Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Peptide inhibitors of botulinum neurotoxin serotype A: design, inhibition, cocrystal structures, structure–activity relationship and pharmacophore modeling

Gyanendra Kumar,[a]‡ Desigan Kumaran,[a]‡ S. Ashraf Ahmed[b] and Subramanyam Swaminathan[a]*

[a]Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA, and [b]Department of Molecular Biology, Integrated Toxicology Division, US Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702, USA

‡ These authors made equal contributions to this work.

Correspondence e-mail: swami@bnl.gov

Received 14 December 2011
Accepted 26 January 2012

PDB References:
Balc424–RRGF, 3qw5;
Balc424–RYGC, 3qw6;
Balc424–RRFC, 3qw7;
Balc424–CRGC, 3qw8.

Clostridium botulinum neurotoxins are classified as Category A bioterrorism agents by the Centers for Disease Control and Prevention (CDC). The seven serotypes (A–G) of the botulinum neurotoxin, the causative agent of the disease botulism, block neurotransmitter release by specifically cleaving one of the three SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor) proteins and induce flaccid paralysis. Using a structure-based drug-design approach, a number of peptide inhibitors were designed and their inhibitory activity against botulinum serotype A (BoNT/A) protease was determined. The most potent peptide, RRGF, inhibited BoNT/A protease with an IC_{50} of 0.9 μ*M* and a K_i of 358 n*M*. High-resolution crystal structures of various peptide inhibitors in complex with the BoNT/A protease domain were also determined. Based on the inhibitory activities and the atomic interactions deduced from the cocrystal structures, the structure–activity relationship was analyzed and a pharmacophore model was developed. Unlike the currently available models, this pharmacophore model is based on a number of enzyme–inhibitor peptide cocrystal structures and improved the existing models significantly, incorporating new features.

1. Introduction

The seven serotypes of botulinum neurotoxin (BoNT/A–G) secreted by various anaerobic Gram-positive spore-forming bacteria of the genus *Clostridium* are members of a family of clostridial neurotoxins (CNT) that also includes tetanus toxin (Simpson, 1989). They are synthesized as ∼150 kDa single-chain polypeptides and released as disulfide-linked heterodimers after proteolytic cleavage. The dimer is composed of an ∼100 kDa heavy chain (HC) and an ∼50 kDa light chain (LC) (Montecucco & Schiavo, 1995). The HC itself comprises two ∼50 kDa domains. The C-terminal domain (H_C), the ganglioside and protein receptor-binding domain, plays a role in binding to the cell membrane and internalization of the toxin into cholinergic neurons, and the N-terminal domain (H_N) facilitates the release of the LC from endosomes into the cytosol (Montecucco, 1986; Black & Dolly, 1986*a*,*b*). The LC is a zinc-dependent endopeptidase that cleaves and inactivates SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor) proteins. SNARE proteins are essential for the release of neurotransmitters, and cleavage of SNARE protein(s) by BoNT disrupts the release of acetylcholine from synaptic terminals, leading to flaccid paralysis or botulism (Montecucco & Schiavo, 1995). The BoNT catalytic domains share ∼36% sequence identity across serotypes and show specificity with respect to their substrates. Serotypes A and E cleave SNAP-25 (synaptosomal-associated protein of

25 kDa), serotypes B, D, F and G cleave VAMP (vesicle-associated membrane protein) and serotype C cleaves both SNAP-25 and syntaxin (Binz *et al.*, 1994; Schiavo *et al.*, 1992; Schiavo, Shone *et al.*, 1993; Schiavo, Rossetto *et al.*, 1993; Schiavo, Malizio *et al.*, 1994; Blasi *et al.*, 1993).

The lethal dose (LD_{50}) of BoNTs for humans is in the range 0.1–1 ng kg^{-1}, making them the most poisonous biological substances known (Schiavo, Rossetto *et al.*, 1994). Their widespread use as therapeutic and cosmetic agents also increases the risk of accidental overdose. Since they can easily be produced and delivered *via* aerosol medium, these agents are considered to be among the most deadly of all potential bioweapons (Josko, 2004) and are classified as such by the Centers for Disease Control and Prevention. The only available antidote against BoNTs is equine antitoxin and it can only target the toxins at the extracellular level. It cannot reverse the paralysis caused by botulism (Cai & Singh, 2007). Efforts to develop recombinant antibodies are also in progress (Marks, 2004). The only post-intoxication treatment currently available is mechanical respiration. Since the metalloprotease activity of BoNT LC can last for several weeks depending upon the serotype (Meunier *et al.*, 2003), the intense inpatient medical care needed to treat botulism is impractical; hence, the development of potent and effective inhibitors of the metalloprotease activity of BoNTs is a priority. In this direction, crystal structures of various BoNT serotypes have been determined by us and others (Lacy *et al.*, 1998; Swaminathan & Eswaramoorthy, 2000; Arndt *et al.*, 2005, 2006; Agarwal *et al.*, 2004, 2005*a*,*b*, 2009; Agarwal & Swaminathan, 2008; Kumaran, Rawat, Ahmed *et al.*, 2008; Kumaran, Rawat, Ludivico *et al.*, 2008). Based on the atomic-level details provided by these structures, several peptide and nonpeptide inhibitors have been discovered (Zuniga *et al.*, 2008; Kumaran, Rawat, Ludivico *et al.*, 2008; Silvaggi *et al.*, 2007; Burnett, Ruthel *et al.*, 2007; Park *et al.*, 2006).

BoNT/A is the most potent of all the clostridial neurotoxins and we recently mapped the interactions of its substrate at the active site of the catalytic domain (residues 1–424; Balc424) using the SNAP-25 peptide (197)QRATKM(202) (Kumaran, Rawat, Ahmed *et al.*, 2008). We also discovered substrate-based tetrapeptide inhibitors and analyzed their interactions with the active-site residues at near-atomic resolution (Kumaran, Rawat, Ludivico *et al.*, 2008). Following this, we conducted detailed studies on their solubility, stability and toxicity in neuronal cells. These tetrapeptides were able to inhibit BoNT/A cleavage of endogenous SNAP-25 in mouse-brain lysates. They were also able to penetrate N2A and BE(2)-M17 neuronal cell lines without adversely affecting metabolic functions. They were active in cultured chick motor neurons and rat and mouse cerebellar neurons for more than 40 h. They inhibited BoNT/A protease activity in the neurons in a dose- and time-dependent manner. Thus, we established these tetrapeptides as excellent candidates for drug discovery against botulism (Hale *et al.*, 2011). Here, we report the design, inhibition, cocrystal structures and structure–activity relationship of a number of new peptide inhibitors. The binding modes and binding energies of the designed molecules were predicted using docking simulations. The inhibitory activity was tested using an ultra-performance liquid chromatography-based assay and the cocrystal structures were determined using X-ray crystallography. The structure–activity relationship of these inhibitors is discussed and a pharmacophore model has been developed.

2. Materials and methods

2.1. Cloning, expression and purification of BoNT/A

Balc424 was cloned, overexpressed in *Escherichia coli* and purified to homogeneity using size-exclusion chromatography as described previously (Kumaran, Rawat, Ludivico *et al.*, 2008). The purified enzyme in 20 m*M* HEPES, 2 m*M* DTT, 200 m*M* NaCl pH 7.4 was stored at 193 K until use. Amides of the peptides were custom-synthesized by Peptide 2.0 Inc., Chantilly, Virginia, USA and the two modified peptides were synthesized by Ezbiolab, Carmel, Indiana, USA. Stock solutions of the peptides were prepared in the above-mentioned buffer.

