant 2006). Recombination events and mutation rate were detected by RDP 3.0 software

using Recomb2007 algorithm (Martin et al. 2010). Hierarchical network parameter for detecting recombination events was computed by CYTOSCAPE 3.0.1 software (Shannon et al. 2003).

Genetic Diversity and Darwinian Selection

Evolutionary genetic analyses (*R, I, d, S, π, D*) were performed with programs in the MEGA. Gene and amino acid substitution patterns across this family were calculated by the maximum composite likelihood method. The purpose of Tajima's test was to identify sequences which do not fit the neutral theory model of equilibrium between mutation and genetic drift (Tajima 1989). Sites subject to Darwinian selection and evolutionary rates were calculated by a selective strength computed from Ka/Ks ratio and dN/dS ratio (ω). The Ka/Ks ratio was calculated with Ka/Ks calculator from different evolutionary models (Zhang et al. 2006). Codon-based likelihood in synonymous and non-synonymous substitutions was used to estimate dN/dS ratios (ω) using HyPhy 1.0 program (Pond et al. 2005).

Estimation of Evolutionary Rate

A similarity scoring matrix calculated from estimated values was detected a protein homolog which was evolutionarily related to C2 toxin as described earlier (Lio and Goldman 1999; Pearson 1998). Global alignment of two protein sequences was carried out by EMBOSS ALIGN tool using different PAM matrices with the Needleman–Wunsch algorithm (http://www.ebi.ac.uk/Tools/psa/). Evolution rate was estimated between C2 toxin and its homologs based on evolutionary relatedness and sequence similarity between amino acids as reflected in substitutions matrixes (Udaya Prakash et al. 2010).

Identification of Mutable Residues

The crucial residues involving in structural and functional stability of this toxin were identified from the literature. Conserved motifs in the sequences were identified from the NCBI conserved domain database (Marchler-Bauer et al. 2015). The mutable residues that bring changes in the function of C2 toxin were predicted from its structure by HotSpot Wizard (Pavelka et al. 2009). Rate4Site 2.01 program was employed to estimate the evolution rate of detecting mutable residues from a probability of its multiple sequence alignment (Mayrose et al. 2004).

Analysis of Coevolution Sites in Local Structural Environment

Secondary structural elements of this toxin and its features including solvent accessibility and H-bonding were predicted with JOY server (Mizuguchi et al. 1998). A coevolution (coupled mutation) site on the local structural environment was inferred from its sequence by a CMAT program with profile-based adjustment and mutual information (Jeong and Kim 2012). Network parameter was calculated for detecting coevolved pairs by CYTOSCAPE 3.0.1.

Analysis of Structural Stability upon Point Mutation

CUPSAT was used to predict the mutant residues that stabilize its structure upon point mutation and to assess the amino acid environment of the mutation site based on amino acid-atom potentials (Parthiban et al. 2006). Structural stability and virulence state of resulted mutants were predicted with SDM server using a statistical potential energy function (Worth et al. 2011). The structural stability score was calculated from the energy function by using environment-specific aminoacid substitution frequencies within homologous protein families.

Analysis of Structural Constraints in Avirulent Toxins

The protein folding rate of all structural classes in C2 toxin and its avirulent mutants was predicted from the sequences by FoldRate program using multiple regression technique (Gromiha et al. 2006). Avirulent mutants with more structural stability were chosen by capturing near-native ensembles of their backbone conformations using Backrub Ensemble (Smith and Kortemme 2008). Structural flexibility in avirulent toxins was analyzed by PredyFlexy server using a structural alphabet and combining classical X-ray B-factor data (De Brevern et al. 2012).

Analysis of Immunogenic Nature of Avirulent Toxins

Using the potential hydrophilic regions, immunogenic nature (antibody epitopes) of avirulent toxins was predicted from their amino acid sequences. The hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptide chains were selected as parameters to predict the location of continuous antibody epitopes. Antibody epitope prediction tool in the Immune Epitope Database (http://tools.immuneepitope.org/bcell/) was used to predict the sites that produce an immunogenic response against C2 toxin and its avirulent toxins, according to a semi-empirical method described by Kolaskar and Tongaonkar (1990) (Kolaskar and Tongaonkar 1990).

Results

Molecular Phylogeny of C2 Across the BTA Family

A dataset for the BADPRTs superfamily consists of 28 complete sequences, related to C2 toxin, obtained from different family members (Supplementary Fig. S1). C2I and C3 are related together and shared their phylogenetic resemblance with Ia and pertussis toxin S5. A phylogenetic proximity is found between C2II and LTB. Cholix toxin is a unique protein and corresponds with C3-like exotoxins. A dataset for inferring phylogenetic relationship of BTA family comprises 35 sequences whose crystallographic structures are to be solved to date (Fig. 1). C2I, C2II, and C3 are clustered separately in the phylogenetic tree in which C2I is evolutionarily related to Ia and VIP2 whereas C2II is similar to Ib. This suggested that both components of C2 and iota toxins share a common evolutionary origin. In contrast, C3 toxin shows a phylogenetic closeness to Toxin2A and EDIN (Epidermal cell differentiation inhibitor) of this family, but not related to C2 toxin as such.

Functional Divergence of C2 Across the BTA Family

Our results described that all sites evolve at a roughly same rate across C2II/Toxin2A cluster and many sites evolve slowly, but a few evolve rapidly among clusters (Table 1). Apart from a cluster of C2II/Toxin2A, a strong rate of heterogeneity among sites is detected across other homologous clusters. Type I functional divergence after gene duplication results in altered functional constraints with different evolutionary rate between duplicate genes of C2II/

