BIOCHEMISTRY
including biophysical chemistry & molecular biology
Rapid Report

Biochemistry **2010,** *49,* 10767–10769 10767
DOI: 10.1021/bi1017792

Conformational Dynamics of the Anthrax Lethal Factor Catalytic Center†

Georgios A. Dalkas,‡,∥ Christos T. Chasapis,‡,∥ Petros V. Gkazonis,‡ Detlef Bentrop,§ and Georgios A. Spyroulias*,‡

‡*Department of Pharmacy, University of Patras, GR-26504 Patras, Greece, and* §*Institute of Physiology II, University of Freiburg, D-79108 Freiburg, Germany.* ∥*These authors contributed equally to this work.*

Received October 15, 2010

ABSTRACT: Anthrax lethal factor (LF) is a zinc-metalloprotease that together with the protective antigen constitutes anthrax lethal toxin, which is the most prominent virulence factor of the anthrax disease. The solution nuclear magnetic resonance and in silico conformational dynamics of the 105 C-terminal residues of the LF catalytic core domain in its apo form are described here. The polypeptide adopts a compact structure even in the absence of the Zn^{2+} cofactor, while the 40 N-terminal residues comprising the metal ligands and residues that participate in substrate and inhibitor recognition exhibit more flexibility than the C-terminal region.

Anthrax is an infectious disease of animals and humans caused by the Gram-positive bacterium *Bacillus anthracis*. Cutaneous anthrax is rarely lethal for humans, while inhalation of *B. anthracis* spores is often fatal if the infection is not properly diagnosed and treated with antibiotics during the early stages of infection. After anthrax spores are inhaled, they adhere to alveolar macrophages and then germinate. Bacteria migrate to lymph nodes, where they rapidly multiply and excrete a tripartite exotoxin comprised of protective antigen (PA, 83 kDa), lethal factor (LF, 90 kDa), and calmodulin-activated edema factor adenylyl cyclase (EF, 89 kDa) (*1*). LF is a highly specific Zn-protease that cleaves members of the mitogen-activated protein kinase kinase family near their amino termini, leading to the inhibition of one or more important signaling pathways (*2*). According to LF X-ray models (*3*), the catalytic site and the Zn^{2+} cofactor of the enzyme are found in C-terminal domain IV (residues 552−776). Zn^{2+} is bound to the protein via the consensus sequence of the *gluzincin* subfamily (HExxH-linker-E, where x is any amino acid) (*4*). The 105 C-terminal residues of domain IV ($LF_{672-776}$ hereafter) contain this motif and most of the active site of LF, including key residues for substrate recognition and inhibitor binding. The $LF_{672-776}$ construct is expressed in high yield in *Escherichia coli* as a soluble protein (*5*).

Here, we present the solution nuclear magnetic resonance (NMR) structure and dynamics of $LF_{672-776}$ in its metal-free form. The protein is monomeric in solution, based on gel filtration chromatography and ^{15}N relaxation data. It adopts a compact structure even in the absence of the metal cofactor (Figure 1), exhibiting remarkable similarities with the catalytic site of full-length LF as seen in X-ray structures, both in its Zn(II)-loaded form and in its Zn(II)-free form (*3*). The structure is dominated by six α-helices (4α4−4α9 according to the helix numbering of the crystal structures), with the highly conserved HExxH Zn(II)-binding motif being located in the N-terminal 4α4 helix.

The almost complete assignment of $apoLF_{672-776}$ has been deposited in the BioMagResBank (entry 16735). According to circular dichroism (CD) and NMR data (NOEs, $^3J_{HNH\alpha}$, and chemical shift index), the helical content of $LF_{672-776}$ in solution is in the range of 53−58% (∼64% in X-ray structures). The solution NMR structure of $apoLF_{672-776}$ was determined from 1181 NOEs, 56 angle constraints, and 28 H-bonds (PDB: 2L0R). The compact fold is based on 209 long-range NOE constraints, many of them between side chains of hydrophobic residues in the LF catalytic site (Figure 1A,B). In contrast, only a few NOEs were identified in the Asp701−Phe714 region (consecutive loops L1 and L2), which links helix 4α4 to helix 4α5 (Figure 1B).