2.2. Enzyme-activity and inhibition assays

The proteolytic activity of Balc424 and its inhibition by peptide inhibitors were determined by ultra-performance liquid chromatography (UPLC) as described previously (Rowe *et al.*, 2010). Briefly, a 17-residue synthetic peptide corresponding to residues 187–203 of SNAP-25 was used as the substrate for a BoNT/A protease assay. The 30 µ*M* reaction mixture consisted of 0.9 m*M* substrate peptide, 0.2 mg ml^{-1} BSA, 0.0025 mg ml^{-1} Balc424 and 50 m*M* Na HEPES, 0.25 m*M* $ZnCl_2$, 5 m*M* dithiothreitol pH 7.4. The amounts of cleaved products and uncleaved substrate were measured after separation by reverse-phase UPLC using a Waters Acquity UPLC system equipped with *Empower Pro* software. A 1.7 mm C18 column (2.1 × 50 mm) with 0.1% trifluoroacetic acid as solvent *A* and 70% acetonitrile, 0.1% trifluoroacetate as solvent *B* at a flow rate of 0.5 ml min^{-1} was used to resolve the substrate and the products.

2.3. Docking simulations

The docking program *AutoDock* 3.05 was used for docking simulations (Morris *et al.*, 1998).

2.3.1. Preparation of the receptor and ligand molecules. The crystal structure of Balc424 (PDB entry 3c88; Kumaran, Rawat, Ludivico *et al.*, 2008) was used for docking simulations. H atoms were added and energy-minimized using the *MOE* (*Molecular Operating Environment*) 2009.10 molecular-modeling program. The three-dimensional coordinates were converted to mol2 format with MMFF94 charges using the *MOE* suite of programs. The receptor file was then prepared using *AutoDock Tools* with columns of fractional volumes and solvation parameters added (Sanner, 1999). Similarly, ligand molecules were also prepared using *MOE* and *AutoDock Tools*. H atoms and MMFF94 charges were added using *MOE* followed by defining torsions using *AutoDock Tools* prior to docking.

Table 1
Crystal data and refinement statistics of Balc424 peptide-inhibitor complexes.

Values in parentheses are for the last shell.

Peptide	RRGF	CRGC	RYGC	RRFC
Unit-cell parameters				
a (Å)	50.70	50.70	51.07	49.84
b (Å)	66.27	66.28	66.38	66.31
c (Å)	64.57	64.77	64.81	64.85
β (°)	98.3	98.4	98.4	98.4
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Resolution range (Å)	50–1.6 (1.65–1.60)	50–1.6 (1.65–1.60)	50–1.6 (1.65–1.60)	50–1.5 (1.54–1.50)
No. of unique reflections	49235	51397	55851	63078
Completeness (%)	88.3 (57)	91.5 (62)	98.6 (91)	93.8 (70)
$R_{\rm merge}$† (%)	8.5 (31)	7.8 (30)	5.7 (43)	6.4 (34)
$\langle I/\sigma(I)\rangle$	11.9 (2.0)	15.8 (2.0)	15.0 (1.5)	15.1 (1.5)
Refinement statistics				
Resolution (Å)	50–1.6	50–1.6	50–1.6	50–1.50
R factor‡/$R_{\rm free}$ (%)	21.3/23.2	19.0/21.7	20.6/22.2	19.5/20.9
R.m.s. deviations from ideality				
Bond lengths (Å)	0.005	0.005	0.005	0.005
Bond angles (°)	1.2	1.18	1.2	1.2
Average B factors (Å^2)				
Main chain	25.3	22.1	21.7	19.3
Side chain	26.8	24.0	23.6	21.3
Waters	30.8	28.9	28.8	26.9
Ions	44.6	37.8	41.2	39.7
Peptide inhibitor	30.1	27.4	24.0	24.1
No. of atoms				
Protein	3422	3416	3416	3416
Waters	272	331	303	395
Ions	16	17	17	17
Peptide inhibitor	38	28	34	40
Ramachandran plot analysis (%)				
Favoured	89.9	89.7	88.4	90.0
Allowed	10.1	10.3	11.6	10.0
Disallowed	0.0	0.0	0.0	0.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity over symmetry equivalents. ‡ R factor $= \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$.

2.3.2. Docking simulations. Grid maps for docking simulations were generated with 71, 61 and 61 grid points (with 0.375 Å spacing) in the x, y and z directions, respectively, centered in the active site using the *AutoGrid* program. The 12-10 and 12-6 Lennard–Jones parameters (supplied with the program package) were used to model hydrogen bonds and van der Waals interactions, respectively. The distance-dependent dielectric permittivity of Mehler and Solmajer was used in the calculation of electrostatic grid maps (Mehler & Solmajer, 1991). The genetic algorithm (GA) and Lamarckian genetic algorithm with pseudo Solis and Wets modification (LGA/pSW) methods were used with default parameters. For each inhibitor, 50 docking runs were conducted and cluster analysis was performed on the docked conformations using an r.m.s. tolerance of 2.0 Å.

2.4. Crystallization and data collection

Balc424 was cocrystallized with peptide inhibitors (RRGF, CRGC, RYGC and RRFC) by the sitting-drop vapor-diffusion method at room temperature. Prior to crystallization, stock solutions of the peptides were prepared in water and added to the concentrated protein solution (∼15 mg ml^{-1}). Balc424 was mixed with a 10–20-fold molar excess of the peptide inhibitors for cocrystallization trials. The Balc424–peptide inhibitor mixtures were incubated at 277 K for 1 h. The best complex crystals were obtained at protein:peptide stoichiometric ratios of 1:15, 1:12, 1:10 and 1:18 for RRGF, CRGC, RYGC and RRFC, respectively. The crystallization condition was 20–25%(*w*/*v*) PEG 3350, 0.3 *M* ammonium sulfate, 50 m*M* HEPES buffer pH 6.5–7.5. Thick rectangular plate-like crystals were obtained in 5–10 d and were flash-cooled with liquid nitrogen using 25% ethylene glycol as a cryoprotectant. X-ray intensity data were collected on the X29 beamline of the National Synchrotron Light Source (NSLS) using an ADSC Quantum 315 detector. The crystals diffracted to at least 1.6 Å resolution and belonged to space group $P2_1$ with one molecule in the asymmetric unit (Table 1). Data were indexed, processed and scaled using the *HKL*-2000 suite (Otwinowski & Minor, 1997).

2.5. Structure determination

The crystal structures of the complexes were determined by rigid-body refinement followed by simulated annealing of the native structure (PDB entry 3bwi; Kumuran, Rawat, Ludivico *et al.*, 2008). Both the composite OMIT maps and the difference Fourier maps clearly showed well defined electron density allowing unambiguous modeling of the inhibitors (Fig. 1). However, to eliminate bias in the electron density the protein was initially modeled and refined in the absence of the inhibitor. In the RRGF, RRFC and RYGC complex structures residual densities elsewhere were modeled as sulfate ions. Extra densities found near (∼2.1 Å) the thiol group of the N-terminal cysteine in CRGC and the C-terminal cysteines in RRFC and RYGC were modeled as sodium ions. This sodium ion in the CRGC structure further interacted with the bound ethylene glycol that was added as cryoprotectant. The graphics program *O* was used to model the enzyme and the inhibitor in the electron density (Jones *et al.*, 1991). The refinement procedures, which included rigid-body refinement, simulated annealing, water picking, minimization *etc.*, were carried out using *CNS* (Brünger *et al.*, 1998). The refined models were validated using the Ramachandran plot and *PROCHECK* (Laskowski *et al.*, 1993). The crystallographic data, refinement statistics and quality of the model are given in Table 1.