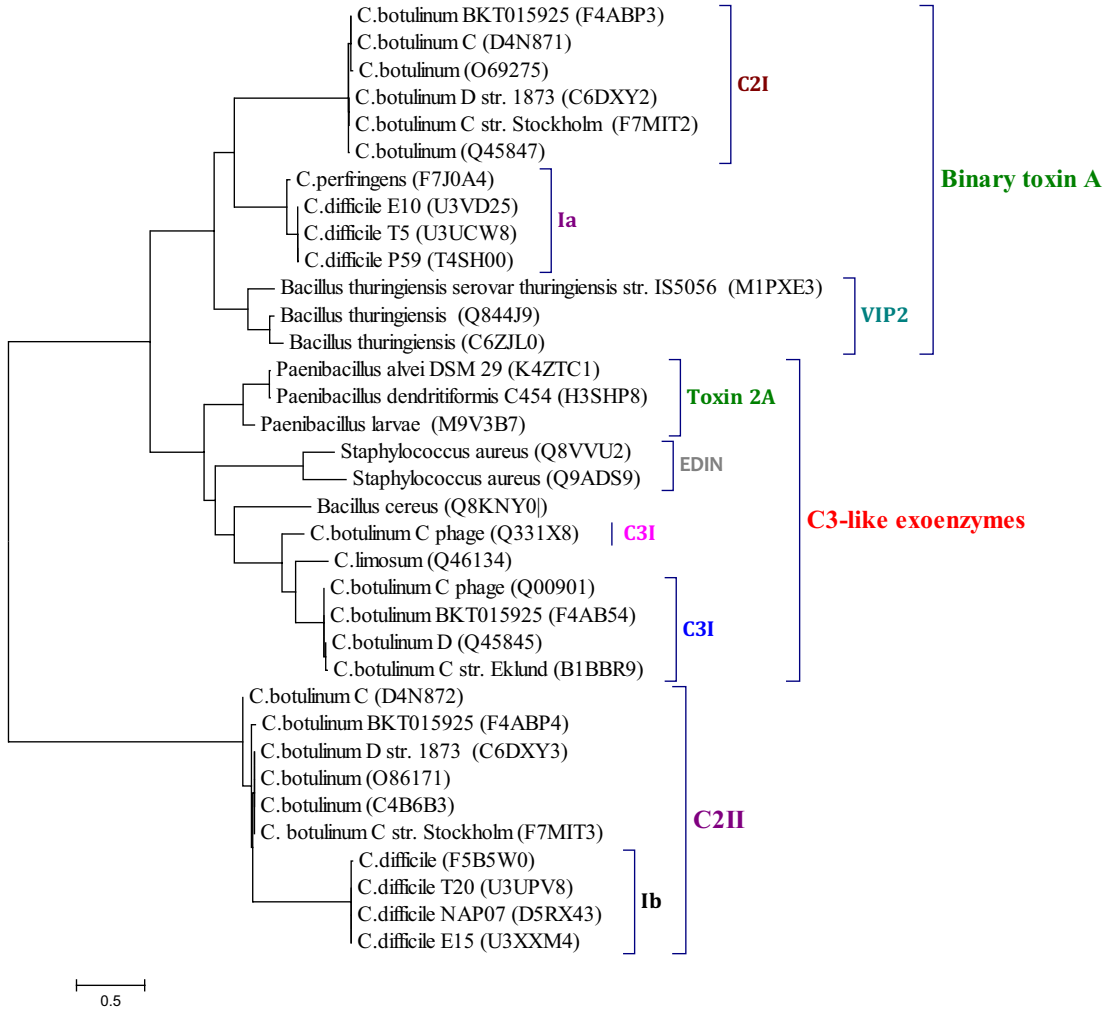

Fig. 1 Phylogenetic tree of *Clostridium botulinum* C2 toxin reconstructed from the similarity sequences of available crystallographic structures. Branch lengths are proportional to evolutionary distances. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed

Table 1 Coefficients of functional divergence between homologous clusters of binary toxin A family, estimated by diverge

Cluster	θ_I	θ_{II}	α_1	α_2
C2II/VIP2	0.628	0.564	0.570	–
C2II/C3	–	0.554	−12.34	0.651
C2II/Toxin2A	1.0	0.455	2.078	–
C2I/Ia	1.0	–	−12.01	–
C2I/C3	1.0	1.0	−9.509	−18.83
VIP2/C3	–	0.520	−16.01	−0.642
VIP2/Toxin2A	0.876	0.388	1.306	–
Ia/C3	1.0	1.0	−13.78	−20.97

θ_I and θ_{II} are the coefficients of type I and type II functional divergence, respectively. The α_1 is the gamma shape parameter for rate variation among sites in type I divergence and α_2 for type II divergence

Toxin2A cluster and C2I/Ia cluster. No altered functional constraints, but radical change in amino acid property is detected between duplicated genes of C2II/C3 and VIP2/C3 clusters that may be arising due to type II functional divergence. Neutral selection is imposed on the functional divergence across the clusters of C2II/VIP2, C2I/C3, VIP2/Toxin2A, and Ia/C3.

Recombination Rate of C2 Across the BTA Family

Evolutionary preference for recombination robustness is much stronger than mutational robustness (Supplementary Table S1; Fig. S2). Recombination analysis of our study predicted two major recombination networks resulting across this family, where the C2II gene in *C. botulinum* type C may be acquired from the C2II gene of D phage through C3I gene of C phage by horizontal gene transfer event. A gene encoding Bcer toxin from *B. cereus* has served as a major parent to combine many segregation sites in directing the recombination process in the evolution of C2, C3, C3stau2, Ia, and EDIN. It clearly implied that C2 and C3 toxins might be evolved separately for their unique functions from those phages or bacterial origins at different evolution rates via a short recombination path.

Evolution Rate of C2 Across the BTA Family

We have employed different substitution matrices (10–500) to calculate its evolution rate from homologs as shown in Fig. 2. As a result of substitution score, a strong sequence conservation is observed between C2I/C2II; C2II/Ib; C3I/C3II. It reflected that evolution rate and acceptable substitution mutations are lower than as anticipated in other toxins. The highly acceptable mutations are also detected between C2I/Ia; C2II/VIP2; C3II/Toxin2A. Substitution rate and sequence conservation are predicted to be low

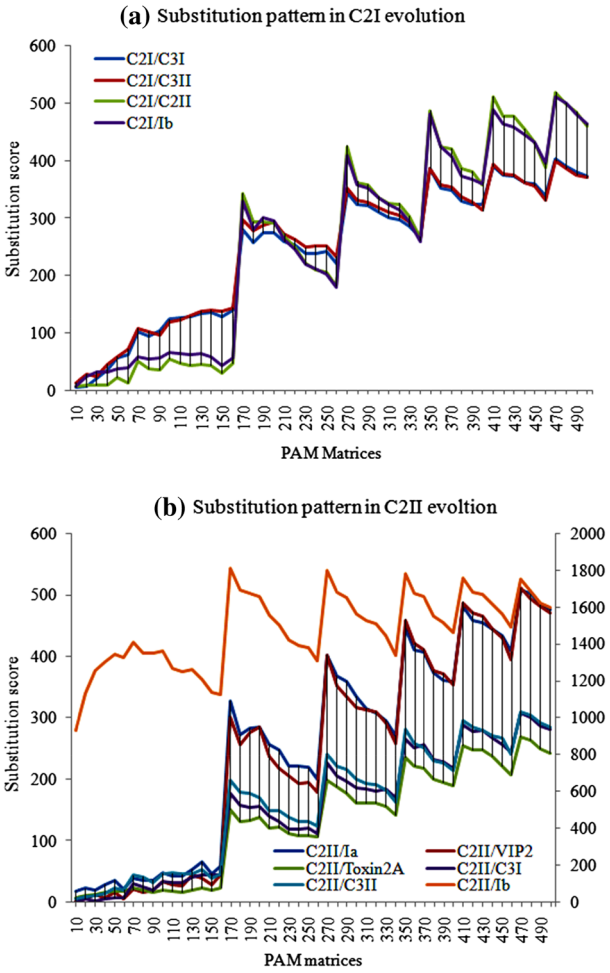

Fig. 2 Amino acid substitution patterns for comparison of *Clostridium botulinum* C2 toxin with C3 toxin and binary toxin A family

between C2I/C3I; C2II/C3II; C2II/Toxin2A. Moreover, C2 shares its residue substitutions within the iota toxin, whereas residues in C3 substitute with VIP2 and Toxin2A.

