



NMR chemical shift assignment of a constitutively active fragment of the antitermination protein LicT

Yinshan Yang, Nathalie Declerck, H el ene D em en e

► To cite this version:

Yinshan Yang, Nathalie Declerck, H el ene D em en e. NMR chemical shift assignment of a constitutively active fragment of the antitermination protein LicT. *Biomolecular NMR Assignments*, Springer, In press, 10.1007/s12104-019-09914-w . hal-02322846

HAL Id: hal-02322846

<https://hal.archives-ouvertes.fr/hal-02322846>

Submitted on 28 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

NMR chemical shift assignment of a constitutively active fragment of the antitermination protein LicT.

Yinshan Yang¹, Nathalie Declerck^{1,2} and Hélène Déméné¹

¹ Centre de Biochimie Structurale (CBS), INSERM, CNRS, Univ Montpellier, France.

² INRA, département MICA, 78352 Jouy-en-Josas, France.

Corresponding authors:

Hélène Déméné and Nathalie Declerck

Helene.demene@cbs.cnrs.fr

+33 (0) 4 67 41 77 01

Author contributions: ND and HD designed the study and prepared protein samples; YY and HD analyzed the data; YY, ND and HD wrote the paper.

Acknowledgments

The CBS is a member of the French Infrastructure for Integrated Structural Biology (FRISBI), a French infrastructure supported by the National Research Agency (ANR-10-INBS-05) and is a GIS-IBIsA platform.

Abstract

LicT belongs to an essential family of bacterial antitermination proteins which bind to nascent mRNAs in order to stimulate transcription of sugar-metabolizing operons. As most of other antitermination proteins involved in carbohydrate metabolism, LicT is composed of a N-terminal RNA-binding module (CAT) and two homologous regulatory modules (PRD1 and PRD2). The activity of the CAT effector module is controlled by antagonist phosphorylations by the phosphotransferase system (PTS) on conserved histidines of the two C-terminal PRDs in response to available carbon sources. Previous studies on truncated and mutant constructs have provided partial structural insight into the mechanism of signal transduction between the N-terminal RNA-binding domain and the two regulation modules. However, no structure at atomic resolution has been ever solved that contain the RNA-binding domain and a regulation module. We report the NMR assignment of a constitutively active fragment of LicT, named D99A-CAT-PRD1 or CAT-PRD1*. This fragment is composed of the RNA-binding module and the first N-terminal regulation module which bears the mutation of Asp99 to an asparagine. Is dimeric as the native protein, with a 40 kD molecular weight. The D99N mutation is sufficient to endow this fragment with a high RNA-binding constitutive activity, in a phosphorylation-free context. The assignment reported here should set the base of future NMR investigation of signal transduction between the regulatory module and the effector module in the active state of the protein, and in the long term enable the structural study of the full length protein structure in interaction with its target RNA.

Keywords: antitermination activity; BglG-like antiterminator; constitutive activity; RNA-binding.

Abbreviation: CAT, Co-AntiTerminator,; PRD PTS Regulation domain; PTS, PhosphoTransferase System ; RAT, Ribonucleic AntiTerminator.

Biological context.

The transcriptional antiterminator LicT from *Bacillus subtilis* belongs to the BglG family of proteins that regulate gene expression. When activated, they bind to and stabilize a specific RNA hairpin sequence called Ribonucleic AntiTerminator (RAT) located in the 5' untranslated region of their mRNAs target gene. This binding inhibits the formation of a competing terminator hairpin which promotes otherwise the premature transcription termination upstream of the genes of interest. This family of antitermination proteins contains

