Biochemical and Biophysical Research Communications 407 (2011) 400–405

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Subsite specificity of anthrax lethal factor and its implications for inhibitor development

Feng Li [a,b,1], Simon Terzyan [a], Jordan Tang [a,*]

[a] Protein Studies, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA
[b] Department of Biochemistry and Molecular Biology, University of Oklahoma Health Science Center, Oklahoma City, OK 73104, USA

ARTICLE INFO

Article history:
Received 24 February 2011
Available online 21 March 2011

Keywords:
Bacillus anthracis
Lethal factor
Specificity
Inhibitor

ABSTRACT

The lethal factor of *Bacillus anthracis* is a major factor for lethality of anthrax infection by this bacterium. With the aid of the protective antigen, lethal factor gains excess to the cell cytosol where it manifests toxicity as a metalloprotease. For better understanding of its specificity, we have determined its residue preferences of 19 amino acids in six subsites (from P3 to P3′) as relative k_{cat}/K_m values (specificity constants). These results showed that lethal factor has a broad specificity with preference toward hydrophobic residues, but not charged or branched residues. The most preferred residues in these six subsites are, from P1 to P3′, Trp, Leu, Met, Tyr, Pro, and Leu. The result of residue preference was used to design new substrates with superior hydrolytic characteristics and inhibitors with high potency. For better use of the new findings for inhibitor design, we have modeled the most preferred residues in the active site of lethal factor. The observed interactions provide new insights to future inhibitor designs.

1. Introduction

Anthrax is a disease caused by the infection of *Bacillus anthracis*, a Gram-positive spore-forming bacterium usually found in the soil [1]. The infectious spores of the bacterium can enter human body through the gastrointestinal tracts (ingestion), skin (cutaneous), and lungs (inhalation) and produce distinct clinical symptoms [2]. Inhalational anthrax is the most dangerous form to human and is usually fatal. There has been an increasing interest in the development of treatment for anthrax especially in view that *B. anthracis* spores have been used as a bio-weapon. Therefore, the understanding of the lethal mechanism of anthrax is of basic scientific importance.

Once *B. anthracis* spores are inhaled into the host, they rapidly germinate and proliferate in the circulation system. Vegetative *B. anthracis* secretes three plasmid-encoded toxin proteins named protective antigen (PA), lethal factor (LF) and edema factor (EF). They work together to cause most of the pathological consequences in the host. PA can combine with LF to form lethal toxin (LeTx) or with EF to form edema toxin (EdTx). These complexes gain entrance to cells through receptor binding to PA, a process ultimately delivers LF and EF to the cytosol. The pathological activities of these toxins are manifested in the cytosol by their enzymic activities. EF is an adenylate cyclase which causes increased level of cAMP in the cells. LF is a metalloproteinase and is by far the most toxic component of the *B. anthracis* infection [3,4]. The administration of low dosage of LT is lethal to experimental animals [1]. In human inhalation anthrax, the elimination of bacteria by antibiotics was frequently insufficient to rescue the patients [5]. Such clinical failure was thought to have caused by the presence of active LF in the cells. The best established cellular targets of LF are the members of MAP kinase kinase (MAPKK) family and the inactivation of these enzymes may account for some of the toxicity of LF [6–8]. However, other protein substrates of LF have also been proposed [9]. These observations suggest that LF is a potential therapeutic target of anthrax for the development of small molecular inhibitor drugs and the full understanding of LF specificity would be beneficial to this end.

The catalytic unit of LF which performs substrate recognition and hydrolysis is formed by three of the four domains in LF. The catalytic active site comprises a bound Zn atom and three histidine side chains. From the crystal structure of substrate peptide bound to LF [10], the binding cleft is large enough to accommodate several amino acid residues (subsites). Preliminary specificity of the subsites has been derived from the alignment of sequences around the LF cleavage sites of MAPKK enzymes [11]. The lack of clear consensus residues in the subsites (Table 1) suggests that LF has a broad amino acid preference in nearly all the subsites, an assumption supported by kinetic data on synthetic peptide substrates [12]. Although peptide inhibitors of LF based on cleavage site sequence of MAPKKs have shown good potency [13], they are too large in

* Corresponding author. Fax: +1 405 271 7249.
E-mail addresses: fengli6870@gmail.com (F. Li), Jordan-tang@omrf.org (J. Tang).
[1] Department of Microbiology and Molecular Genetics, University of Texas Health Science Center, 6431 Fannin Street, Houston, TX 77030, USA.

0006-291X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2011.03.033

Table 1
Comparison of the most preferred amino acids with those of LF biological substrates.

	P3	P2	P1	P1′	P2′	P3′
MAPKK-1	P	T	P	I	Q	L
MAPKK-2	V	L	P	A	L	T
MAPKK-3	D	L	R	I	S	C
MAPKK-4	A	L	K	L	N	F
MAPKK-6	G	L	K	I	P	K
MAPKK-7	T	L	Q	L	P	S
Most preferred amino acid	W(100) F (29) Y (20)	L(100) F(68) Y(42)	M(100) A(70) P(68)	Y(100) L(66) I(39)	P(100) A(72) L(69)	L(100) I(81) M(76)

LF cleaves it substrates between P1 and P1′ (scissile bond). The positions on the left side of the scissile bond are named P1, P2, P3 and so on; and those on the right sides are named P1′, P2′, P3 and so on. For each subsite, the amino acid with the highest initial velocity was assigned a number of 100, and others are the percentage of their initial velocity versus the highest one.

molecular size to be useful in clinical settings. Detailed knowledge on LF subsite specificity would provide insights for the design of small potent inhibitors with pharmacological potentials. Here we report the residue preferences in six subsites of LF (from P3 to P3′) determined as kinetic parameter, relative k_{cat}/K_m. The protein–substrate interactions was also studied by molecular modeling of binding modes of the most preferred residues in these subsites.

