Toxicon xxx (2017) 1–8

Contents lists available at ScienceDirect

Toxicon

journal homepage: www.elsevier.com/locate/toxicon

The hypothetical protein P47 of *Clostridium botulinum* E1 strain Beluga has a structural topology similar to bactericidal/permeability-increasing protein

Kwok-ho Lam [a, 1], Ruifeng Qi [a, 1], Shun Liu [a], Amelie Kroh [b], Guorui Yao [a], Kay Perry [c], Andreas Rummel [b], Rongsheng Jin [a, *]

[a] Department of Physiology and Biophysics, University of California, Irvine, CA, USA
[b] Institut für Toxikologie, Medizinische Hochschule Hannover, Hannover, Germany
[c] NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Argonne National Laboratory, Argonne, IL, USA

ARTICLE INFO

Article history:
Received 27 July 2017
Accepted 13 October 2017
Available online xxx

Keywords:
Botulinum neurotoxin
Crystal structure
Bactericidal/permeability-increasing protein
Toxin complex
orfX gene cluster
P47
Lipid binding

ABSTRACT

Botulinum neurotoxins (BoNTs) are causative agents of the life-threatening disease botulism. They are naturally produced by species of the bacteria *Clostridium botulinum* as stable and non-covalent complexes, in which the BoNT molecule is assembled with several auxiliary non-toxic proteins. Some BoNT serotypes, represented by the well-studied BoNT serotype A (BoNT/A), are produced by *Clostridium* strains that carry the *ha* gene cluster, which encodes four neurotoxin-associated proteins (NTNHA, HA17, HA33, and HA70) that play an important role to deliver and protect BoNTs in the gastrointestinal tract during oral intoxication. In contrast, BoNT/E- and BoNT/F-producing strains carry a distinct gene cluster that encodes five proteins (NTNHA, P47, OrfX1, OrfX2, and OrfX3, termed the *orfX* cluster). The structures and functions of these proteins remain largely unknown. Here, we report the crystal structure of P47 resolved at 2.8 Å resolution. Surprisingly, P47 displays a structural topology that is similar to bactericidal/permeability-increasing (BPI) like proteins, which were previously identified only in eukaryotes. The similarity of a hydrophobic cleft of P47 with the phospholipid-binding groove of BPI suggests that P47 might be involved in lipid association to exert its function. Consistently, P47 associates and induces aggregation of asolectin-containing liposomes in a protein- and lipid-concentration dependent manner. These findings laid the foundation for future structural and functional studies of the potential roles of P47 and OrfX proteins in facilitating oral intoxication of BoNTs.

1. Introduction

Botulinum neurotoxins (BoNTs) are among the most poisonous toxins, causing the life-threatening disease botulism. They specifically target the presynaptic neurons at neuromuscular junctions, where they degrade SNAREs (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors) to block neurotransmission and cause muscle paralysis. BoNTs are classified by Centers for Disease Control and Prevention (CDC) as tier I select agents due to their potential use as bioweapons. Nevertheless, BoNT/A is one of the top selling drugs that is widely used for the treatment of several neuromuscular diseases and as popular cosmetics.

There are seven major serotypes of BoNTs (designated A to G), among which BoNT/A, B, E and F are detected in human patients (Rossetto et al., 2014). Naturally occurring BoNT is synthesized together with several nontoxic neurotoxin-associated proteins (NAPs) in the form of a progenitor toxin complex (PTC). All BoNTs carry non-toxic-non-haemagglutinin (NTNHA) that forms a 1:1 heterodimer with BoNT, the minimally functional PTC (M-PTC) (Gu et al., 2012; Sagane et al., 2012; Gu and Jin, 2013; Matsui et al., 2014; Eswaramoorthy et al., 2016). NTNHA is crucial to protect the toxin in the acidic and protease-rich gastrointestinal (GI) environment during oral intoxication. In addition, BoNT serotypes A1, B, C, D, G, which are encoded in the *ha* toxin gene cluster, are produced alongside with hemaglutinin-17, 33 and 70 (HA17, 33, 70), which further assemble with the M-PTC to form the large PTC (L-PTC) (Amatsu et al., 2013; Benefield et al., 2013; Lee et al., 2013, 2014a,b;