2.6. Pharmacophore modeling

Pharmacophore models were developed and elucidated using the molecular-modeling program *MOE*. Briefly, the cocrystal structures of all of the peptide inhibitors in complex

with Balc424 were superimposed and the highly consistent interactions contributing toward inhibitory activity were shortlisted. Based on the properties of these interactions, the features of the pharmacophore were illustrated.

3. Results and discussion

We have previously reported the inhibition of Balc424 by four substrate-analogue tetrapeptides (RRGC, RRGL, RRGI and RRGM) mimicking the P1–P1′–P2′–P3′ stretch of the substrate (Kumaran, Rawat, Ludivico *et al.*, 2008) and two hexapeptides (QRATKM and RRATKM; Kumaran, Rawat, Ahmed *et al.*, 2008). In the former study (Kumaran, Rawat, Ludivico *et al.*, 2008), in which we tested four residues of varying size and properties at the P3′ position, that with cysteine (RRGC) showed improved affinity over the other three in terms of K_i values against purified Balc424. In order to understand the structure–activity relationship and to find better peptide-based drug candidates, we designed and tested a number of new peptide and modified-peptide inhibitors by varying residues at all four positions: P1, P1′, P2′ and P3′. We also varied the length of the peptides to find the optimum size for a peptide-based inhibitor. We employed the structure-based drug-discovery (SBDD) approach supplemented with docking simulations to design potential candidates. In our previously determined cocrystal structures of tetrapeptide inhibitors with Balc424, the N atom of the N-terminal amino group replaced the nucleophilic water and coordinated to zinc. It also formed a hydrogen bond to Glu224, which acts as a base in the catalytic activity. The guanidinium group of the first arginine (P1) made a salt bridge with Glu164 at the S1 subsite. The side chain of the P1′ arginine was stacked between Arg363 and Phe194, with its guanidinium group forming a salt bridge with Asp370. Also, the P1 carbonyl oxygen formed a hydrogen bond to the side-chain hydroxyl of Tyr366 in addition to coordinating to the zinc. The P1′ carbonyl O atom formed a hydrogen bond to the side chain of Arg363. The C-terminal amide O atom formed a hydrogen bond to the backbone N atom of Asp370, and the side chain of the C-terminal residue was nestled in the hydrophobic-cum-aromatic pocket formed by Pro206, Leu207, Tyr250, Tyr251, Met253, Leu256, Phe369 and Phe423.

Based on our lead peptide RRGC, we designed a number of new peptides and conducted docking studies to predict their

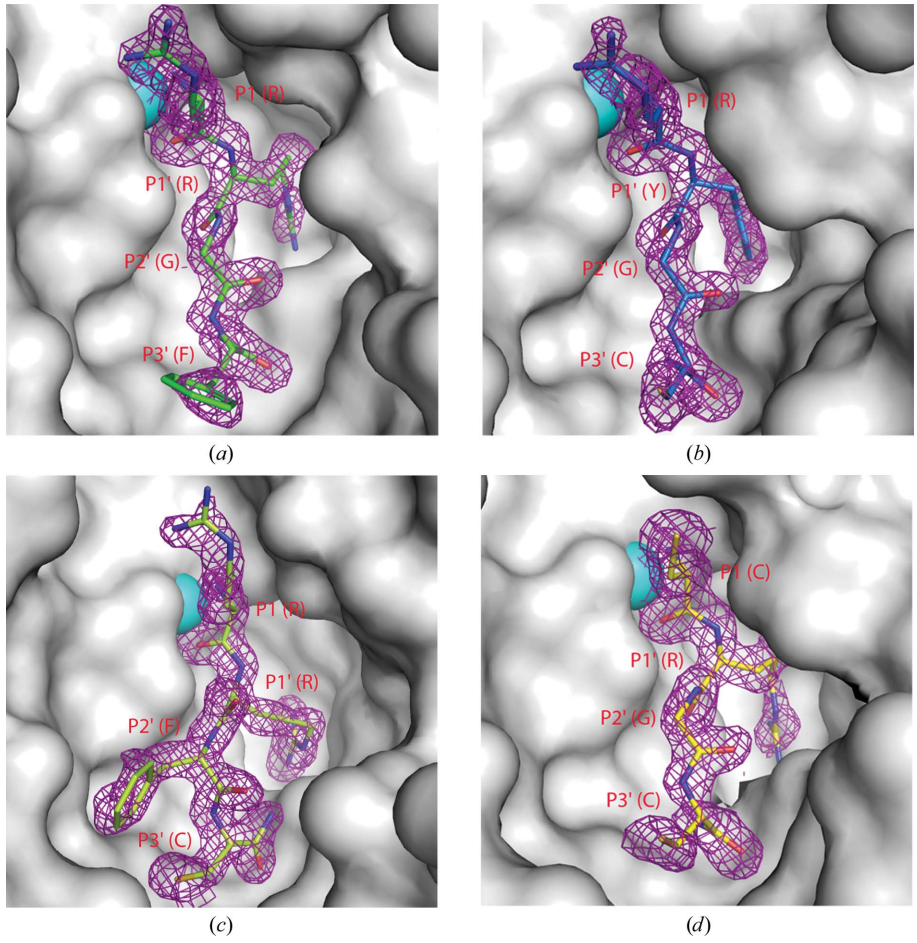

Figure 1
Electron-density maps. σ_A-Weighted ($2|F_o| - |F_c|$) composite OMIT maps for RRGF, RYGC, RRFC and CRGC are shown in (*a*), (*b*), (*c*) and (*d*), respectively, and are contoured at the 1σ level. Peptides are shown as stick models. The Zn atom is shown as a cyan sphere. The active-site cavity of the enzyme is shown as a surface representation (gray). This figure was prepared with *PyMOL* (DeLano, 2002).

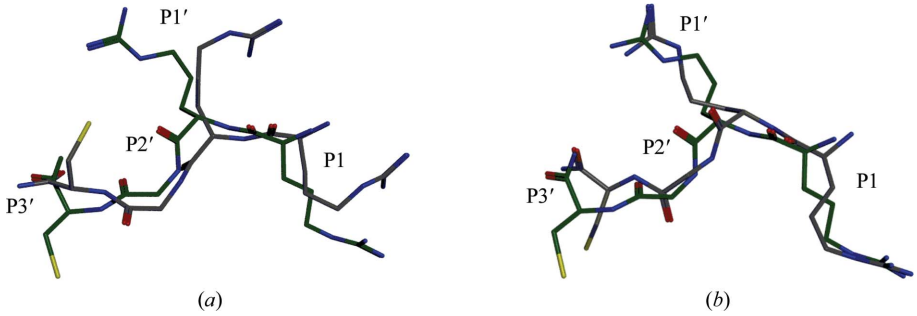

Figure 2
Comparison of docked and cocrystallized peptide inhibitor RRGC. The cocrystallized conformation is shown with green C atoms, while those of the docked conformation are shown in gray. (*a*) The docked conformation without waters in the receptor active site; the r.m.s.d. is 2.81 Å. (*b*) The docked conformation with waters in the receptor active site as found in the cocrystallized structure; the r.m.s.d. is 1.52 Å.

Table 2
Docking of peptide inhibitors with BoNT/A LC using *AutoDock* 3.05.

During the docking simulations, the protein receptor was kept rigid while the peptide molecules were fully flexible. 1 cal = 4.186 J.