over fifty members involved in carbohydrate metabolism control in Gram-positive and Gram-negative bacteria. LicT, which is studied here, mediates the induction of the *B.subtilis licS* gene and the *bglPH* operon controlling the utilization of aryl- β -glucosides and β -glucans involved in plant cell wall formation (Krüger and Hecker 1995; Schnetz et al. 1996). As most of BglG-like antitermination proteins, it is composed of three modules: the N-terminal RNA-binding effector module (CAT), and two homologous regulation modules named PRD1 and PRD2 (for PTS Regulation modules). The RNA binding activity of the CAT domain is under the control of phosphorylation on conserved histidines of the PRDs by the PTS. Previous studies have revealed the structural features of isolated fragments of the full-length protein. The isolated CAT module is constitutively active and NMR studies on its complex with the cognate RAT sequence has revealed the structural basis of RNA recognition (Yang et al. 2002). The isolated CAT domain is dimeric and recognizes the two asymmetric internal loops of the antiterminator hairpin in a complex of 2:1 stoichiometry. Crystallographic studies of PRD1-PRD2 fragments, either native or bearing phosphomimetic mutations, showed homodimers in a compact and open configuration respectively (van Tilbeurgh et al. 2001; Graille et al. 2005). Using NMR residual dipolar coupling constants, we had established a crude model of the native CAT-PRD1 fragment, where the CAT module exhibits an open conformation compatible with the decrease of affinity of this fragment for RNA (Ducat et al. 2002; Déméné et al. 2008). However, no structure at atomic resolution of a fragment containing the effector module and a regulation module, have ever been reported for this important protein family, neither in the inactive nor in the active conformation. We report here the assignment of a mutant D99N CAT-PRD1 fragment, known to be constitutively active (Lindner et al. 2002). This fragment is 169 amino acid long (hexa histidine tag excluded) but analysis of ultracentrifugation and size-exclusion chromatography data showed that it behaves as a dimer in solution (Déméné et al. 2008). 98% of backbone NH amide pairs have been assigned. This work should pave the way to the structure elucidation of this protein and its interaction with target RNAs.

Methods and experiments

Protein production and purification

The CAT-PRD1* protein fragment was produced from a pET15 derivative plasmid carrying the *licT* (1-167) gene fragment from *B. subtilis* LicT (Uniprot P39805) encoding the 167 N-terminal residues corresponding to the CAT RNA binding domain (residues 1-52), the first regulatory module (PRD1, residues 69-167) and the linker region in between. The protein was produced with a C-terminal His-Tag (LQHSHHHH) for

purification need. Introduction of the activating mutation substituting an Asn residue for Asp99, was performed using the QuikChange kit (Agilent) and checked by DNA sequencing and mass spectrometry. CAT-PRD1* was then expressed in *E. coli* BL21 (DE3) cells using M9 minimal medium (1 L) with a phosphate buffer concentration of 33 mM supplemented with 1 g $^{15}\text{NH}_4\text{Cl}$, 3 g $^{-1}$ ^{13}C -D-glucose, 1 mg thiamine and biotin, 5 mM MgSO_4 , 0.2 mM CaCl_2 and metal traces at 37°C. Expression was induced by addition of 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG) when an OD_{600} 0.6 was reached.

Cells were harvested after overnight growth at 20°C, by centrifugation at 6,000 g for 30 min (induction length: 16 hours). Cells were resuspended in 10 mM TRIS buffer (pH 8), 200 mM NaCl, 0.1 mg.mL $^{-1}$ lysozyme, DNase (10 $\mu\text{g.mL}^{-1}$). Benzamidine (1 mM.L $^{-1}$) was added as protease inhibitor before lysis by sonication. The cell lysate was centrifuged at 35 000 G for 30 min at 4 °C. The soluble fraction was isolated and was supplemented with imidazole to a final concentration of 20 mM. CAT-PRD1* was purified using a Ni-NTA resin column of 5 mL. Subsequently, it was washed 2 times with 50 mL of 10 mM TRIS buffer (pH 8), 200 mM NaCl, 1 mM Benzamidine, 1 mM Beta mercaptoethanol (buffer A), interspersed with a 5 mL wash with 50 mM Tris pH 8, 500 mM NaPO_4 . The protein was eluted with an imidazole gradient in buffer A. Selected fractions of high purity and protein content were then concentrated in a 10 kDa centrifugal filter unit and loaded into a Superdex 75 Hiload 26/60 column (GE Healthcare) equilibrated with buffer B (15 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM DTT, 0.5 mM EDA and 0.2 mM benzamidine). Fractions containing high content of D99N CAT-PRD1 at high purity were pooled and concentrated to a concentration of 6 mg/mL. The purified proteins were vacuum concentrated and kept at -80 °C until use. With this protocol, we obtained a 40 mg yield of protein for 1 culture liter.

