Crystal Structures of Anthrax Toxin Lethal Factor Bound to an Optimized Substrate and Candidate Small Molecule Inhibitors

by Thiang Yian Wong, Robert Schwarzenbacher and Robert C. Liddington
The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Anthrax Toxin, together with its bacterial capsule, is a major virulence factor in Anthrax[1]. The virulent strain of *Bacillus anthracis* is an encapsulated gram-positive, rod-shaped, spore-forming bacterium that produces and exports the three Anthrax Toxin proteins, Protective Antigen (PA), Lethal Factor (LF) and Edema Factor (EF). The infectious agent of anthrax, the bacillus spore, may be introduced into a mammalian host through inhalation, ingestion and subcutaneous wounds. Once in a susceptible host organism, the spores germinate within macrophages and their vegetative cells proliferate rapidly[2]. The explosive release of bacilli from the host macrophages simultaneously introduces the three toxin components into the bloodstream. Continuous rapid bacterial proliferation and toxin production by the bacilli could lead to very high concentrations of toxin accumulation in the bloodstream. Although antibiotics may be effective in clearing the bacteria from the bloodstream, high level of toxins may remain in the blood circulation for a further three to four days. In the bloodstream, PA attaches to the host cell surface via the anthrax toxin receptor (ATR)[3]. Furin, a cell surface protease, then removes a 20kD fragment of PA and initiates the formation of the pore-forming complex of the anthrax toxin. Seven copies of the processed PA (now a 63 kD form of PA, PA_{63}), together with seven copies of the ATR, form a seven-membered ring[4]. The PA_{63} heptamers combine with either of two enzymes, LF (a zinc metalloprotease) or EF (an adenylate cyclase), to form either of the two anthrax toxins, Lethal Toxin (LeTx) or Edema Toxin (EdTx). The whole toxin assembly is taken into an intracellular compartment through "endocytosis". Acidification within the endosome through natural cell processes triggers a

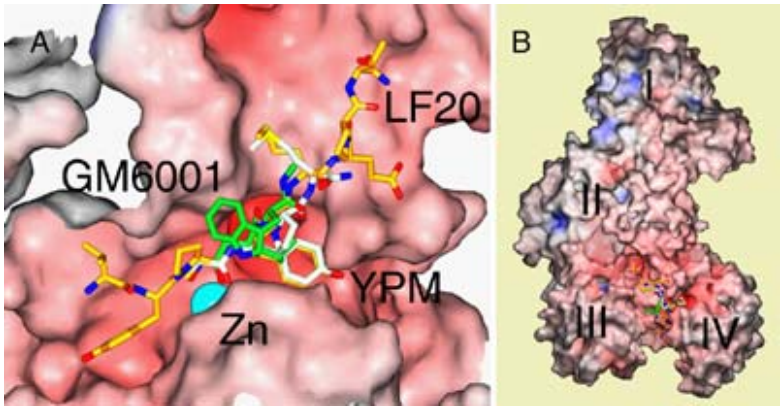

Figure 1. Crystal structures of LF-LF20, LF-thioacetylYPM-zinc, and LF-GM6001-zinc complexes.

The molecular surface of LF is colored by charge (red = negative; blue = positive), with Zn^{2+} shown as a solid sphere (cyan) and the model of the peptide or inhibitor shown in stick representation. (A) The superposed individual complex structures of the three target molecules: LF20 (yellow gold) in the absence of Zn^{2+}, resolution limit 2.85 Å; SHAc-YPM (tinted white, labeled "YPM"), resolution limit 3.50 Å; and GM6001 (green), resolution limit 2.70 Å. The model of bound LF20 shows the sequence VYPYPMEPT (residues 8 to 16 of the 20-residue long LF20). This was the ordered region, and the electron density was clearly visible in difference maps ($2F_o$-F_c and F_o-F_c) calculated from crystal X-ray diffraction data. All three molecules are shown bound in the substrate-binding groove of LF, using the surface calculated for LF-LF20. The targets are all bound in the same N to C peptide orientation. (B) An overview of LF bound to the targets LF20, GM6001, and SHAc-YPM, superposed and colored as in (A). The molecular surface was calculated from the LF-LF20 complex. The domains in LF are labeled I, II, III and IV. The catalytic site is in Domain IV, where the zinc atom (not seen in this figure) is bound. These figures were prepared using SPOCK[14].