	Protein receptor without waters		Protein receptor with waters	
Peptide	R.m.s.d. (Å)	ΔG† (kcal mol^{-1})	R.m.s.d. (Å)	ΔG† (kcal mol^{-1})
RRGC	2.81	−12.03	1.52	−10.01
RRGF	1.96	−10.19	1.47	−14.76
RRFC	2.28	−10.28	1.37	−12.64
CRGC	1.71	−10.12	0.80	−15.11
RRGCM	2.35	−7.68	1.95	−7.51
RYGC	1.97	−14.20	1.35	−15.07

† Estimated free energy of binding.

potential binding modes and binding energies using *AutoDock* 3.05. Unlike small-molecule inhibitors, accurately predicting the binding modes and binding energies of peptides is challenging owing to the larger size and enormous flexibility of the peptides. During the docking simulations, the peptides were kept flexible and the protein was kept rigid. To begin with, we docked RRGC (for which we already had a cocrystal structure) to test the accuracy of docking simulations by *AutoDock* on our protein of interest, Balc424. As is usually performed, we stripped the protein of ions (except for the catalytic zinc) and water molecules before running docking simulations. At the end of the simulation, the docked RRGC had a root-mean-square deviation (r.m.s.d.) of 2.81 Å with the cocrystallized RRGC in the crystal structure (Fig. 2 and Table 2). On careful analysis, we found that one reason for the deviation of the docked conformation from the experimental structure could be the removal of certain water molecules at the active site of the prepared receptor that make hydrogen bonds with peptide inhibitor as well as protein atoms. In the crystal structure, a number of interactions between the peptide inhibitors and Balc424 are mediated through these water molecules. We found that in other cocrystallized tetrapeptides (RRGL, RRGI and RRGM) a few water molecules are also conserved and are consistently involved in mediating the inhibitor–protein interactions (Fig. 3). Also, compared with the native structure some water molecules were displaced upon binding of peptide inhibitors, as one would expect, but certain others remained conserved and some new water molecules were also recruited to facilitate hydrogen bonding between the peptide inhibitor and the Balc424 active-site residues. Therefore, we conducted another set of docking simulations keeping the conserved water molecules as part of the protein receptor. This led to an improvement in the accuracy of the docking-pose prediction, and the r.m.s.d. between the docked and the cocrystallized RRGC conformations improved to 1.52 Å from 2.81 Å. After this, we conducted docking simulations on various designed peptides with and without water molecules and in a number of cases improved binding modes were achieved in the presence of these conserved water molecules. This trend indicated that water molecules play an important role in the inhibition of Balc424 by these peptide inhibitors. This fact could be utilized in further improving the inhibition activity by incorporating

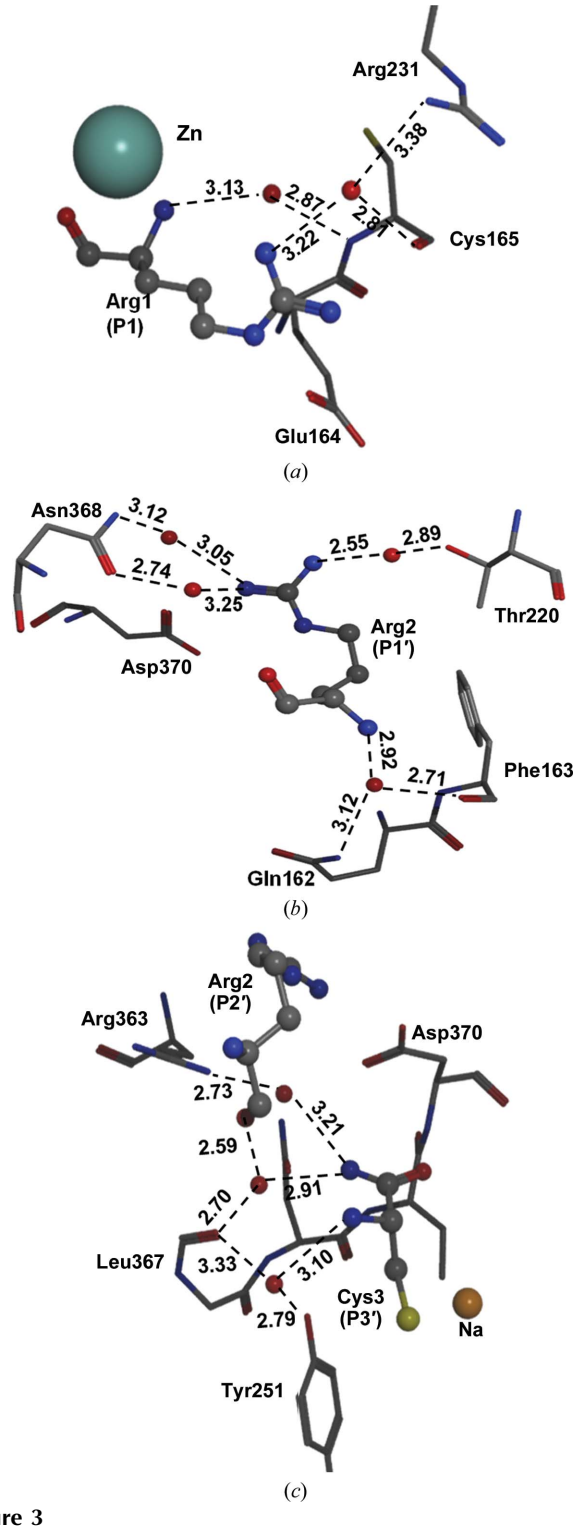

Figure 3
Role of water molecules in the inhibition of Balc424 by peptide inhibitors. (*a*) P1 site, (*b*) P1′ site, (*c*) P3′ site. Balc424 residues are shown as sticks, inhibitor residues and water molecules as ball-and-stick models and Zn as a sphere. Only water-mediated hydrogen bonds are depicted. All atoms are shown in elemental colors.

new functional groups on the inhibitor to mimic these water molecules and their interactions with the Balc424 active site.

Following the docking simulations, a number of promising peptides were obtained and tested for their inhibitory activity

against Balc424 in a UPLC-based assay. We also tested two modified peptides, RFGC with *p*-nitro and *p*-chloro modifications on phenylalanine. A list of all of these peptides and their corresponding IC_{50} values is given in Table 3. The most potent peptide inhibitor RRGF has an IC_{50} of 0.9 μ*M*.

In terms of binding energy, a comparison of peptide inhibitors that we were able to cocrystallize with Balc424 against our simulations shows that the correlation between predicted binding energies of the peptide inhibitors and experimentally deduced IC_{50} values may not be very strong. However, it does seem to be useful in designing inhibitors. RRGF did turn out to be a better inhibitor than RRGC, while RRFC and CRGC showed inhibitory potencies close to that of RRGC. One exception is RYGC, which did not show good inhibition as predicted by docking simulations (Tables 2 and 3), although such false positives are not uncommon.

3.1. Structure–activity relationship (SAR) of position P1

We tested three variations at this position: CRGC, WRGC and QRGC. All of these peptides showed a decrease in inhibitory activity over RRGC, with IC_{50} values of 8, 10 and 14 μ*M*, respectively. This indicates that the salt-bridge interaction between the P1 arginine and Glu164 remains the strongest interaction. However, some inhibitory activity could be reinstated when the P3′ cysteine was replaced by a phenylalanine. This phenylalanine nestles comfortably in a pocket created by hydrophobic and aromatic residues at the active site (CRGF, IC_{50} = 1.5 μ*M*).