Genetic Diversity of C2 Across the BTA Family

A site-specific rate change is related to functional divergence during protein evolution. Genetic diversity analysis shows that gene (π 0.479) and protein (π 0.452) sequences are diverged within the BTA family at a slow recombination/mutability rate (Table 2). Phylogenetic distance for gene diversity (d 3.254; R 1.11) is higher than protein diversity (d 2.929). Darwinian positive selection acting on gene function (D 4.3) is rather than on protein function (D 3.3). The synonymous substitution rate of C2 is higher than non-synonymous substitution rate to select and purify its function from BTA family. However, the selective strength (Ka/Ks; Ω) supports the fitness of the C2 function as such by neutral evolution.

Table 2 Estimates of genetic diversity and Darwinian selection for binary toxin A family

Genetic parameters	
Protein diversity	
Phylogenetic distance (*d*)	2.929
Invariant sites (+I)	0.969
Phylogenetic diversity	4.425
Number of segregating sites (*S*)	1151
Nucleotide/amino acid diversity (π)	0.452
Tajima test statistic (*D*)	3.306
Selective strength (Ka/Ks)	0.785
Gene diversity	
Transition/transversion ratio (*R*)	1.11
Phylogenetic distance (*d*)	3.254
Recombination/mutability rate (Δ)	0.07
Number of segregating sites (*S*)	3022
Nucleotide diversity (π)	0.479
Tajima test statistic (*D*)	4.386
Non-synonymous substitution rate (d*N*)	5.223
Synonymous substitution rate (d*S*)	6.480
Selective strength (d*N* − d*S*; normalized Ω)	0.146

General time reversible with gamma distribution was found as an evolution model

Coevolution on Local Structural Environment of C2

Coevolution is significantly acted on the residues (64%) in helices of the secondary structure of C2I, which in turn, to change coevolved pairs as β-strands (Fig. 3). About 15% changes is accounted between residues in helices and residues in coils due to coupled residue mutations. No major adjustments in C2 are observed in other structural environment at secondary level. Coevolution sites (65%) are majorly accounted in residues involving in the solvent accessible surface area (SASA). Conversely, the local structural environment may be altered from buried surface area to SASA and vice versa (30%). About 38% of the residues is not contributed in the H-bonding. H-bonding (24%) and non-hydrogen bonding (25%) residues destabilize H-bonding frequency in coevoled pairs. Coevolution on H-bonding residues radically rearranges the H-bonding pattern with non-H-bonding pairs. As a result, the order of primary coevolution sites is detected as 393 > 348 > 39 > 29 > 16 > 73 from the coevolution network with shortest path of 4%.

Screening and Selection of Avirulent Toxins

Conserved domain analysis indicated that C2I has two functional motif regions such as [Y252]–[R265]–[R299]–

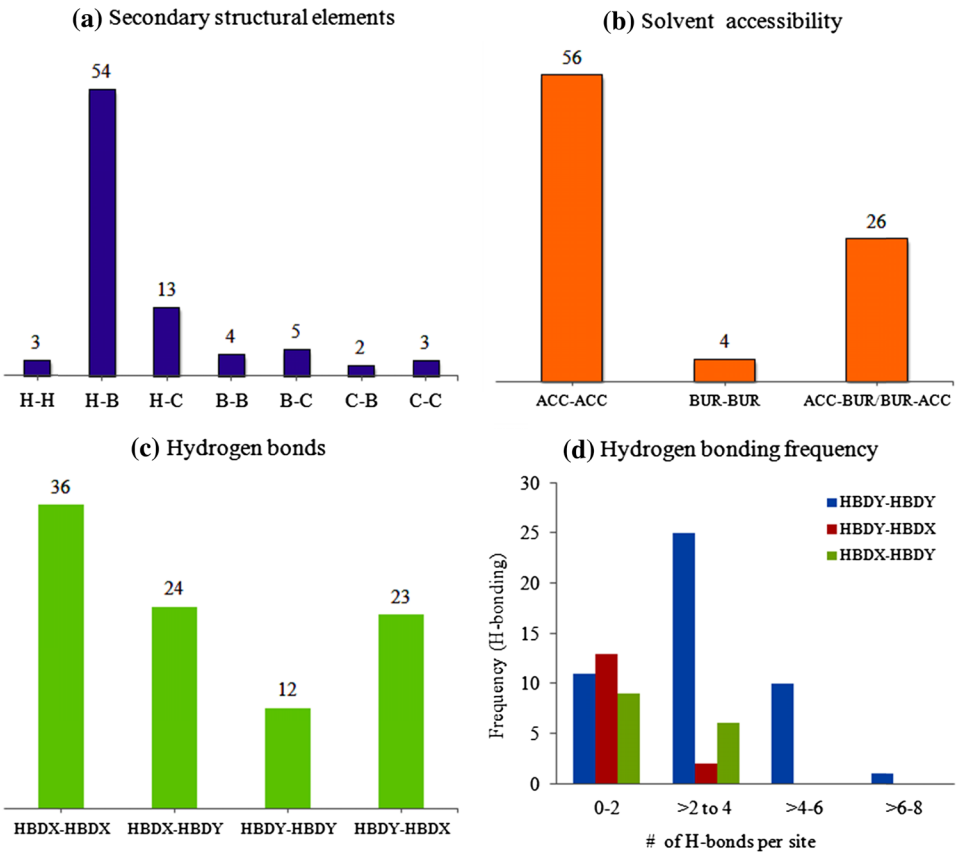

Fig. 3 Calculation of secondary structural types (**a**), solvent accessibility (**b**), H-bonds (**c**), and H-bonding frequency (**d**) for coevolved sites identified in the C2I structure. Within the coevolved site pairs, if both residues are buried or accessible, they are shown as 'BUR–BUR' or 'ACC–ACC', respectively. 'HBDY–HBDY' and 'HBDX–HBDX' indicate cases where both coevolved residues are involved or not involved in H-bonding correspondingly. 'HBDX–HBDY/HBDY–HBDX' indicates cases where at least one residue is involved in H-bonding. Values in the parenthesis show mean and standard error of estimated from the distribution of structural property values for randomly selected non-coevolved residue pairs. *H* α-Helix, *B* β-strand, *C* coil

[S348T349S350]–[S361]–[E389]–[F398] and [F384]–[E387]–[E389]. Arg is identified as a site for conformational flexibility of the ligand binding pocket. Evolution rate is calculated to be low for the residues selected for point mutation, suggested that these residues contribute a crucial role in its molecular function (Supplementary Fig. S3). We have screened 29 mutants from wild-type C2I and 21 mutants from C2II, which are detailed in Supplementary Tables S2, S3. Only 17 mutants from wild-type C2I and 16 mutants from C2II were selected as avirulent mutants with no radical changes in their secondary structural elements, but overall native-like structures are destabilized by point mutations. Finally, four C2I and five C2II avirulent mutants with more structural stability ($-\Delta\Delta G$) were selected for further evaluation (Table 3). It shows that the exchange of the STS motif in C2I causes a decrease in structural stability compared to E389K and F398Y mutants ($\Delta\Delta G$ −1.94 to −1.65 kcal/mol). Channel/pore-forming property of C2II is substantially changed in its mutants. F456S and F456A have shown more structural stability ($\Delta\Delta G$ −1.96 to −1.79 kcal/mol) without change in H-bonding patterns. Solvent accessibility of C2I and C2II avirulent mutants is drastically reduced as low as C2 toxin. The local structural environment in C2I avirulent mutants is stabilized by unsaturated H-bonding resulted from saturated H-bonding, but H-bonding pattern in C2II mutants is not destabilized as such.