conformational change in the PA_{63} heptamer, causing it to form a channel that punctures through the membrane of the endosome and facilitates the translocation of the active enzymes into the host cell cytosol (interior). LeTx is specifically cytotoxic towards the host's white blood cells called macrophages, with LF targeting five of seven members of the Mitogen-Activated Protein Kinase Kinase (MAPKK) family[5], a protein family involved in the cell signaling essential for normal cell growth and differentiation. Proteolysis of the MAPKKs at their N-termini by LF rapidly blocks signals to recruit other immune cells to fight the bacterial infection[6]. EF, which is a calmodulin-activated adenylyl cyclase, can increase the concentration of adenosine 3',5'-cyclic monophosphate (cyclic AMP or cAMP), a secondary messenger that regulates cell function inside the host cell, to abnormal levels[7]. Accumulation of fluids within and between cells occurs due to this abnormality, leading to edema. In advanced anthrax intoxication, especially in highly fatal cases of inhalation anthrax where LF is the greatest source of damage, an antitoxin such as an inhibitor against LF, would be critical as part of a combined therapy with other existing treatments (eg. antibiotics, anti-PA antibodies, critical care).

To date, the crystal structures of all three proteins, PA[4], LF[8] and EF[9] that make up the Anthrax Toxin components have been solved. Previously, we had reported the crystal structure of LF, and LF bound to the N-terminus fragment of its natural substrate, MAPKK2[8] (also see SSRL Science Highlights, April 2002). In collaboration with researchers at Harvard Medical School and United States Army Medical Research Institute of Infectious Diseases (USAMRIID), we have now obtained crystal structures of LF bound to an optimized substrate peptide and also of LF bound to several small molecule inhibitors[10, 11]. Data collection from the protein crystals was done at the Stanford Synchrotron Radiation Laboratory (SSRL), Menlo Park, California, on the macromolecular crystallography Beamlines 9-1 and 7-1, and also at the National Synchrotron Light Source (NSLS), Brookhaven, New York.

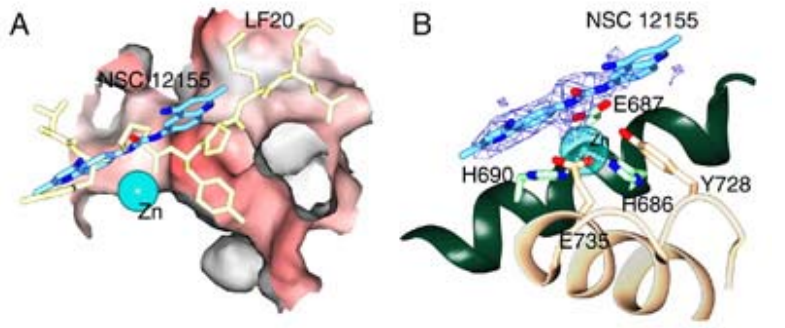

Figure 2. X-ray crystal structure of the LF-NSC 12155-Zn complex.

(A) Molecular surface of LF colored by charge (red = negative; blue = positive), with Zn^{2+} as a solid sphere in cyan. The model of the inhibitor molecule NSC 12155 (carbons are in light blue), and an optimized peptide with the sequence VYPYPME in light yellow, both shown in stick representation. The inhibitor NSC 12155 binds to LF in the active site at the hydrophobic neck of the pocket, where Tyrosine 4 of sequence VYPYPME (equivalent of Tyrosine 11 of LF20) is highly selected as the optimized residue to bind in this position. The binding position of NSC 12155 overlaps that of the optimized peptide sequence at residues 2 and 3 (YP). This figure was prepared using SPOCK[14].