* Corresponding author.
E-mail address: r.jin@uci.edu (R. Jin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.toxicon.2017.10.012
0041-0101/

Yao et al., 2014). For BoNT/A, the HA proteins assemble into a dodecameric ~470 kDa complex that facilitates absorption of BoNT/A in the intestine by recognizing cell-surface glycans on the intestinal epithelium and disrupting the integrity of the E-cadherin-mediated epithelial cell junction (Sugawara et al., 2010; Lee et al., 2014a,b; Lee et al., 2015; Lam and Jin, 2015). Intriguingly, bacteria that express BoNT/E, F, and A2-A4 do not encode HA proteins (Jacobson et al., 2008; Popoff and Connan, 2014; Williamson et al., 2016). Instead they carry a distinct gene cluster, known as the *orfX* toxin gene cluster that encodes NTNHA as well as four other proteins (OrfX1, 2, 3, and P47). The structural and functional roles of these proteins remain poorly understood.

P47 is a hypothetical protein with a molecular weight of ~47 kDa, whose gene is positioned upstream of *ntnha* in the *orfX* gene cluster in BoNT/E1-E12 producing *C. botulinum* (Li et al., 1998; Hill and Smith, 2013). *P47* is the only gene co-transcribed with *ntnha* and *bont/E* and is selectively expressed at the transition phase together with *bont/E* (Dineen et al., 2004; Kubota et al., 1998), suggesting they may be involved in BoNT/E function. *C. botulinum* transformed with *p47* antisense mRNA did not affect BoNT/E synthesis, indicating it unlikely serves as a regulator of the operon (Couesnon et al., 2006). P47 was detected in the crude extract of BoNT/A2-producing *C. botulinum*, suggesting it might be secreted, although this is yet to be reported for BoNT/E or F (Lin et al., 2010; Hines et al., 2005).

Further upstream of *p47* are three hypothetical proteins known as OrfX1, 2, and 3 encoded in a separate operon of *C. botulinum*. It was suggested that these proteins encoded in the *orfX* cluster may directly or indirectly interact with M-PTC/E (Kukreja and Singh, 2007), but the physiological relevance of such interactions and the biological functions of OrfXs and P47 remain mysterious. Interestingly, amino acid sequence analyses of OrfX2 and OrfX3 showed that they have internal segments that share similarities to P47, suggesting that they may carry a P47-like protein fold (Lin et al., 2010). Therefore, we decided to focus our research on P47 first, which will provide new insights into the function of the whole *orfX* gene cluster.

2. Materials and methods

2.1. Cloning, expression and purification

Genomic DNA of *Clostridium botulinum* E1 strain Beluga (Accession code: ACSC01000002) was isolated. P47 gene (locus_tag CLO_2650) was amplified by PCR using the corresponding primers and cloned into the pQE (Qiagen) expression vector with an N-terminal His_6 tag fused to a thrombin cleavage site and a C-terminal Strep tag, yielding pH6tP47-ES. The recombinant H6tP47-ES (Protein ID ZP_04823541.1) was expressed in the *E. coli* strain BL21-star (DE3) (Calsbad, CA, USA). Bacteria were grown at 37 °C in LB (Luria-Bertani) medium containing 100 μg/ml ampicillin. The temperature was decreased to 18 °C when OD_{600} reached 0.4–0.6. Expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 16 h. The cells were harvested by centrifugation and stored at −20 °C until use. The selenomethionine (Se-Met) labeled H6tP47-ES was expressed as previously described (Van Duyne et al., 1993). Bacteria were grown in M9 minimal media supplemented with amino acids (lysine, threonine, phenylalanine, leucine, isoleucine, and valine) and Se-Met.