Protein Fold Rate of Avirulent Toxins

The folding rate on structural classes, particularly a mixed class ($\alpha+\beta$ and α/β), of C2I, is decisively changed the folding process of its avirulent mutants (Fig. 4). The fast fold rate is detected in the mixed classes of T349H, E389K, and F398Y compared with C2I. The minor and major changes in fold rate are detected in all-β class and all-α class of C2I mutants, respectively. The fast fold rate is resulted in all-α class, which was higher than all-β class of its mutants. C2II mutants not including E399C and F456A accomplish the fast fold rate in many structural classes, and no major changes detected in the fold rate on all-α class. The fold rate on all structural classes of C2II is faintly lower than F456S and F456H.

Backbone Conformation of Avirulent Toxins

We generated different conformational ensembles of the C2 toxin carrying a single amino acid mutation, which are shifted from native proteins (Supplementary Fig. 4). The best backbone ensembles are captured for S350R (−413 kcal/mol) and F398Y (−411 kcal/mol) from C2I structure. E389K (−405 kcal/mol) and T349V (−400 kcal/mol) also yield the stable ensembles with insignificant conformational flexibility. Similar to C2I mutants, we obtained stable backbone conformations for F456H (1083), F456A (1089), and E399C (1064) from C2II structure (1076 kcal/mol). F456S (587 kcal/mol) has not shown significant near-native shifts similar to the C2II structure due to localized structural flexibility. It is suggested that C2II mutants apart from F456S are able to reduce the channel/pore-forming property.

Conformational Sub-Space of Avirulent Toxins

Structural fluctuations (Cα-RMSF) were calculated between C2 and its avirulent mutants to know whether any constraints in conformational sub-dynamic space at a mutated residue (Table 4; Supplementary Fig. S5, S6). S350R (0.188–0.266) and E389K (−0.328 to −0.438) are structurally deviated from C2I with small extent in

Table 3 Amino acid mutations that stabilize the overall structure of C2 avirulent toxins

Mutant	SASA	Wild-type		Mutant type		Potential energy
		SA (%)	HB	SA (%)	HB	$\Delta\Delta G$ (kcal/mol)
C2I mutants						
T349H	0.89	2.0	NHB	1.0	NHB	−0.47
S350R	3.38	16	SHB	22	NHB	−0.28
E389K	0.33	14	SHB	12	USHB	−1.94
F398Y	0.63	6.0	NHB	2.0	NHB	−1.65
C2II mutants						
E307S	29	29	USHB	35	NHB	−0.62
E399C	58	58	SHB	47	SHB	−0.68
F456A	76	76	NHB	55	NHB	−1.79
F456S	76	76	NHB	53	NHB	−1.96
F456H	76	76	NHB	47	SHB	−0.57

SASA solvent accessible surface area, *SA* solvent accessibity, *HB* H-bonding, *NHB* no H-bonding, *SHB* saturated H-bonding, *USHB* unsaturated H-bonding

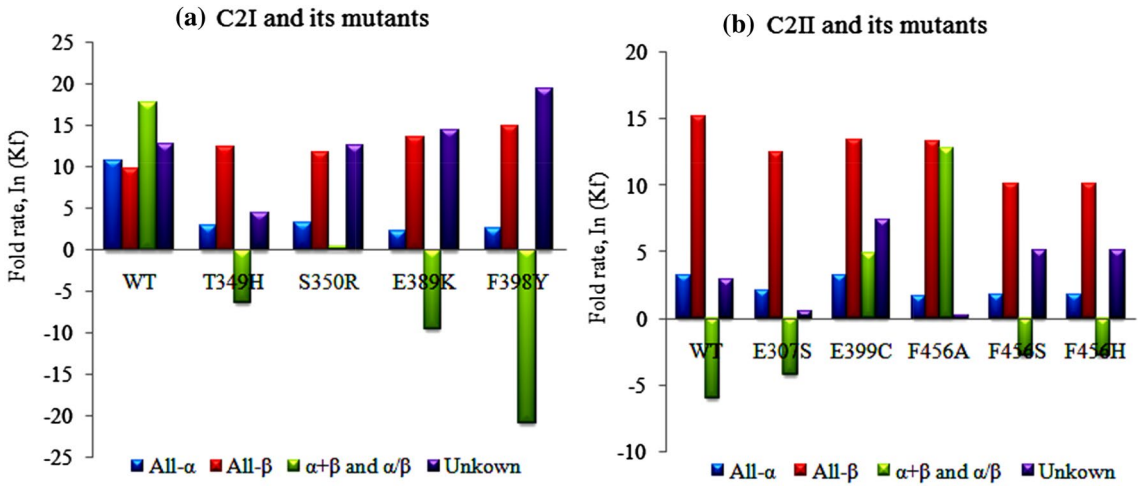

Fig. 4 Fold rate prediction for comparison of structural classes in C2 toxin and its avirulent mutants

Table 4 Impact of point mutation on the structural fluctuation (RMSF) and antigenic determinant sites in the mutated residues of C2I and C2II subunits

Mutant	Root mean square fluctuation (Å)		Immunogenic determinant score	
	Wild-type	Mutant type	Wild-type	Mutant type
C2I mutants				
T349H	−0.099	−0.099	1.064	1.063
S350R	0.188	0.266	1.183	1.114
E389K	−0.328	−0.438	0.796	0.834
F398Y	−0.123	−0.123	0.947	1.024
C2II mutants				
E307S	0.266	0.266	0.993	1.091
E399C	1.374	1.233	0.969	1.033
F456A	0.359	0.359	1.154	1.159
F456H	0.359	0.712	1.152	1.203
F456S	0.359	0.542	1.151	1.269

the side-chain of neighboring amino acids. We found no structural fluctuation in T349H and F398Y compared to C2I. E307S (0.266–0.477), E399C (1.374–1.233), F456H (0.359–0.712), and F456S (0.359–0.542) of C2II have shown conformational sub-dynamic flexibility in the side-chain of amino acids which are near to mutated site. However, no structural deviation is detected between F456A and C2II. This indicated that coevolutionary forces acting on the first mutated site may destabilize or stabilize the amino acid network of these mutants to withhold their native function.

Analysis of Immunogenic Nature of Avirulent Toxins

Antigenic determinants of C2 toxin were predicted and compared with its avirulent mutants (Table 4; Supplementary Fig. S7). It shows that E389K has more hydrophobicity at the residues near to mutated one. We found a more hydrophobic region in the S350R, indicating that it has more antigenic determinants in that region. The antigenic nature of F456S and C2II is almost similar to each other compared to other mutants. Antigenic determinants in E399C, F456A, and F456H predicted to be more than C2II, suggestive of their suitability as immunogens.