3.2. SAR of position P1′

The residue at P1′ has access to the deepest pocket of the otherwise wide-open active site of Balc424 and plays a crucial role in substrate binding. The site is lined with the active-site residues Phe194, Thr215, Thr220, Arg363 and Asp370. The planar guanidinium group of the arginine stacks between Phe194 and Arg363, while forming a salt bridge with Asp370. In order to exploit the stacking interaction of aromatic rings, we replaced the arginine (P1′) with tyrosine, two modified phenylalanines (*p*Cl and *p*NO_2) and tryptophan. Although the stacking interaction of the tyrosine side chain predicted by docking simulations was confirmed by the cocrystal structure (described in a later section), the predicted binding energy does not match the actual inhibitory activity (RYGC, IC_{50} = 300 μ*M*). RYGC turns out to be a poor inhibitor, probably owing to the absence of the salt bridge. The inhibitory activity could be partially recovered by modifications at the *para* position of the phenylalanine [RF(*p*Cl)GC, IC_{50} = 108.9 μ*M*; RF(*p*NO_2)GC, IC_{50} = 123.0 μ*M*]. RWGC showed much better inhibition (IC_{50} = 17.0 μ*M*) over RYGC probably because tryptophan is a larger aromatic residue and also has an NH group in its side chain which participates in forming a hydrogen bond with Asp370, thus restoring some affinity for Asp370. Hence, a positively charged group that can form a salt bridge with Asp370 appears to be essential at the P1′ position.

Table 3
Peptide inhibitors of BoNT/A LC with their corresponding IC_{50} values against BoNT/A as deduced from a UPLC-based assay.

Peptide No.	Peptide	IC_{50}† (μ*M*)
1	RRGC	1.5 ± 0.1
2	RRGF	0.9 ± 0.1
3	CRGF	1.5 ± 0.1
4	RRFC	1.8 ± 0.1
5	RRYC	5.4 ± 0.2
6	CRGC	8.0 ± 1.2
7	WRGC	10.0 ± 1.1
8	QRGC	14.0 ± 1.1
9	RWGC	17.0 ± 0.7
10	RRYY	24.1 ± 4.2
11	RRGT	27.0 ± 1.8
12	RRAT	71.0 ± 4.8
13	RYFC	81.0 ± 13.0
14	RF(*p*Cl)GC	108.9 ± 45.5
15	RF(*p*NO_2)GC	123.0 ± 66.3
16	RYGC	300.0 ± 3.5
17	CRRGF	7.0 ± 0.5
18	CRRGC	43.0 ± 11.4
19	RRGCM	26.1 ± 7.7
20	RRKRLL	28.7 ± 4.1

† Average of five measurements.

3.3. SAR of position P2′

We tested two new tetrapeptides with variations of the glycine at the P2′ position in RRGC to phenylalanine (RRFC) and tyrosine (RRYC) and one in combination with a variation at P1′ (RYFC). While RRFC (IC_{50} = 1.8 μ*M*) showed inhibition similar to that of RRGC (IC_{50} = 1.5 μ*M*), RRYC showed decreased inhibition (IC_{50} = 5.4 μ*M*). Also, RRGT gave better inhibition over RRAT. The cocrystal structure of RRFC with Balc424 showed that there is a flip in the main chain of the peptide and the side chain of phenylalanine is pushed out towards the solvent-exposed area. Also, a tyrosine at P1′ decreased the activity further (RYFC, IC_{50} = 81 μ*M*), probably owing to the lack of the salt-bridge interaction.

3.4. SAR of position P3′

The residue at the P3′ position binds in a pocket lined by the aromatic residues Tyr250, Tyr251, Phe369 and Phe423 and the hydrophobic residues Val69, Val70 and Leu256. The C-terminal amide O atom of the tetrapeptide inhibitors forms a hydrogen bond to the backbone N atom of Asp370. We replaced the cysteine in RRGC by phenylalanine and threonine. While RRGF shows the highest inhibitory activity (IC_{50} = 0.9 μ*M*) of all the peptide inhibitors tested, inhibition by RRGT (IC_{50} = 27 μ*M*) was worse than that by RRGC. Other replacements at this site in combination with variations at the P2′ site (RRAT, IC_{50} = 71 μ*M*; RRYY, IC_{50} = 24.1 μ*M*) resulted in deterioration of inhibitory activity compared with RRGF. Overall, an aromatic/hydrophobic moiety at this site appears to be appropriate for good inhibitory activity given the nature of this pocket.

3.5. Length of peptide inhibitors

In order to deduce the optimum length of an ideal peptide inhibitor for BoNT/A, we tried removing residues as well as

adding new residues at the two termini of the standard tetrapeptide. Removal of the N-terminal arginine (–RGC), the C-terminal cysteine (RRG–) or both (–RR–) led to a substantial decrease in inhibitory activity (Hale *et al.*, 2011). Similarly, the addition of cysteine to the N-terminus of RRGF reduced the inhibitory activity sevenfold (CRRGF, IC_{50} = 7.0 μ*M*) and to that of RRGC reduced the inhibitory activity ∼30-fold (CRRGC, IC_{50} = 43.0 μ*M*). Addition of methionine at the C-terminus of RRGC reduced the inhibitory activity by 17-fold (RRGCM, IC_{50} = 26.1 μ*M*) and a new hexapeptide RRKRLL gave an IC_{50} of 28.7 μ*M*. Hence, a tetrapeptide appears to be the optimum length for further improvement in the inhibitory activity and the development of peptidomimetic drug candidates.