The six α-helices of $apoLF_{672-776}$, observed also in LF X-ray models, are arranged with the same topology as in the crystal structure (Figure 1B). The root-mean-square deviation (rmsd) between the solution structure of $apoLF_{672-776}$ and the Zn(II)-LF crystal structure [Protein Data Bank (PDB) entry 1J7N] is 1.26 Å for the backbone and 2.39 Å for all heavy atoms, while a comparison with the apoLF−MEK2 peptide complex (PDB entry 1JKY) (*3*) shows rmsd values of 1.37 and 2.53 Å, respectively (Figure 1B and Table S3 of the Supporting Information). In the solution structure, helices 4α4−4α9 are rather well-defined through helix diagnostic NOEs and backbone dihedral angle constraints (Figure S3 of the Supporting Information), with the exception of helix 4α5 and to a lesser extent helix 4α9. The former follows the highly disordered and flexible Asp701−Asp706 and Asn710−Phe714 segments, while the latter is the C-terminal helix. Helix 4α7 of $apoLF_{672-776}$ is tightly packed among helices 4α4, 4α6, 4α8, and 4α9 and forms an angle of 127° with helix 4α4 to adopt a so-called "X" topology, typical for the conserved two-helix active site structure of *gluzincin* proteases (*4*). These two helices provide the three Zn^{2+}-binding residues, i.e., His686 and His690 in helix 4α4 and Glu735 in helix 4α7. In $apoLF_{672-776}$, these two histidines along with His745 are found to be in the charged state according to $^1H-^{15}N$ heteronuclear single-quantum coherence (HSQC) NMR spectra optimized for the detection of imidazole ring protons [analogous signals of His749 were too weak to be detected possibly because of exchange broadening (Figure S4 of the Supporting Information)]. Thus, the zinc binding site of $apoLF_{672-776}$ has the same net electrostatic charge of +2 as in Zn(II)-LF. This similarity in the electrostatics of the catalytic site could well explain the remarkable structural similarity between the Zn(II)-LF X-ray structures and the apoLF solution structure presented here.

†P.V.G. and G.A.S. acknowledge a UPAT "C. Karatheodori" Research Grant (C.181) and the "Herakleitos II" Research and Education Program.

*To whom correspondence should be addressed. Phone: (+30) 2610-969950. Fax: (+30) 2610-969950. E-mail: G.A.Spyroulias@upatras.gr.