Discussion

A coordinated change of amino acid residues from more conserved to highly variable determines the molecular diversity of a protein during evolution (Engelhardt et al. 2011). Conserved domain-specific feature and a core of β-strands surrounded by α-helices reconcile a family-specific function of bacterial toxins and BADPRTs (Aktories et al. 2011; Chellapandi et al. 2013; Yeang and Haussler 2007). Phylogenetic analysis implied that C21 adaptor domain structure corresponds to similar domains of Ia and VIP2, and D1–D3 domain structure of C2II relates to PA and Ib in accordance with earlier studies (Aktories and Barth 2004; Tsuge et al. 2003, 2008). C2I and C2II share a common evolutionary origin like iota toxin, which might be evolved from BTA family at different evolutionary rates. Even if C2II is structurally and functionally similar to PA, D4 of C2II determines its receptor specificity and recognition (Barth et al. 2004; Varughese et al. 1999).

Phylogenetic lineage-specific changes can resolve the functional specificity of C2 on its protein target, and to maintain a newly evolved functional feature from BTA family. NAD-binding core and catalytic mechanism of C2I resembles to C3, but its overall domain structure is dissimilar to C3. It may be a reason of that both of these toxins are belonged to two distinct families of ADP-ribosylation superfamily (Simon et al. 2014).

C2II/Toxin2A and C2I/Ia clusters are functionally evolved across the BTA family by duplication of these genes via Type I functional divergence and neutral selection. Gamma shape parameter (α) described the degree of heterogeneity rate across the BTA family. A protein fold may be reused during divergent or convergent evolution, which determines the functional range of a newly evolved protein (Gerlt and Babbitt 2001). Accordingly, NAD-binding core in C2 might be derived from C3 gene duplication by convergent evolution. Substrate-binding specificity slowly originated from sequence divergence within the BTA family by divergent evolution. The sequence-structure compatibility of a protein was analyzed by studying intergenic recombination events (Xia and Levitt 2002). C2 toxin gene locus is located in either plasmid or phage genome of *C. botulinum*. C2 toxin gene might be acquired by horizontal transfer mechanism that determines its virulence state. The number of sequence changes occurring along a lineage is more variable than expected with a constant rate of independent substitutions (Choi and Hannehall 2013). Evolutionary pattern analysis suggested that C2I and C2II have a common evolutionary pattern and evolved at the same rate. At the sequence level, it coordinates and stabilizes its prior structures (C2I–C2IIa) to have a function on the cell surfaces. The probable mutations being fixed in C2 structure depend on the residue position in accordance with Studer et al. (2013).

Coevolution sites are important constraints to understand the structural and functional evolution of a protein (Marks et al. 2012; Sandler et al. 2013). In our study, coevolution was an important force to be acted on the residues in α-helices, leading to change its local structural environment to β-sheets in coevolved pairs. Conformational flexibility in C2 structure was resulted by coevolved pairs as similar to C3 exotoxin (Prathiviraj et al. 2015). Even if these avirulent mutants have stable energetic structures similar to C2, a residue-coupling mutation may restrain the binding affinity to NAD^+ or a protein substrate in accordance with earlier work on HIV-1 env protein (Travers et al. 2007). Since most of the functional sites are localized in its coiled and buried surface areas, correlated mutations in those sites may bring structural as well as functional changes (Aktories and Barth 2004; Barth et al. 1998; Blöcker et al. 2000; Sterthoff et al. 2010). Coevolution forces on the first residues may disorder or revolutionize its H-bonding pattern of coevolved pairs, leading to bring a radical change in H-bonding frequency. Therefore, coevolutionary constraints may prefer to stabilize its local structural environment and retained the non-toxic function.

The residues involving in catalytic and pathological mechanisms of the C2 are evolutionarily conserved across ADP-ribosyltransferases and also contributed in preserving the overall structural stability (Chavan et al. 1992; Han et al. 2001; Han and Tainer 2002; Kowarsch et al. 2010; Ménétrey et al. 2002, 2008). A structure–function link of ADP-ribosyltransferases was analyzed using several molecular constraints (Chellapandi 2014; Prathiviraj et al. 2015). Our study revealed that functional divergence of C2 structure was far slower than that of sequence. Thus, various structural constraints in the avirulent mutants of our study were detected to evaluate them as suitable avirulent toxins/immunogens.

Finding a suitable residue for replacing mutable one is an important measure to determine the structural and functional changes of a protein during the evolutionary process (Arenas et al. 2013). The structures of these avirulent mutants are stabilized by optimizing amino acid-atom potentials and torsion angle distribution (Gromiha et al. 2006). The SASA of these mutants are adjusted towards their buried surface area to hold the molecular function. Structure–function integrity may be retained by sustaining buried surface area and no H-bonding for detrimental binding. It suggested that avirulent mutants may analogously conserve their native-like structures. Secondary structural elements, particularly α-helices and β-sheets are not altered, but the loops/coils are reorganized from the native state, which was agreed with the previous work done by Sarabojia et al. (2005).

A conformational flexibility in fold states is a crucial constraint to design several proteins (Hammes et al. 2011; Ho and Agard 2010; Kuzmanic and Zagrovic 2010). Studies on protein folding rates are a measure to understand the pathogenic nature of mutants (Huang and Gromiha 2010, 2012). The fold rate on the structural classes may influence on the folding stability of these mutants compared to C2I. Our study suggested that there is no major impact of fold rates on the structural classes of protein folding process, but minor changes predicted in C2 mutants. It may be resulted due to the conserved NAD-binding β-sandwich toxin fold in wild-type C2 through all stages of the evolution (Blöcker et al. 2000; Han et al. 2001). Avirulent mutants with rapid folding rate do not require extensive evolutionary optimization, which was agreed with Riddle et al. (1997). Consequently, we revealed that avirulent mutants follow the protein folding rate from a common blueprint (β-sandwich toxin fold) of C2, which may suppose to carry out the fold recognition, process, stability and post translation modification in the expressed host. It also

described that evolutionary constraint manifest the fast and efficient folding rates of these mutants and correlate with stability rather than contact order. Our hypothesis were correlated with earlier work (Monsellier and Chiti 2007).

Immunogens and vaccines can be designed by successful capturing near-native shifts in a protein backbone in response to mutational changes (Correia et al. 2014; Kuroda et al. 2012; Sevy and Meiler 2014). Avirulent mutants derived from C2I are structurally stable, and the resulted ensembles shown near native-like shifts in a protein backbone of C2I. F456S has shown a localized structural flexibility, indicating that its channel-forming property can be reduced when compared to C2II. The stability of a protein backbone structure and solvent accessibility can be affected by point mutations (Maguidet al. 2006; Sikosek and Chan 2014; Worth et al. 2009). Accordingly, we found several conformational flexibilities in the backbone structure of these mutants by mutational constraints in C2 toxin.