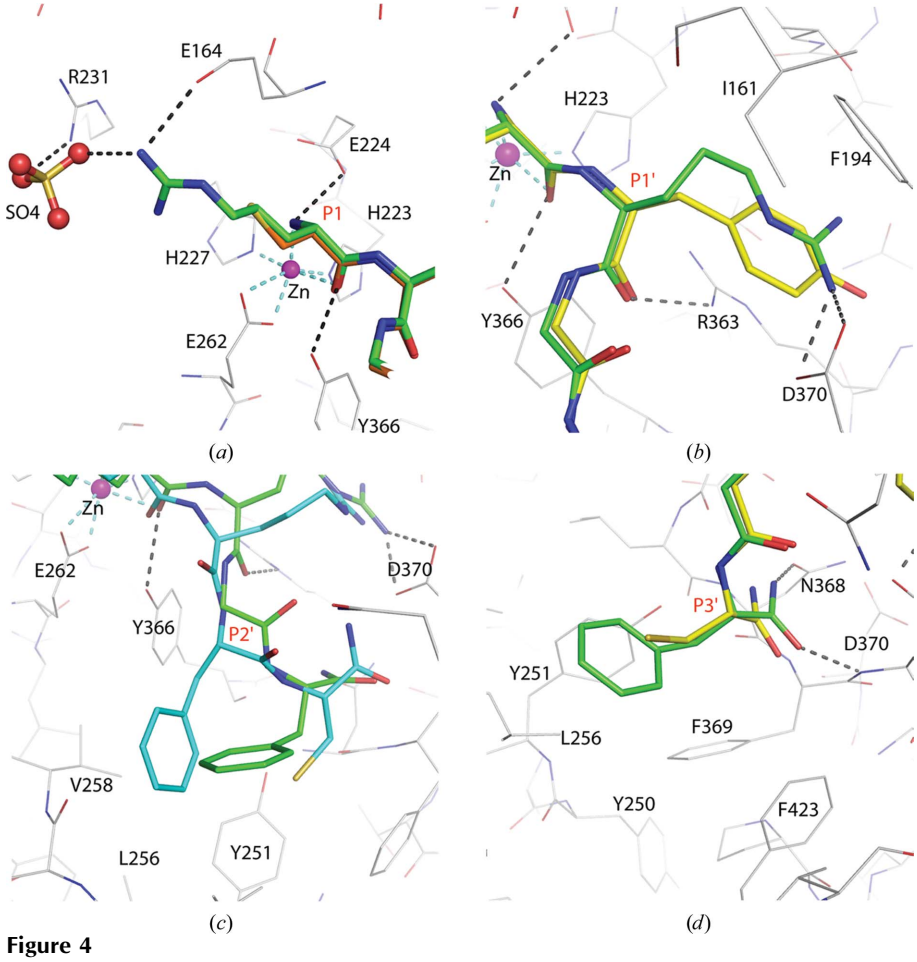

Figure 4
Molecular interactions between peptide inhibitors and Balc424. (*a*), (*b*), (*c*) and (*d*) show close-up views of the interactions between RRGF and Balc424 at the P1, P1′, P2′ and P3′ positions, respectively. Superposition of RRGF with CRGC, RYGC, RRFC and CRGC at the P1, P1′, P2′ and P3′ positions, respectively, are also shown: (*a*) RRGF and CRGC at the P1 position, (*b*) RRGF and RYGC at the P1′ position, (*c*) RRGF and RRFC at the P2′ position, (*d*) RRGF and CRGC at the P3′ position. The enzyme, peptide inhibitors and sulfate ion are shown as line, stick and ball-and-stick models, respectively. The bonds and C atoms of RRGF, CRGC, RYGC and RRFC are shown in green, gold, yellow and cyan, respectively. Magenta, red, blue and brown colors represent Zn, O, N and S atoms, respectively. Coordination bonds and hydrogen bonds are shown in cyan and gray, respectively. *PyMOL* (DeLano, 2002) was used to prepare these figures.

3.6. Structures of peptide-inhibitor complexes and their mode of action

3.6.1. Molecular interactions at the P1 site. In all of the cocrystal structures of Balc424 with peptide inhibitors (RRGF, CRGC, RYGC and RRFC) the overall structure and active-site architecture of the enzyme are similar to the native structure (Kumaran, Rawat, Ludivico *et al.*, 2008). The binding mode and orientation of all of the peptide inhibitors are similar to the previously reported structures (Kumaran, Rawat, Ludivico *et al.*, 2008; Kumaran, Rawat, Ahmed *et al.*, 2008) except for a few differences that are discussed here. In all four complexes the zinc ion is hexa-coordinated by four ligands from the enzyme residues His223, His227 and Glu262 (bidentate) and two from the inhibitors (Fig. 4*a*). The N-terminal free amino group of the peptide inhibitors replaces the nucleophilic water and makes a hydrogen bond to the activating base Glu224. This replacement of the nucleophilic water by the amino group of the inhibitor might be critical for its inhibitory action. The carbonyl O atom of the P1 residue forms a coordination bond to the zinc ion and a hydrogen bond to Tyr366.

In the complex structures of Balc424 with RRGF, RRFC and RYGC, the P1 arginine is exposed to the solvent and its guanidinium group interacts with Glu164 and a bound sulfate ion through ionic interactions (Fig. 4*a*). This sulfate ion further makes a salt bridge with Arg231. The sulfate ion could be from the crystallization condition, which contained ammonium sulfate. In the CRGC complex, the cysteine at the P1 position does not interact with any of the enzyme residues directly, although its thiol group interacts with Asn162 *via* a sodium ion. The sodium ion also interacts with a bound ethylene glycol. Since the P1 cysteine lacks the salt-bridge interaction with Glu164 and does not interact with any of the other residues directly, the CRGC peptide inhibitor shows lower inhibition (IC_{50} = 8.0 μ*M*) compared with the RRGC peptide inhibitor (Kumaran, Rawat, Ludivico *et al.*, 2008).

3.6.2. Molecular interactions at the P1′ site. The P1′ arginine in the peptide inhibitors RRGF, RRFC and CRGC and the P1′ tyrosine in RYGC reside in the S1′ subsite formed by residues such as Asp370, Thr220, Thr215, Arg363, Phe163, Phe164 and Ile161 (Fig. 4*b*). The guanidinium group of the P1′ arginine makes stacking interaction with Arg363 and Phe194 and a salt bridge with Asp370 in the RRGF, RRFC and CRGC complex structures. These interactions are critical for positioning the substrate for catalysis since mutations of Asp370 and Arg363 in the BoNT/A light chain reduce the catalytic

activity (Ahmed *et al.*, 2008; Binz *et al.*, 2002; Chen & Barbieri, 2006). These interactions were also observed in our earlier studies in addition to in the complex structures of L-arginine hydroxamate and the peptidomimetic inhibitors DNP-DAB-RWT-DAB-ML [DNP-DAB, 4-(2,4-dinitrophenylamino)-2-aminobutanoic acid; DAB, 2,4-diaminobutanoic acid] and JTH-NB72-39 (Kumaran, Rawat, Ahmed *et al.*, 2008; Kumaran, Rawat, Ludivico *et al.*, 2008; Silvaggi *et al.*, 2007; Zuniga *et al.*, 2008, 2010). In the *N*-acetyl CRATKML complex structure, the guanidinium group of the P1′ arginine forms a bidentate interaction with the carbonyl O atom of Ile161, a cation–π interaction with Phe194 and a hydrogen bond to the carboxylate group of Asp370 owing to the coordination of the S^{γ} atom of the P1 cysteine to the zinc ion (Silvaggi *et al.*, 2008). In the RRFC structure, the C^{α} position of the P1′ arginine is shifted by about 1.5 Å when compared with the RRGF and CRGC structures, although it maintains the salt-bridge and stacking interactions. This shift in conformation may be a consequence of the effect of phenylalanine substitution at the P2′ position. In the RYGC structure, the P1′ tyrosine occupies the S1′ site formed by the same residues as in other inhibitor complexes but with the loss of the salt-bridge interaction with Asp370. The low inhibition (IC_{50} = 300 μM) may be a consequence of the loss of the salt-bridge interaction since this is critical.

3.6.3. Molecular interactions at the P2′ site. In the RRFC structure, the C^{α} position of the P2′ phenylalanine is shifted by about 1.5 Å towards the surface compared with the other Balc424–tetrapeptide complex structures. The phenyl ring of the P2′ phenylalanine makes hydrophobic contacts with Val258, Tyr251 and Leu256 (Fig. 4*c*). The shift in the backbone position that includes P1′ and P2′ towards the surface and the flipping of the peptide bond may be a consequence of this hydrophobic interaction. This interaction is absent in the RRGF, CRGC and RYGC structures in which the P2′ residue is glycine.

3.6.4. Molecular interactions at the P3′ site. The C-terminal cysteine at P3′ in the case of RRFC, CRGC and RYGC, and the phenylalanine in RRGF bind in the hydrophobic pocket formed by residues Leu256, Tyr250, Tyr251, Arg363, Phe369 and Phe423 (Fig. 4*d*). These hydrophobic interactions, as observed previously, are very similar to those in the Balc424–tetrapeptide complex structures with leucine, isoleucine or methionine at the P3′ position (Kumaran, Rawat, Ludivico *et al.*, 2008). The increase in inhibition activity for RRGF compared with other peptides, including RRGC, may be a consequence of the strength of this hydrophobic interaction with the phenylalanine side chain. The C-terminal amide O and N atoms form hydrogen bonds to the backbone N atom of Asp370 and O atom of Asn368, respectively.