2. Materials and methods

2.1. Design of the substrate mixtures

Peptide mixtures were designed and synthesized based on a peptide template as RGKKKVLR* ILLN (in which the star denotes the cleaving site) which was known to be cleaved by LF. For characterization of each of the six subsites studied, a peptide mixture composed of 19 equal molar peptides which were differed only by one amino acid at a single subsite was designed and synthesized in an appropriate cycle of solid-phase peptide synthesis (Synpep, Dublin, CA). Because limiting the number of peptides in a mixture facilitated their identification [14], the 19 peptides were grouped into four sets of substrate mixtures according to their molecular weights of all the amino acid studied. Thus, 24 substrate mixtures in total were required for characterization of all the six subsites. A substrate with known k_{cat}/K_m will be added to each mixture to work as an internal standard.

2.2. Initial rate determination

Substrate mixtures were dissolved in DMSO at 10 mM as stocks and were further diluted to 10 μM in 20 mM HEPES, pH 7.4, 0.1 mg ml^{-1} BSA. Reactions were initiated at 37 °C by the addition of 10 nM LF. Aliquots were removed at different time points and reactions were stopped by the addition of formic acid (10% of reaction system in volume). Quantitative analysis was conducted by ESI-LC MS. The system was composed of an Agilent 1100 HPLC, a Clipeus 1 × 50 mm 5 mm C-18 column, and a Bruker Microtof ESI-LC MS. The HPLC buffers used were: A – 99.5% H_2O, 0.5% formic acid, and B – 99.5% ACCN, 0.5% formic acid. Separations were conducted using a 5–50% B gradient over 4 min at a flow rate of 200 ml/min. Ion detection was accomplished using the time of flight instrument in positive reflector mode with ion detection between 200 and 2000 m/z through an ESI interface. Data were analyzed by Quant Analysis software equipped with the ESI mass spectrometer to obtain peak areas of the substrates and their corresponding products in a given reaction.

Each of the libraries was incubated with LF and the hydrolysis was analyzed in ESI-LC MS. To work out the hydrolytical conditions and monitor procedures, a mode peptide with the sequence of RGKKKVLRILLN was incubated with LF and analyzed. In order to compare their hydrolysis rates, relative product formation was calculated as a ratio of the peak area of a product to the sum of the areas of both the product and its corresponding substrate. Data were plotted in terms of relative product formation as a function of reaction incubation time. Initial velocity was obtained from nonlinear regression analysis representing the initial 15% formation of product. The relative initial velocity of different substrate is equal to the relative k_{cat}/K_m.

2.3. Determination of kinetic parameters

The kinetic experiments for investigating K_m and k_{cat} for synthesized LF fluorogenic substrates were performed in a buffer of 0.1 M HEPES, pH 7.4 at 37 °C. Reactions were performed in black flat-bottomed 96-well plates by mixing LF (10 nM) with varying concentrations of the fluorogenic substrates. The hydrolysis of substrate was monitored by continuously measuring the increase of fluorescence intensity using a TECAN 200, a fluorescence microplate reader. An excitation wavelength of 325 nm and an emission wavelength of 393 nm were used to monitor the changes of fluorescence intensity. The reaction rate of substrate hydrolysis for each substrate concentration was obtained as the initial velocity which was calculated as the ratio of initial 15% product formation to the reaction time. The final reaction curve was plotted as reaction rate as a function of substrate concentrations. The kinetic parameters, K_m and $V_{max,}$ were gained by fitting the reaction curve using GraFit 5 (Erithacus Software, Horley, Surrey, UK), a nonlinear regression analysis software. K_{cat} was calculated as the ratio of V_{max} to the initial LF concentration [14,15].

2.4. Investigation of inhibition constants of LF inhibitors

Both the kinetic assays for investigating inhibition constants (K_i) and data analysis were similar to 2.3 except that investigation of inhibition constants of LF inhibitors required the presence of LF inhibitor at different concentrations. The final result was plotted as different initial velocities as a function of varying concentrations of LF inhibitor. The K_i values were obtained by fitting the curve using software of GraFit 5 [16].

2.5. Molecular modeling of subsite binding

A hexapeptide containing the most preferred residues was modeled into the active site of LF using the X-ray structure of LF complexed with a MAPKK peptide (PDB access code 1pww) [12]. Firstly, the side chains of the original peptide were replaced with the optimized residues and then the substrate and surrounding LF residues (residues A297-I345, L368-Q417, P551-S776) were subjected to 20 cycles of conjugate gradient energy minimization calculations using program CNS/_Crystallography & NMR System [17]. During the process, the total energy decreased from 3058.5 to 2767.7, finally, the residues were fixed and examined for their contacts in graphic models.

3. Results

3.1. Determination of substrate side-chain preference

We examined the residue preference of LF on six central subsites, S3, S2, S1, S1′, S2′, and S3′, which bind three residues on each side of the scissile bond. These subsites are the major determinants for substrate binding of proteases in general and the interaction of small potent inhibitors would not be expected to involve subsites