Conformational reorganization of binding sites is particularly important to design a protein-based vaccine (De Oliveira et al. 2011; Lapelosa et al. 2009; Topchiy et al. 2013). Herein, we demonstrated how single amino acid substitution can increase the conformational flexibility in the structures of mutants (Lapelosa et al. 2009; Lukman et al. 2010). Mutational constraints in conformational sub-dynamic space determine the substrate recognition across the BTA family and NAD-binding core among the BAD-PRTs superfamily. We also found coevolutionary constraints that are acting on the amino acid network of many avirulent mutants to stabilize their structures as well as withhold the native function in accordance with previous studies (Chellapandi et al. 2013; Chellapandi 2014; Prathiviraj et al. 2015; Travers et al. 2007). Geometric constraints that shape patterns of antigenic variation in Influenza H3N2 Hemagglutinin and bacterial toxins (Foley 2015; Meyer and Wilke 2015). Even a single amino acid substitution/mutation may bring antigenic evolution in many human viruses (Liang et al. 2010; Yoo and Deregt 2001). In our study, we found a drastic change in antigenic peptides predicted from avirulent toxins compared to C2 toxin. Thus, we suggested that evolution constraints-free immunogenic peptides can raise the humoral immunity in mammalian cells.

Conclusions

Our study is helpful in understanding the role of functional mutable residues in the ADP ribosyltransferase activity of C2I and channel and pore-forming properties of C2II on actin. Evolutionary genetic analyses described that the virulence state of C2 toxin is constrained by the residues present in conserved NAD-binding core, family-specific domain structure, and functional motifs. Structure–function integrity of C2 can be destabilized by single amino acid substitution in these regions. We have screened and selected nine avirulent mutants from C2 toxin showing more stability in their local structural environments with low evolutionary constraints. A coevolution force is also a concern of acting on stabilization of the local structural environment as well as conformational sub-dynamic space of C2 toxin upon a point mutation. Avirulent mutants have low NAD-binding specificity and pore/channel-forming capability by altering their H-bonding patterns and frequency. These mutants are structurally stable with low conformational flexibility, followed common blue print of rapid folding process. Antigenic determinants in avirulent toxins with lack of catalytic and pore-forming function is comparatively better than C2 toxin, suggestive of using them as immunogens for treating *C. botulinum* infected poultry and veterinary animals. Nevertheless, these avirulent toxins should be experimentally validated by protein engineering and cell culture techniques for successful protein-based subunit vaccines/immunogens candidates.

Acknowledgements Life Science Research Board-Defense Research and Development Organization (Sanction No. DLS/81/48222/LSRB-249/BTB/2012), New Delhi, India, is duly acknowledged for financial support.

Compliance with Ethical Standards

Conflict of interest The authors confirm that this article content has no conflicts of interest.

References

Aktories K, Barth H (2004) *Clostridium botulinum* C2 toxin—new insights into the cellular up-take of the actin-ADP-ribosylating toxin. Int J Med Microbiol 293:557–564

Aktories K, Lang AE, Schwan C, Mannherz HG (2011) Actin as target for modification by bacterial protein toxins. FEBS J 278:4526–4543

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Arenas M, Dos Santos HG, Posada D, Bastolla U (2013) Protein evolution along phylogenetic histories under structurally constrained substitution models. Bioinformatics 29:3020–3028

Barth H, Preiss JC, Hofmann F, Aktories K (1998) Characterization of the catalytic site of the ADP-ribosyltransferase *Clostridium botulinum* C2 toxin by site-directed mutagenesis. J Biol Chem 273:29506–29511

Barth H, Aktories K, Popoff MR, Stiles BG (2004) Binary bacterial toxins: biochemistry, biology, and applications of common *Clostridium* and *Bacillus* proteins. Microbiol Mol Biol Rev 68:373–402

Benson EL, Huynh PD, Finkelstein A, Collier RJ (1998) Identification of residues lining the anthrax protective antigen channel. BioChemistry 37:3941–3948

Blöcker D, Barth H, Maier E, Benz R, Barbieri JT, Aktories K (2000) The C terminus of component C2II of *Clostridium botulinum* C2 toxin is essential for receptor binding. Infect Immun 68:4566–4573

Blöcker D, Pohlmann K, Haug G, Bachmeyer C, Benz R, Aktories K, Barth H (2003a) *Clostridium botulinum* C2 toxin: low pH-induced pore formation is required for translocation of the enzyme component C2I into the cytosol of host cells. J Biol Chem 278:37360–37367

Blöcker D, Bachmeyer C, Benz R, Aktories K, Barth H (2003b) Channel formation by the binding component of *Clostridium botulinum* C2 toxin: glutamate 307 of C2II affects channel properties in vitro and pH-dependent C2I translocation in vivo. BioChemistry 42:5368–5377

Chavan AJ, Nemoto Y, Narumiya S, Kozaki S, Haley BE (1992) NAD^+ binding site of *Clostridium botulinum* C3 ADP-ribosyltransferase Identification of peptide in the adenine ring binding domain using 2-azido NAD. J Biol Chem 267:14866–14870

Chellapandi P (2014) Structural-functional integrity of hypothetical proteins identical to ADP-ribosylation Superfamily upon point mutations. Protein Pept Lett 21:22–35

Chellapandi P, Shree SS, Bharathi M (2013) Phylogenetic approach for inferring the origin and functional evolution of bacterial ADP-ribosylation superfamily. Protein Pept Lett 20:1054–1065

Choi SS, Hannehall S (2013) Three independent determinants of protein evolutionary rate. J Mol Evol 76:98–111

Correia BE et al (2014) Proof of principle for epitope-focused vaccine design. Nature 507:201–206

Cunha CE, Moreira GM, Salvarani FM, Neves MS, Lobato FC, Dellagostin OA, Conceição FR (2014) Vaccination of cattle with a recombinant bivalent toxoid against botulism serotypes C and D. Vaccine 32:214–216

De Oliveira CAF, Grant BJ, Zhou M, McCammon JA (2011) Large-scale conformational changes of trypanosoma cruzi proline racemase predicted by accelerated molecular dynamics simulation. PLoS Comput Biol 7:e1002178

De Brevern AG, Bornot A, Craveur P, Etchebest C, Gelly JC (2012) PredyFlexy: flexibility and local structure prediction from sequence. Nucleic Acids Res 40:W317–W322

Domenighini M, Rappuoli R (1996) Three conserved consensus sequences identify the NAD-binding site of ADP-ribosylating enzymes, expressed by eukaryotes, bacteria and T-even bacteriophages. Mol Microbiol 21:667–674

Domenighini M, Magagnoli C, Pizza M, Rappuoli R (1994) Common features of the NAD-binding and catalytic site of ADP-ribosylating toxins. Mol Microbiol 14:41–50

Eckhardt M, Barth H, Blöcker D, Aktories K (2000) Binding of *Clostridium botulinum* C2 toxin to asparagine-linked complex and hybrid carbohydrates. J Biol Chem 275:2328–3234