3.7. A pharmacophore model based on peptide inhibitors

Several pharmacophore models have been proposed based on inhibition studies and docking simulations (Burnett, Opsenica *et al.*, 2007; Burnett, Ruthel *et al.*, 2007; Burnett *et al.*, 2003, 2005, 2009; Hermone *et al.*, 2008). Burnett and co-workers performed docking and molecular-dynamics simulations on 2-mercapto-3-phenylpropionyl-RATKML to predict the binding mode and designed a number of analogs. They proposed an improved model containing (i) an aromatic substituent with planar-conjugated positive-ionizable functional groups; (ii) a positive-ionizable component; (iii) two planar pharmacophore groups separated by 6.5–13 Å; (iv) alkenes, imines, amides and azo linkages as pharmacophore planes and (v) a compound length constraint of 23.0 Å. They used this model to screen the NCI compound library, resulting in the identification of some new low-micromolar inhibitors (Burnett, Ruthel *et al.*, 2007). They further improved the pharmacophore model, adding new features, and discovered new 4-amino-7-chloro-quinoline-based inhibitors, but the inhibitory activity did not improve (Burnett, Opsenica *et al.*, 2007), suggesting that the pharmacophore model has not yielded the desired results. The pharmacophore model was then reduced to four components and database searches yielded new inhibitors with similar inhibitory activities (IC_{50} = 3.0 μM; Hermone *et al.*, 2008). In

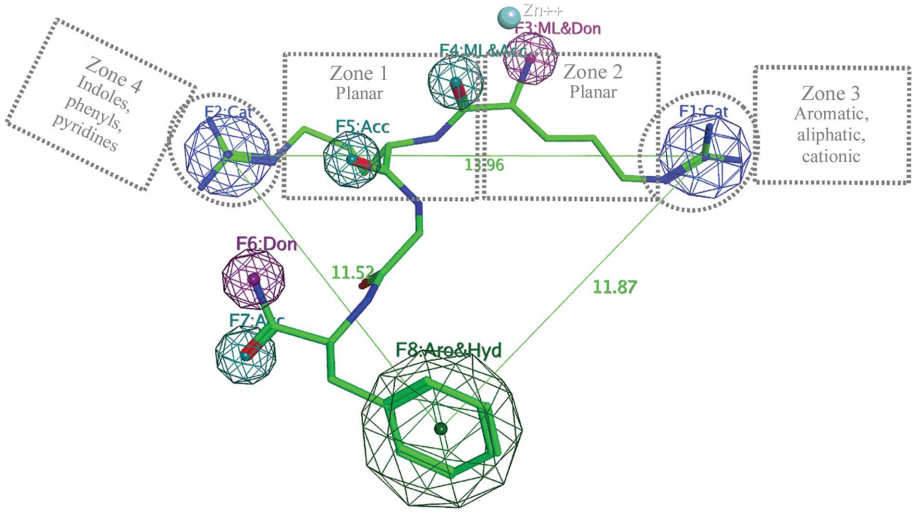

Figure 5
Three-dimensional pharmacophore model showing various features: (i) two positive-ion features (F1:Cat and F2:Cat), with the latter also having a π-stacking interaction; (ii) metal chelators (F3:ML&Don and F4:ML&Acc), with F3 also acting as a hydrogen-bond donor and F4 as a hydrogen-bond acceptor; (iii) a hydrogen-bond acceptor (F5:Acc); (iv) a hydrogen-bond donor–acceptor pair (F6:Don and F7:Acc) and (v) an aromatic/hydrophobic feature (F8:Aro&Hyd). The pharmacophore model was generated using the molecular-modeling program *MOE* (*Molecular Operating Environment*) 2009.10. Linear distances among various pharmacophore features are also shown in Å (green lines). RRGF was used as a reference (stick model) and Zn is shown as a cyan sphere. For comparison, the previously proposed pharmacophore model is shown in gray dashed lines (Nuss *et al.*, 2010).

recent reports, the same pharmacophore model has been improved further and redefined into three zones (Burnett *et al.*, 2009, 2010). Here, we describe a three-dimensional pharmacophore model based on the cocrystal structures of peptide inhibitors with Balc424. The pharmacophore features derived in this model are based not only on the inhibitor functionalities but also on the basis of the specific interactions that these functionalities are involved in as provided by cocrystal structures. The pharmacophore model (Fig. 5) contains (i) two positive-ion features, one of which also has a π-stacking interaction; (ii) metal chelators; (iii) a hydrogen-bond acceptor; (iv) a hydrogen-bond donor/acceptor pair and (v) an aromatic/hydrophobic feature.

Coming to the rationale behind these pharmacophore features, based on the inhibitor–protein cocrystal structures the two positively charged ion features make salt bridges with two negatively charged amino acids (F1 and F2; Fig. 5), the first with Glu164 and the other with Asp370. In the case of the latter there is also a stacking interaction. However, this stacking of the guanidinium group between the side chains of Arg363 and Phe194 is less significant compared with the salt bridge (since the inhibitory activity of RYGC is lower than that of RRGC). In the absence of cocrystal structures with inhibitors, the previously proposed pharmacophore models lacked rationales for the various features and the proposed distance between the two salt bridges also varied. Fig. 5 shows the exact distances required among the prominent features. While metal chelators are not considered to be ideal in drug molecules, here we have observed that in addition to coordinating with the zinc ion the two metal-chelator groups also participate in hydrogen-bonding interactions. The main-chain C$=$O group of the P1′ arginine makes a hydrogen bond to the side-chain OH of Tyr365 and the NH_2 group makes a hydrogen bond to the side-chain carboxylate group of Glu224. The hydrogen-bond acceptor feature (F5) makes a hydrogen bond to the guanidinium group of Arg363. The hydrogen-bond donor–acceptor pair (F6 and F7) is also consistently found across all of our peptide inhibitors and appears to be essential for the binding of peptide inhibitors in an otherwise shallow binding site. The carboxyl group forms a hydrogen bond to the main-chain NH of Asp370 and the terminal amide group forms a hydrogen bond to the side chain of Asn368. The aromatic/hydrophobic feature (F8) is crucial in the sense that it drives the inhibitor to the exposed hydrophobic/aromatic surface of the Balc424 active site; hence, smaller peptides such as RR that lack the hydrophobic/aromatic moiety, although capable of forming salt bridges, show a significant loss of inhibition. The present work based on the three-dimensional structures of the BoNT/A LC active site and the available BoNT/A LC–inhibitor complexes adds new features to the BoNT/A pharmacophore model.

In summary, using the SBDD approach assisted by docking simulations we designed and tested a number of natural and modified peptide inhibitors of BoNT/A. Since these tetrapeptide inhibitors have good solubility and specificity, excellent access to their site of action inside the neuronal cells and are stable in the cellular environment (Hale *et al.*, 2011), they make good candidate molecules for drug discovery. The structure–activity relationship provides important guidelines for the design of better peptidomimetic as well as nonpeptide small-molecule inhibitors of BoNT/A LC. We also determined cocrystal structures of some of these peptides. Based on the inhibition studies and cocrystal structures, we proposed a three-dimensional pharmacophore model that would be useful for the screening of small-molecule libraries for the identification of lead molecules.

This research was supported by an award from DTRA BO742081 under DOE prime contract No. DEAC02-98CH10886 with Brookhaven National Laboratory. We thank PXRR at the NSLS (BNL) for data-collection facilities.

References

Agarwal, R., Binz, T. & Swaminathan, S. (2005*a*). *Biochemistry*, **44**, 8291–8302.

Agarwal, R., Binz, T. & Swaminathan, S. (2005*b*). *Biochemistry*, **44**, 11758–11765.

Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T. & Swaminathan, S. (2004). *Biochemistry*, **43**, 6637–6644.