(B) The inhibitor NSC 12155 bound in the active site of LF. The difference map, $2F_o$-F_c, calculated at 2.9-Å resolution limit, is contoured at 1.0 σ. A portion of NSC 12155 appears non-rigid owing to a rotatable bond, with patchy electron density trace at 1.0 σ level, although almost full electron density coverage is seen for this portion at a contour level of 0.6 σ. Inhibitor molecule (light blue), zinc-coordinating residues (H686, H690, E735) and catalytic residues (E687, Y728) are in stick representation. The Cα atoms of residues 680-694 (green, background) and 726-742 (beige, foreground) are in ribbon representation. The Zn^{2+} ion (cyan) is a lined sphere, and its hydrogen bonds with His686, His690 and Glu735 are represented as aligned small white spheres. These figures were prepared using SPOCK[14] (http://mackerel.tamu.edu/spock/).

In the work with colleagues at Harvard Medical School, the structural basis for LF target recognition was investigated initially through finding an optimized substrate for LF[10]. Two sets of peptide libraries with randomized amino acid sequence in 13 residue-long peptides were generated, with either the sequence N-terminal or C-terminal to the peptide cleavage site randomized. Each residue that contributed to increased substrate catalysis rate (increased turnover) at the individual specific position on the peptide substrate was scored. A consensus peptide sequence was built upon these results and we designed an optimized peptide substrate for further experimental work. A 20-residue long peptide was synthesized, using the consensus 13-residue sequence identified, flanked by the amino acid sequence found in the natural substrate for LF, MAPKK2. This peptide was named LF20 (MLARRKKVYPYPMEPTIAEG). LF20 provided a probe to illustrate the shape, extent and limits of the substrate-binding and catalytic site of LF. Based on the consensus sequence, a small peptidic inhibitor named SHAc-YPM, consisting of a zinc-binding thioacetyl moiety coupled to the three residues C-terminal to the cleavage site in LF20, was also created (K_i = 11 ± 3 μM). Looking at other known metalloprotease inhibitors with similar peptidic sequences to SHAc-YPM, GM6001[12, 13], a commercially available potent inhibitor of matrix metalloproteases (MMPs, a family of cell surface proteins in mammalian cells implicated in cancer tumour growth), MMP1 and MMP2, was assayed for and showed significant inhibition towards LF (K_i = 2.1 ± 0.2 μM).

Simultaneously, in collaboration with colleagues at USAMRIID, who performed a search for an effective LF inhibitor in a high-throughput format, using the National Cancer Institute (NCI) Diversity Set (http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html), several candidate small molecule compounds were identified[11] and set up for LF-inhibitor complex crystals. One LF-inhibitor complex that gave the best crystal data was that of LF bound to the hit, NSC 12155, a quinolone compound that gave the highest inhibition constant in this screen (K_i = 0.5 ± 0.18 μM).

Crystal structures of LF-LF20, in the absence and presence of zinc[10], and LF-SHAc-YPM[10], LF-GM6001[10] and LF-NSC12155[11], all in the presence of zinc, were obtained to explore the three-dimensional relationship between LF and its binding targets. The striking features that were evident from the crystal structures were that all of the peptides and small molecule inhibitors recognized and bound to LF via the hydrophobic substrate-binding groove in the immediate vicinity of the catalytic zinc. LF20 induced a conformational change in the active site of LF, and Tyr11, when bound in the S1' hydrophobic pocket in LF causes the pocket to expand, in a phenomena termed "induced fit". This would not have been known if we did not have the consensus peptide LF20 to probe the binding site of LF, and thus presents a novel look at the elasticity of LF in its active site. The small molecule inhibitors on the other hand, provided us with an indication of the criteria for inhibitor design, which guides us to conclude that the main determinants for most effective inhibition is through zinc-deprivation by a strong metal-chelating moiety, like the hydroxamate group, coupled to a hydrophobic and stereochemically well-fitted small molecule structure. The structure of LF-NSC12155 with[11] and without zinc (unpublished data), indicates that it may even be sufficient to have a compound with hydrophobic interactions alone (without metal chelation) to effectively inhibit the catalytic activity of LF.