Published on Web 12/01/2010

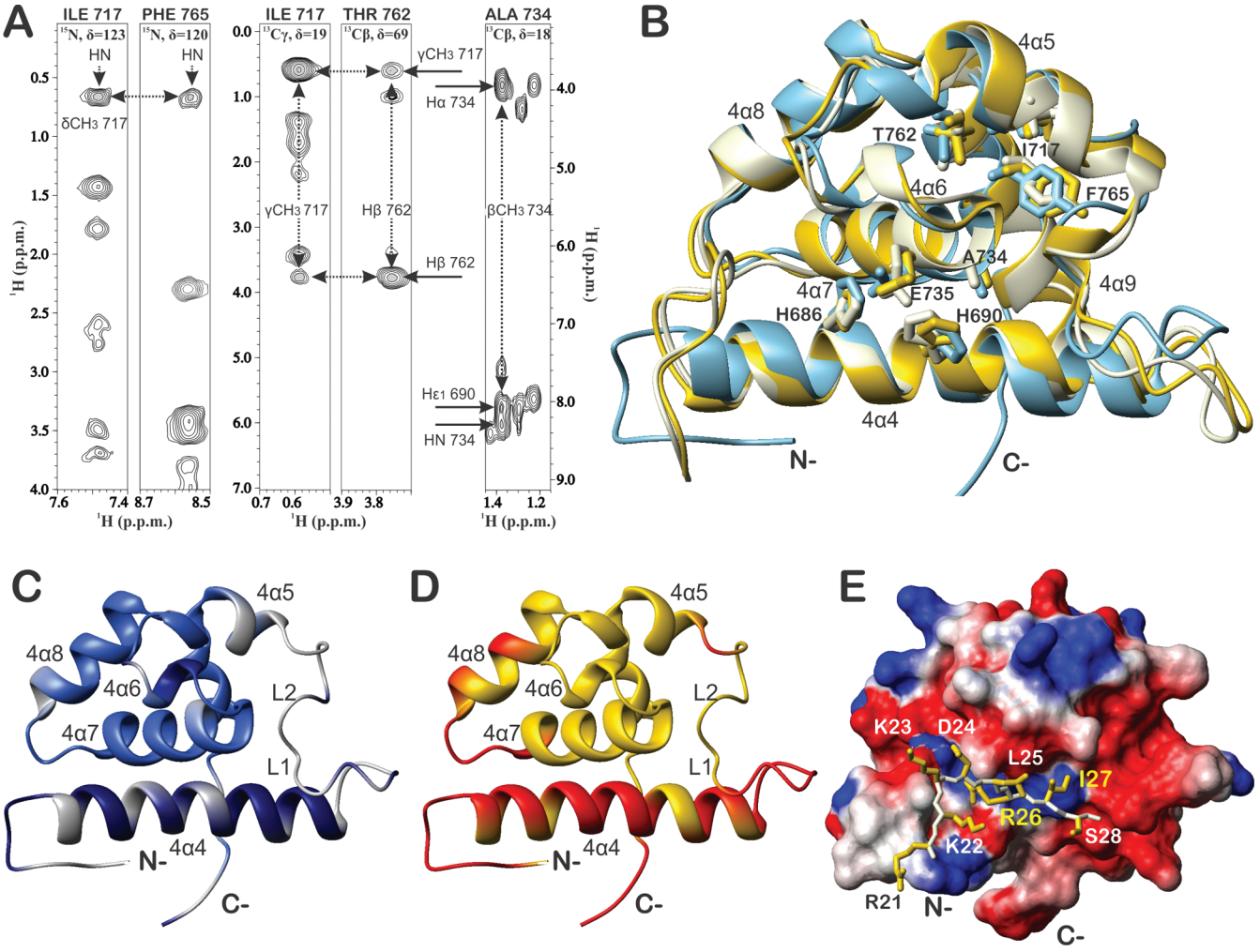

FIGURE 1: (A) Characteristic long-range NOEs defining the hydrophobic core of $LF_{672-776}$. (B) Superimposed LF catalytic core domain structures: X-ray Zn(II)-LF in light gray (PDB entry 1J7N), X-ray apoLF in yellow (PDB entry 1JKY), and NMR apoLF in sky blue (PDB entry 2L0R). (C) Backbone dynamics are illustrated in colors according to S^2 order parameters: royal blue for residues for which $S^2 > 0.6$, dark blue for residues for which $S^2 < 0.6$, and light gray for residues for which S^2 values are not available. (D) Backbone dynamics according to MD. Atomic rms fluctuations of >2.0 Å are colored red. (E) LF−MKK3 complex built through homology modeling using the solution structure of $apoLF_{672-776}$ and the MKK3 LF-bound conformation according to ref *7*. Molecular surface electrostatic potentials of less than $-10kT$ are colored red, potentials of greater than $10kT$ blue, and neutral potentials white (k is the Boltzmann constant). The MKK3 Arg26 (P1) guanidinium group is oriented toward the acidic catalytic site channel and surrounded by glutamates 682, 687, and 735. The carbonyl group of the scissile bond may be oriented toward the positive potential raised by the two charged histidines 686 and 690. This figure was generated with MOLMOL (*8*).

Glu687 in helix 4α4 is the conserved glutamate residue in the HExxH motif of all *gluzincin* proteases (acting as a base to activate the Zn^{2+}-bound water during catalysis). The side chains of these residues are solvent-exposed with the exception of His686, which is oriented toward the catalytic pocket, and exhibit rather similar orientations in the solution structure of $apoLF_{672-776}$ and the crystal structures of Zn(II)-loaded (PDB entry 1J7N) and Zn(II)-free (PDB entry 1JKY) LF. The same holds true for the side chains of Tyr728 and Glu739. These two residues along with Glu687 belong to the secondary shell of Zn^{2+}, and both experimental (*6*) and in silico data (*7*) suggest a significant role of Tyr728 in LF activity. Specifically, computational studies of LF−MKK complexes show that Tyr728 actively participates in substrate harboring through interactions with the carbonyl group of the scissile bond, while in the crystal structure, Tyr728 is hydrogen bonded to a Zn^{2+}-coordinated water molecule. Glu739 is also a key residue of the S1 LF catalytic site subpocket and forms a salt bridge with Arg742 in the majority of the NMR models in a manner similar to that reported for the LF crystal structures. The negative charge density in this site (four positions downstream Zn^{2+}-bound Glu735) is rather conserved in the vast majority of *gluzincins*, and this position is occupied by an aspartate in thermolysin and angiotensin-I converting enzyme. Its role, according to simulated LF−MKK complexes, is to interact efficiently with P1 Arg of LF substrates MKK3b and MKK4 (Figure 1E) determining the orientation of the peptide inside the enzyme's catalytic groove (*7*). All these residues orient their side chains in the NMR models in a manner similar to that in X-ray structures toward the catalytic channel of the enzyme. Additionally, helices 4α4 and 4α7 and the helix 4α4−helix 4α5 linker (including loops L1 and L2) are rich in acidic residues [Glu682, Glu687, Asp693, Asp694, Asp701, Asp706, Glu733, Glu735, and Glu739 (Figure 1E)]. The carboxyl moieties of these residues are exposed toward the catalytic channel, thus contributing to the substrate subpockets (S4−S1) that accommodate the basic residues of MEKs and MKKs.