For purification of native and Se-Met labeled pH6tP47-ES, proteins were bound to a Ni-NTA (nitrilotriacetic acid, Qiagen) affinity column in a buffer containing 50 mM Tris, pH 8.0, 400 mM NaCl, 10 mM imidazole, and 1 mM TCEP, and subsequently eluted in the same buffer containing 300 mM imidazole. The eluted fractions were pooled and dialyzed for 16 h at 4 °C against a buffer containing 20 mM Tris, pH 8.8, 20 mM NaCl, and 1 mM TCEP and then further purified by MonoQ ion-exchange chromatography (GE Healthcare) in a buffer containing 20 mM Tris, pH 8.8 and eluted with a NaCl gradient. The peak fractions were then subjected to Superdex 200 size-exclusion chromatography (SEC) in a buffer containing 20 mM Hepes, pH 6.8, 150 mM NaCl, 1 mM TCEP and subsequently concentrated to ~7 mg/ml using Amicon Ultra centrifugal filters (Millipore) and stored at −80 °C until use.

2.2. Crystallization and X-ray diffraction data collection

Initial crystallization screens for P47 with the intact His- and Strep-tag were carried out using a Gryphon crystallization robot (Art Robbins Instrument, Sunnyvale, CA, USA) and commercial high-throughput crystallization screen kits from Hampton Research or Qiagen. After extensive manual optimization, the best P47 crystals were grown by hanging-drop vapor diffusion at 18 °C, in which the protein (2.5 mg/ml) was mixed in 1:1 (v/v) ratio with a reservoir solution containing 12% Polyethylene glycol (PEG) 6000 (Hampton Research), 0.1 M Hepes, pH 7.5. Micro-seeding was employed to obtain single crystals.

The P47 crystals were cryoprotected in the original mother liquor supplemented with 15% (v/v) ethylene glycol and flash-frozen in liquid nitrogen. The crystals belong to space group $P4_12_12$, with unit cell dimensions of a = b = 74.08 Å, c = 308.18 Å; $\alpha = \beta = \gamma = 90°$. The best native data set was collected at 2.8 Å. The X-ray diffraction data were collected at 100 K at beam line 24-ID-E, Advanced Photon Source (APS), using detector ADSC Q315. All data sets were processed and scaled using XDS (Kabsch, 2010) and CCP4i (Potterton et al., 2003). Data collection statistics are summarized in Table 1.

2.3. Structure determination

Using the Se-Met derivative crystals, we solved the structure of P47 to 2.8 Å by SAD (single-wavelength anomalous dispersion)

Table 1
Data collection and refinement statistics.

	Se-Met P47	P47
Data collection		
Wavelength (Å)	0.97918	0.97918
Resolution (Å)	51.34–2.88 (3.04–2.88)	43.32–2.80 (2.95–2.80)
Space group	$P\,4_12_12$	$P\,4_12_12$
Cell dimensions		
a, *b*, *c* (Å)	74.1, 74.1, 308	74.1, 74.1, 308.2
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90
Completeness (%)	100.0 (100.0)	99.08 (100.0)
Redundancy	6.8 (6.9)	5.5 (5.6)
R_{merge} (%)	15.9 (143.3)	10.2 (166.3)
Mean $I/\sigma(I)$	14.4 (2.4)	12.48 (1.6)
Autosol FOM	0.281	
Refinement		
Resolution (Å)		43.32–2.80
No. reflections		21, 917
R_{work}/R_{free} (%)		23.10/27.48
No. atoms		
Protein		3, 250
Ligand/ion		0
Water		0
B-factors (Å^2)		
Protein		105.4
Ligand/ion		–
Water		–
r.m.s. deviations		
Bond lengths (Å)		0.009
Bond angles (°)		1.246

Values in parentheses are for the highest resolution shell.

Please cite this article in press as: Lam, K.-h., et al., The hypothetical protein P47 of *Clostridium botulinum* E1 strain Beluga has a structural topology similar to bactericidal/permeability-increasing protein, Toxicon (2017), https://doi.org/10.1016/j.toxicon.2017.10.012