Due to the diffraction limits and large unit cell dimensions of the protein crystals possibly grown for LF, data collection at a synchrotron facility like SSRL greatly enhances our ability to obtain data of high quality and resolution to achieve the results for our studies. Crystal structures will be highly important to guide future drug-designing efforts for LF inhibitors. The continuing work on the inhibitors for LF will be towards generating chemical derivatives with better inhibitory capabilities and cell permeability for effective delivery of the potential drug molecule to the centre of LF activity in the event of anthrax intoxication. We are at

present working on this aspect of our research, and will be frequent users of the SSRL for its synchrotron radiation source.

References:

1. Leppla SH in *Comprehensive Sourcebook of Bacterial Protein Toxins* 2nd edn (eds Alouf JA & Freer J) 243-263 (Academic, London, 1999).
2. Hanna PC, Acosta D & Collier RJ. On the role of macrophages in anthrax. *Proc. Natl. Acad. Sci. USA* **90**, 10198-10201 (1993).
3. Bradley KA, Mogridge J, Mourez M, Collier RJ & Young JAT. Identification of the cellular receptor for anthrax toxin. *Nature* **414,** 225-229 (2001).
4. Petosa C, Collier RJ, Klimpel KR, Leppla SH & Liddington RC. Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833-838 (1997).
5. Vitale G, Bernadi L, Napolitani G, Mock M & Montecucco C. Susceptibility of mitogen-activated protein kinase kinase family members to proteolysis by anthrax lethal factor. *Biochem. J.* **352**, 739-745 (2000).
6. Duesbury NS, Webb CP, Leppla SH, Gordon VM, Klimpel KR, Copeland TD, Ahn NG, Oskarsson AK, Fukasawa K, Paull KD & Vande Woude GF. Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science* **280**, 734-737 (1998).
7. Leppla SH. Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc. Natl. Acad. Sci. USA* **79**, 3162-3166 (1982).
8. Pannifer AD, Wong TY, Schwarzenbacher R, Renatus M, Petosa C, Bienkowska J, Lacy DB, Collier RJ, Park S, Leppla SH, Hanna P & Liddington RC. Crystal structure of the anthrax lethal factor. *Nature* **414**, 229-233 (2001).
9. Drum CL, Yan S-Z, Bard J, Shen Y-Q, Lu D, Soelaiman S, Grabarek Z, Bohm A & Tang W-J. Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415**, 396-402 (2002).
10. Turk BE and Wong TY, Schwarzenbacher R, Jarrell ET, Leppla SH, Collier RJ, Liddington RC & Cantley LC. The structural basis for substrate and inhibitor selectivity of the anthrax lethal factor. *Nat. Struct. Mol. Biol.* **11**, 60-66 (2004).
11. Panchal RG, Hermone AR, Nguyen TL, Wong TY, Schwarzenbacher R, Schmidt J, Lane D, McGrath C, Turk BE, Burnett J, Aman MJ, Little S, Sausville EA, Zaharevitz DW, Cantley LC, Liddington RC, Gussio R & Bavari S. Identification of small molecule inhibitors of anthrax lethal factor. *Nat. Struct. Mol. Biol.* **11**, 67-72 (2004).
12. Grobelny D, Poncz L & Galardy RE. Inhibition of human skin fibroblast collagenase, thermolysin, and *Pseudomonas aeruginosa* elastase by peptide hydroxamic acids. *Biochemistry* **31**, 7152-7154 (1992).
13. Levy DE, Lapierre F, Liang W, Ye W, Lange CW, Li X, Grobelny D, Casabonne M, Tyrrell D, Holme K, Nadzan A & Galardy RE. Matrix metalloproteinase inhibitors: a structure-activity study. *J. Med. Chem.* **41**, 199-223 (1998).
14. Christopher JA. *SPOCK: The Structural Properties Observation and Calculation Kit [program manual]* (The Center for Macromolecular Design, Texas A&M University, College Station, TX, 1998).

SSRL is supported by the Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.