Engelhardt BE, Jordan MI, Srouji JR, Brenner SE (2011) Genome-scale phylogenetic function annotation of large and diverse protein families. Genome Res 21:1969–1980

Fahrer J, Plunien R, Binder U, Langer T, Seliger H, Barth H (2010a) Genetically engineered clostridial C2 toxin as a novel delivery system for living mammalian cells. Bioc Onjug Chem 21:130–139

Fahrer J, Rieger J, Zandbergen GV, Barth H (2010b) The C2-streptavidin delivery system promotes the uptake of biotinylated molecules in macrophages and T-leukemia cells. Biol Chem 391:1315–1325

Foley J (2015) Mini-review: strategies for variation and evolution of bacterial antigens. Comput Struct Biotechnol J 13:407–416

Fujii N, Kubota T, Shirakawa S, Kimura K, Ohishi I, Moriishi K, Isogai E, Isogai H (1996) Characterization of component-I gene of *botulinum* C2 toxin and PCR detection of its gene in clostridial species. Biochem Biophys Res Commun 220:353–359

Gerlt JA, Babbitt PC (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. Annu Rev Biochem 70:209–246

Gil LA, Da Cunha CE, Moreira GM, Salvarani FM, Assis RA, Lobato FC, Mendonça M, Dellagostin OA, Conceição FR (2013) Production and evaluation of a recombinant chimeric vaccine against *Clostridium botulinum* neurotoxin types C and D. PLoS ONE 8:e69692

Gromiha MM, Thangakani AM, Selvaraj S (2006) FOLD-RATE: prediction of protein folding rates from amino acid sequence. Nucleic Acids Res 34:W70–W74

Gu X, Vander Velden K (2002) DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. Bioinformatics 18:500–501

Hammes GG, Benkovic SJ, Hammes-Schiffer S (2011) Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. BioChemistry 50:10422–10430

Han S, Tainer JA (2002) The ARTT motif and a unified structural understanding of substrate recognition in ADP-ribosylating bacterial toxins and eukaryotic ADP-ribosyltransferases. Int J Med Microbiol 291:429–523

Han S, Craig JA, Putnam CD, Carozzi NB, Tainer JA (1999) Evolution and mechanism from structures of an ADP-ribosylating toxin and NAD complex. Nat Struct Biol 6:932–936

Han S, Arvai AS, Clancy SB, Tainer JA (2001) Crystal structure and novel recognition motif of rho ADP-ribosylating C3 exoenzyme from *Clostridium botulinum*: structural insights for recognition specificity and catalysis. J Mol Biol 305:95–107

Ho BK, Agard DA (2010) Conserved tertiary couplings stabilize elements in the PDZ fold, leading to characteristic patterns of domain conformational flexibility. Protein Sci 19:398–411

Holbourn KP, Sutton JM, Evans HR, Shone CC, Acharya KR (2005) Molecular recognition of an ADP-ribosylating *Clostridium botulinum* C3 exoenzyme by RalA GTPase. Proc Natl Acad Sci USA 102:5357–5362

Huang LT, Gromiha MM (2010) First insight into the prediction of protein folding rate change upon point mutation. Bioinformatics 26:2121–2127

Huang LT, Gromiha MM (2012) Real value prediction of protein folding rate change upon point mutation. J Comput Aided Mol Des 26:339–347

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

Jank T, Aktories K (2013) Strain-alleviation model of ADP-ribosylation. Proc Natl Acad Sci USA 110:4163–4164

Jeong JS, Kim D (2012) Reliable and robust detection of coevolving protein residues. Protein Eng 25:705–713

Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276:172–174

Kowarsch A, Fuchs A, Frishman D, Pagel P (2010) Correlated Mutations: a hallmark of phenotypic amino acid substitutions. PLoS Comput Biol 6:e1000923

Krüger M, Skau M, Shehata AA, Schrödl W (2013) Efficacy of *Clostridium botulinum* types C and D toxoid vaccination in Danish cows. Anaerobe 23:97–101

Kuroda D, Shirai H, Jacobson MP, Nakamura H (2012) Computer-aided antibody design. Protein Eng Des Sel 25:507–521

Kuzmanic A, Zagrovic B (2010) Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophys J 98:861–871

Lang AE, Neumeyer T, Sun J, Collier RJ, Benz R, Aktories K (2008) Amino acid residues involved in membrane insertion and pore