Agarwal, R., Schmidt, J. J., Stafford, R. G. & Swaminathan, S. (2009). *Nature Struct. Mol. Biol.* **16**, 789–794.

Agarwal, R. & Swaminathan, S. (2008). *J. Biol. Chem.* **283**, 25944–25951.

Ahmed, S. A., Olson, M. A., Ludivico, M. L., Gilsdorf, J. & Smith, L. A. (2008). *Protein J.* **27**, 151–162.

Arndt, J. W., Chai, Q., Christian, T. & Stevens, R. C. (2006). *Biochemistry*, **45**, 3255–3262.

Arndt, J. W., Yu, W., Bi, F. & Stevens, R. C. (2005). *Biochemistry*, **44**, 9574–9580.

Binz, T., Bade, S., Rummel, A., Kollewe, A. & Alves, J. (2002). *Biochemistry*, **41**, 1717–1723.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Südhof, T. C., Jahn, R. & Niemann, H. (1994). *J. Biol. Chem.* **269**, 1617–1620.

Black, J. D. & Dolly, J. O. (1986*a*). *J. Cell Biol.* **103**, 521–534.

Black, J. D. & Dolly, J. O. (1986*b*). *J. Cell Biol.* **103**, 535–544.

Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H. & Jahn, R. (1993). *EMBO J.* **12**, 4821–4828.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Burnett, J. C., Li, B., Pai, R., Cardinale, S. C., Butler, M. M., Peet, N. P., Moir, D., Bavari, S. & Bowlin, T. (2010). *Open Access Bioinformatics*, **2010**, 11–18.

Burnett, J. C., Opsenica, D., Sriraghavan, K., Panchal, R. G., Ruthel, G., Hermone, A. R., Nguyen, T. L., Kenny, T. A., Lane, D. J., McGrath, C. F., Schmidt, J. J., Vennerstrom, J. L., Gussio, R., Solaja, B. A. & Bavari, S. (2007). *J. Med. Chem.* **50**, 2127–2136.

Burnett, J. C., Ruthel, G. *et al.* (2007). *J. Biol. Chem.* **282**, 5004–5014.

Burnett, J. C., Schmidt, J. J., McGrath, C. F., Nguyen, T. L., Hermone, A. R., Panchal, R. G., Vennerstrom, J. L., Kodukula, K., Zaharevitz, D. W., Gussio, R. & Bavari, S. (2005). *Bioorg. Med. Chem.* **13**, 333–341.

Burnett, J. C., Schmidt, J. J., Stafford, R. G., Panchal, R. G., Nguyen, T. L., Hermone, A. R., Vennerstrom, J. L., McGrath, C. F., Lane, D. J., Sausville, E. A., Zaharevitz, D. W., Gussio, R. & Bavari, S. (2003). *Biochem. Biophys. Res. Commun.* **310**, 84–93.

Burnett, J. C., Wang, C., Nuss, J. E., Nguyen, T. L., Hermone, A. R., Schmidt, J. J., Gussio, R., Wipf, P. & Bavari, S. (2009). *Bioorg. Med. Chem. Lett.* **19**, 5811–5813.

Cai, S. & Singh, B. R. (2007). *Infect. Disord. Drug Targets*, **7**, 47–57.
Chen, S. & Barbieri, J. T. (2006). *J. Biol. Chem.* **281**, 10906–10911.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Hale, M., Oyler, G., Swaminathan, S. & Ahmed, S. A. (2011). *J. Biol. Chem.* **286**, 1802–1811.
Hermone, A. R., Burnett, J. C., Nuss, J. E., Tressler, L. E., Nguyen, T. L., Solaja, B. A., Vennerstrom, J. L., Schmidt, J. J., Wipf, P., Bavari, S. & Gussio, R. (2008). *ChemMedChem*, **3**, 1905–1912.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Josko, D. (2004). *Clin. Lab. Sci.* **17**, 30–34.
Kumaran, D., Rawat, R., Ahmed, S. A. & Swaminathan, S. (2008). *PLoS Pathog.* **4**, e1000165.
Kumaran, D., Rawat, R., Ludivico, M. L., Ahmed, S. A. & Swaminathan, S. (2008). *J. Biol. Chem.* **283**, 18883–18891.
Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R. & Stevens, R. C. (1998). *Nature Struct. Biol.* **5**, 898–902.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Marks, J. D. (2004). *Mov. Disord.* **19**, S101–S108.
Mehler, E. L. & Solmajer, T. (1991). *Protein Eng.* **4**, 903–910.
Meunier, F. A., Lisk, G., Sesardic, D. & Dolly, J. O. (2003). *Mol. Cell. Neurosci.* **22**, 454–466.
Montecucco, C. (1986). *Trends Biochem. Sci.* **11**, 315–317.
Montecucco, C. & Schiavo, G. (1995). *Q. Rev. Biophys.* **28**, 423–472.
Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998). *J. Comput. Chem.* **19**, 1639–1662.
Nuss, J. E., Dong, Y., Wanner, L. M., Ruthel, G., Wipf, P., Gussio, R., Vennerstrom, J. L., Bavari, S. & Burnett, J. C. (2010). *ACS Med. Chem. Lett.* **1**, 301–305.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Park, J. G., Sill, P. C., Makiyi, E. F., Garcia-Sosa, A. T., Millard, C. B., Schmidt, J. J. & Pang, Y.-P. (2006). *Bioorg. Med. Chem.* **14**, 395–408.
Rowe, B., Schmidt, J. J., Smith, L. A. & Ahmed, S. A. (2010). *Anal. Biochem.* **396**, 188–193.
Sanner, M. F. (1999). *J. Mol. Graph. Model.* **17**, 57–61.
Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R. & Montecucco, C. (1992). *Nature (London)*, **359**, 832–835.
Schiavo, G., Malizio, C., Trimble, W. S., Polverino de Laureto, P., Milan, G., Sugiyama, H., Johnson, E. A. & Montecucco, C. (1994). *J. Biol. Chem.* **269**, 20213–20216.
Schiavo, G., Rossetto, O., Benfenati, F., Poulain, B. & Montecucco, C. (1994). *Ann. N. Y. Acad. Sci.* **710**, 65–75.
Schiavo, G., Rossetto, O., Catsicas, S., Polverino de Laureto, P., DasGupta, B. R., Benfenati, F. & Montecucco, C. (1993). *J. Biol. Chem.* **268**, 23784–23787.
Schiavo, G., Shone, C. C., Rossetto, O., Alexander, F. C. & Montecucco, C. (1993). *J. Biol. Chem.* **268**, 11516–11519.
Silvaggi, N. R., Boldt, G. E., Hixon, M. S., Kennedy, J. P., Tzipori, S., Janda, K. D. & Allen, K. N. (2007). *Chem. Biol.* **14**, 533–542.
Silvaggi, N. R., Wilson, D., Tzipori, S. & Allen, K. N. (2008). *Biochemistry*, **47**, 5736–5745.
Simpson, L. L. (1989). *Botulinum Neurotoxins and Tetanus Toxin*. New York: Academic Press.
Swaminathan, S. & Eswaramoorthy, S. (2000). *Nature Struct. Biol.* **7**, 693–699.
Zuniga, J. E., Hammill, J. T., Drory, O., Nuss, J. E., Burnett, J. C., Gussio, R., Wipf, P., Bavari, S. & Brunger, A. T. (2010). *PLoS One*, **5**, e11378.
Zuniga, J. E., Schmidt, J. J., Fenn, T., Burnett, J. C., Araç, D., Gussio, R., Stafford, R. G., Badie, S. S., Bavari, S. & Brunger, A. T. (2008). *Structure*, **16**, 1588–1597.