The backbone dynamics of $apoLF_{672-776}$ were studied through the analysis of ^{15}N R_1 and R_2 relaxation rates and heteronuclear $\{^1H^N\}-^{15}N$ NOEs (Figure S5 of the Supporting Information). The correlation time for isotropic tumbling in solution as

calculated from the R_2/R_1 ratio is 6.8 ± 0.2 ns, clearly indicating that apoLF$_{672-776}$ is in the monomeric state. Model-free analysis of ^{15}N relaxation data as implemented in Tensor2 shows distinct dynamic properties for the N- and C-terminal segments of the polypeptide chain. For the N-terminal residues 675−711, the average S^2 order parameter is 0.26 ± 0.13, whereas S^2 values are much higher for the C-terminal part of apoLF$_{672-776}$ (average value of 0.88 ± 0.17 for residues 717−776) (Figure 1C and Figure S6 of the Supporting Information). Thus, the N-terminal part of the construct is significantly mobile on the subnanosecond time scale, whereas the C-terminal segment appears well-ordered.

The conformational dynamics of apoLF$_{672-776}$ was also probed by a 10 ns unrestrained molecular dynamics (MD) simulation of the average energy-minimized NMR model in explicit water. The MD data on the nanosecond time scale are in good agreement with the ^{15}N relaxation measurements confirming that the N-terminal part of the protein exhibits higher flexibility than the C-terminal region. In particular, the N-terminal helix 4α4 that is a key component of both the Zn^{2+} binding site and the substrate MKK and/or MEK docking site is intrinsically mobile. Although some helix-typical (i, $i+3$,4) H-bonds in helix 4α4 are lost during the MD, it does not collapse during the simulations or lose contact with helix 4α7. This is attributed to several main chain or side chain interactions of helix 4α4 with helix 4α7 and loops L1 and L2. Both His686 and His690 of the Zn^{2+}-binding motif in helix 4α4 interact through H-bonds of their $N^{\delta 1}$ atoms with the side chains of 4α7 Glu735 and 4α6 Thr731, and 4α4 Asp693 and L2 Thr709, respectively. His686 and His690 $N^{\delta 1}$ atoms are found on average 2.70−2.80 Å from carboxyl oxygen atoms $O^{\varepsilon 1/2}$ of Glu735 and $O^{\delta 1/2}$ of Asp693. For more than 60% of the MD time course, Asp693 $O^{\delta 1/2}$ atoms are found at a favorable distance for H-bonding (2.70−2.90 Å) with the $O^{\gamma 1}$ atom of Thr731 in the enzyme's S1′ subpocket and $O^{\gamma 1}$ of Thr709. In agreement with NMR data, these residues are involved in proton−proton interactions detected in ^{13}C- and ^{15}N-edited NOESY spectra and participate in a NOE network between 4α4 His686 and 4α7 residues Ala738, Glu739, and Arg742, surrounding Thr731. In turn, Thr731 protons are found within NOE distance of Glu735. Finally, there are also NOEs between the residues in the 4α4−L2 H-bond network observed during the MD calculation, namely, between 4α4 His690 and S4′ Asp693 and between Asp693 and Leu707 and Val708 preceding L2 Thr709. According to experimental and computational data, Leu707 and Val708 participate in hydrophobic interactions between LF substrates and subpockets S3′−S5′ of the enzyme (*3*, *6*, *7*).