formation of *Clostridium botulinum* C2 toxin. BioChemistry 47:8406–8413
Lapelosa M, Gallicchio E, Arnold GF, Arnold E, Levy RM (2009) In silico vaccine design based on molecular simulations of rhinovirus chimeras presenting HIV-1 gp41 epitopes. J Mol Biol 385:675–691
Leppla SH (1995) Anthrax toxins. In: Handbook of natural toxins, bacterial toxins and virulence factors in disease, vol 8. New York, Marcel Dekker, pp. 543–572
Liang JH, Dai X, Dong C, Meng JH (2010) A single amino acid substitution changes antigenicity of ORF2-encoded proteins of hepatitis E virus. Int J Mol Sci 11:2962–2975
Lio P, Goldman N (1999) Using protein structural information in evolutionary inference: transmembrane proteins. Mol Biol Evol 16:1696–1710
Lukman S, Grant GH, Bui JM (2010) Unraveling evolutionary constraints: a heterogeneous conservation in dynamics of the titin Ig domains. FEBS Lett 584:1235–1239
Maguid S, Fernández-Alberti S, Parisi G, Echave J (2006) Evolutionary conservation of protein backbone flexibility. J Mol Evol 63:448–457
Marchler-Bauer A et al (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226
Marks DS, Hopf TA, Sander C (2012) Protein structure prediction from sequence variation. Nat Biotechnol 30:1072–1080
Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462–2463
Mayrose I, Graur D, Tal BN, Pupko T (2004) Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. Mol Biol Evol 21:1781–1791
Menetrey J, Flatau G, Boquet P, Menez A, Stura EA (2008) Structural basis for the NAD-hydrolysis mechanism and the ARTT-loop plasticity of C3 exoenzymes. Protein Sci 17:878–886
Ménétrey J, Flatau G, Stura EA, Charbonnier JB, Gas F, Teulon JM, Le Du MH, Boquet P, Menez A (2002) NAD binding induces conformational changes in Rho ADP-ribosylating *Clostridium botulinum* C3 exoenzyme. J Biol Chem 277:30950–30957
Meyer AG, Wilke CO (2015) Geometric constraints dominate the antigenic evolution of influenza H3N2 hemagglutinin. PLoS Pathog 11:e1004940
Mizuguchi K, Deane CM, Blundell TL, Johnson MS, Overington JP (1998) JOY: protein sequence-structure representation and analysis. Bioinformatics 14:617–623
Monsellier E, Chiti F (2007) Prevention of amyloid-like aggregation as a driving force of protein evolution. EMBO Rep 8:737–742
Moriishi K, Syuto B, Yokosawa N, Oguma K, Saito M (1991) Purification and characterization of ADP-ribosyltransferases (exoenzyme C3) of *Clostridium botulinum* type C and D strains. J Bacteriol 73:6025–6029
Moriishi K, Syuto B, Saito M, Oguma K, Fujii N, Abe N, Naiki M (1993) Two different types of ADP-ribosyltransferase C3 from *Clostridium botulinum* type D lysogenized organisms. Infect Immun 61:5309–5314
Nagahama M, Morimitsu S, Kihara A, Akita M, Setsu K, Sakurai J (2003) Involvement of tachykinin receptors in Clostridium perfringens beta-toxin-induced plasma extravasation. Br J Pharmacol 138:23–30
Ohishi I, wasaki M, Sakaguchi G (1980) Purification and characterization of two components of botulinum C2 toxin. Infect Immun 30:668–673
Parthiban V, Gromiha MM, Schomburg D (2006) CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res 34:W239–W242
Pavelka A, Chovancova E, Damborsky J (2009) HotSpot Wizard: a web server for identification of Hot Spots in—protein engineering. Nucleic Acids Res 37:376–383
Pearson W (1998) Empirical statistical estimates for sequence similarity searches. J Mol Biol 276:71–84
Petosa C, Collier RJ, Klimpel KR, Leppla SH, Liddington RC (1997) Crystal structure of the anthrax toxin protective antigen. Nature 385:833–838
Pond SL, Frost SD, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679
Prathiviraj R, Prisilla A, Chellapandi P (2015) Structure-function discrepancy in *Clostridium botulinum* C3 toxin for its rational prioritization as a subunit vaccine. J Biomol Struct Dyn 34(6):1317–1329
Pust S, Barth H, Sandvig K (2010) *Clostridium botulinum* C2 toxin is internalized by clathrin- and Rho-dependent mechanisms. Cell Microbiol 12:1809–1820
Riddle DS, Santiago JV, Bray-Hall ST, Doshi N, Grantcharova VP, Yi Q, Baker D (1997) Functional rapidly folding proteins from simplified amino acid sequences. Nat Struct Biol 4:805–809
Sandler I, Abu-Qarn M, Aharoni A (2013) Protein co-evolution: how do we combine bioinformatics and experimental approaches? Mol Biosyst 9:175–181
Sarabojia K, Gromihab MM, Ponnuswamy MN (2005) Relative importance of secondary structure and solvent accessibility to the stability of protein mutants. A case study with amino acid properties and energetics on T4 and human lysozymes. Comput Biol Chem 29:25–35
Schleberger C, Hochmann H, Barth H, Aktories K, Schulz GE (2006) Structure and action of the binary C2 toxin from *Clostridium botulinum*. J Mol Biol 8:705–715
Sevy AM, Meiler J (2014) Antibodies: computer-aided prediction of structure and design of function. Microbiol Spectr. doi:10.1128/microbiolspec.AID-0024-2014
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504
Sikosek T, Chan HS (2014) Biophysics of protein evolution and evolutionary protein biophysics. J R Soc Interface. doi:10.1098/rsif.2014.0419
Simon NC, Aktories K, Barbieri JT (2014) Novel bacterial ADP-ribosylating oxins: structure and function. Nat Rev Microbiol 12:599–611
Smith CA, Kortemme T (2008) Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. J Mol Biol 380:742–756
Sterthoff C, Lang AE, Schwan C, Tauch A, Aktories K (2010) Functional characterization of an extended binding component of the actin-ADP-ribosylating C2 toxin detected in *Clostridium botulinum* strain (C) 2300. Infect Immun 78:1468–1474
Studer RA, Dessailly BH, Orengo CA (2013) Residue mutations and their impact on protein structure and function: detecting beneficial and pathogenic changes. Biochemistry J 449:581–594
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
Takada T, Iida K, Moss J (1995) Conservation of a common motif in enzymes catalyzing ADP-ribose transfer identification of domains in mammalian transferases. J Biol Chem 270:541–544
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for

multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Topchiy E, Armstrong GS, Boswell KI, Buchner GS, Kubelka J, Lehmann TE (2013) T1BT* structural study of an anti-plasmodial peptide through NMR and molecular dynamics. Malar J 12:104

Travers SA, Tully DC, McCormack GP, Fares MA (2007) A study of the coevolutionary patterns operating within the env gene of the HIV-1 group M subtypes. Mol Biol Evol 24:2787–2801

Tseng YY, Liang J (2006) Automated method for predicting enzyme functional surfaces and locating key residues with accuracy and specificity. Conf Proc IEEE Eng Med Biol Soc 1:4552–4555

Tsuge H, Nagahama M, Nishimura H, Hisatsune J, Sakaguchi Y, Itogawa Y, Katunuma N, Sakurai J (2003) Crystal structure and site directed mutagenesis of enzymatic components from *Clostridium perfringens* iota-toxin. J Mol Biol 325:471–483

Tsuge H, Nagahama M, Oda M, Iwamoto S, Utsunomiya H, Marquez VE, Katunuma N, Nishizawa M, Sakurai J (2008) Structural basis of actin recognition and arginine ADP-ribosylation by *Clostridium perfringens* iota-toxin. Proc Natl Acad Sci USA 105:7399–7404

Udaya Prakash NA, Jayanthi M, Sabarinathan R, Kangueane P, Mathew L, Sekar K (2010) Evolution, homology conservation, and identification of unique sequence signatures in GH19 family chitinases. J Mol Evol 70:466–478

Varughese M, Teixeira AV, Liu S, Leppla SH (1999) Identification of a receptor-binding region within domain 4 of the protective antigen component of anthrax toxin. Infect Immun 67:1860–1865

Vitkup D, Sander C, Church GM (2003) The amino-acid mutational spectrum of human genetic disease. Genome Biol 4:R72

Wan Y, Ren X, Ren Y, Wang J, Hu Z, Xie X, Xu J (2014) As a genetic adjuvant, CTA improves the immunogenicity of DNA vaccines in an ADP-ribosyltransferase activity- and IL-6-dependent manner. Vaccine 32:2173–2180

Wiegers W, Just I, Müller H, Hellwig A, Traub P, Aktories K (1991) Alteration of the cytoskeleton of mammalian cells cultured in vitro by *Clostridium botulinum* C2 toxin and C3 ADP-ribosyltransferase. Eur J Cell Biol 54:237–245

Worth CL, Gong S, Blundell TL (2009) Structural and functional constraints in the evolution of protein families. Nat Rev Mol Cell Biol 10:709–720

Worth CL, Preissner R, Blundell TL (2011) SDM—a server for predicting effects of mutations on protein stability and malfunction. Nucleic Acids Res 39:W215–W222

Xia Y, Levitt M (2002) Roles of mutation and recombination in the evolution of protein thermodynamics. Proc Natl Acad Sci USA 99:10382–10387

Yeang CH, Haussler D (2007) Detecting coevolution in and among protein domains. PLoS Comput Biol 3:e211

Yoo D, Deregt D (2001) A single amino acid change within antigenic domain II of the spike protein of bovine coronavirus confers resistance to virus neutralization. Clin Diagn Lab Immunol 8:297–302

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J (2006) Ka/Ks Calculator: Calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics 4:259–263