In conclusion, we have described the solution structure and the dynamical behavior of the monomeric polypeptide that represents the C-terminal region of the LF catalytic core in its metal free state. Despite the absence of the metal cofactor, the NMR model is strikingly similar to the holoenzyme and even exhibits a similar charge density in the catalytic channel, as there is experimental evidence that the zinc binding histidines are in the doubly protonated, charged state. NMR relaxation data in concert with MD simulations support an overall compact fold of the enzyme's catalytic site. However, they reveal a higher flexibility of helix 4α4 comprising the first Zn^{2+}-binding motif and the loops in the sequential stretch of Asp701−Ser711. Additional MD calculations on the NMR model reconstituted with a Zn^{2+} ion and a coordinated H_2O molecule indicate a well-defined helix 4α4 in the holoenzyme during the course of the simulation (10 ns). However, the N-terminus still exhibits a higher conformational plasticity with respect to the rest of the protein (Figure S6C of the Supporting Information). Besides its role in binding the metal cofactor, helix 4α4 is crucial for substrate or inhibitor recognition. Despite its flexibility on the picosecond to nanosecond time scale in apoLF, helix 4α4 remains in the spatial proximity of helix 4α7 through H-bonds, and the LF catalytic site retains its intrinsic negative charge density. The negatively charged residues in the LF catalytic channel control recognition and binding of MEKs and MKKs, which are characterized by a significant number of positively charged residues close to the scissile bond (Figure 1E). Metal reconstitution studies are underway, and a metal-loaded LF$_{672-776}$ construct could provide additional atomic-level structural, dynamical, and functional insights into LF−substrate or LF−inhibitor binding and interaction.

ACKNOWLEDGMENT

We thank the Instrumental Analysis Laboratory of the University of Patras (UPAT), Dr. A. Papakyriakou for his assistance in Zn(II)-LF MDs, and Dr. B. Fakler (University of Freiburg) for continued support.

SUPPORTING INFORMATION AVAILABLE

Experimental methods, NMR data and structural statistics (Tables S1 and S2), comparison of LF$_{672-776}$ NMR and X-ray models (Table S3), 30 NMR models (Figure S1), NOE classification and rmsd per residue (Figure S2), NOE sequential connectivity map and $^3J_{HNH\alpha}$ values (Figure S3), $^2J_{HN}$-optimized ^{1}H−^{15}N HSQC spectrum of LF$_{672-776}$ (Figure S4), ^{15}N relaxation data (Figure S5), S^2 order parameters and MD rms fluctuations (Figure S6), and MD frames (Figure S7). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dixon, T. C., Meselson, M., Guillemin, J., and Hanna, P. C. (1999) *N. Engl. J. Med. 341*, 815–826.
2. Duesbery, N. S., Webb, C. P., Leppla, S. H., Gordon, V. M., Klimpel, K. R., Copeland, T. D., Ahn, N. G., Oskarsson, M. K., Fukasawa, K., Paull, K. D., and Vande Woude, G. F. (1998) *Science 280*, 734–737.
3. Pannifer, A. D., Wong, T. Y., Schwarzenbacher, R., Renatus, M., Petosa, C., Bienkowska, J., Lacy, D. B., Collier, R. J., Park, S., Leppla, S. H., Hanna, P., and Liddington, R. C. (2001) *Nature 414*, 229–233.
4. Sturrock, E. D., Natesh, R., van Rooyen, J. M., and Acharya, K. R. (2004) *Cell. Mol. Life Sci. 61*, 2677–2686.
5. Gkazonis, P. V., Dalkas, G. A., Chasapis, C. T., Vlamis-Gardikas, A., Bentrop, D., and Spyroulias, G. A. (2010) *Biochem. Biophys. Res. Commun. 396*, 643–647.
6. Tonello, F., Naletto, L., Romanello, V., Dal Molin, F., and Montecucco, C. (2004) *Biochem. Biophys. Res. Commun. 313*, 496–502.
7. Dalkas, G. A., Papakyriakou, A., Vlamis-Gardikas, A., and Spyroulias, G. A. (2009) *Protein Sci. 18*, 1774–1785.
8. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51–55.