NMR spectra

The sample concentration were either 0.8 mM protein in 10 mM Tris, 200 mM NaCl, pH 8.0, 10% D_2O , or 0.6 mM protein , in 20 mM phosphate, 200 mM NaCl, pH 6.4, 10% D_2O . Both samples were supplemented with 0.5 mM EDTA and 0.2 mM benzamidine to avoid degradation by proteases.

NMR experiments were performed at 308 K and 303K on a Bruker Avance III spectrometer (Bruker, Rheinstetten, Germany) operating at ^1H frequency of 800.23 MHz, using a 5 mm cryogenic H/ C/N/D probe with Z-axis gradient. Proton chemical shifts were referenced with respect to DSS and the frequency ratios of

$^{15}\text{N}/^1\text{H} = 0.101329118$ and $^{13}\text{C}/^1\text{H} = 0.251449530$ (Wishart et al. 1995). For sequential backbone assignment, a ^{15}N -edited 3D NOESY-HSQC (100ms mixing time) and a 3D TOCSY-HSQC (50 ms mixing time) experiment were recorded on the sample at pH 8 and at 308K and provided initial assignments of most spin systems, which were further confirmed by 3D-HNCA and 3D-HNCOCA. For completion of chemical shifts and further verification, a HNC(O) and a CBCA(CO)NH experiment were recorded as well as a (H)CCH-TOCSY and 2D ^1H - ^{13}C HSQC at the same temperature in the same solvent, but in also D_2O (otherwise same composition). ^{15}N -edited 3D NOESY-HSQC (tm 100ms) and 3D TOCSY-HSQC (tm 50 ms) experiment were repeated at 303K on the sample at pH 6.4 for detection of missing NH resonances. An additional ^1H - ^1H 2D NOESY was recorded in D_2O . Gifa was used for NMR data processing (Pons et al. 1996).

NMR assignment and deposition

The spectrum of ^{15}N -labelled D99N LicT-CAT-PRD1 is shown in Fig. 1. It contains 175 amino acids, including the C-terminal hexahistidine tag. It is a symmetrical dimeric protein of 40 kDa molecular weight, and one set of resonances was observed. The protein revealed to be impossible to deuterated in flask conditions. Hence, we had to work at high temperature, and most experiments were carried out at 308K, on the sample whose spectrum was the best looking, i.e. at pH 8. A set of experiments at recorded pH 6.4 and 303 K provided the few missing resonances. The assignment of residues belonging to the CAT domain (residues 1-53) was facilitated because it resembles to its isolated form (Yang et al. 2002), as this module conserves its structure when the protein has a high affinity to RNA (Déméné et al. 2008). The assignment of the resonances of the regulatory module (70-169) was ascertained by the assignment of sequential NOEs typical of helix ($\text{HN}_i\text{-HN}_{i+1}$ and $\text{H}\alpha_i/\text{HN}_{i+3}$) corresponding to the fold of this region as a helix bundle (van Tilbeurgh et al. 2001). D99-N LicT-CAT-PRD1 contains 169 amino acids, including 2 prolines (hexahistidine tag excluded). Of the 167 expected NH resonances, 163 have been assigned, corresponding to a 97.6% assignment ratio. Of these assignments, N9, E17, K53, D91, E150, E167 and Q169 ^1H - ^{15}N correlation peaks were only observed at pH 6.4. Backbone resonances that could not be definitely assigned included the NH, ^{15}N and ^{13}C (carbonyl) resonances of M1, S92, I93 and V94, the $\text{C}\alpha$ resonance of N166 and Q169 and the ^{13}C resonance of I66, N90, D91, L148, and Q169. 30% (55 out of 160) $\text{C}\beta$ resonances could be identified

from HNCACB or CBCA(CO)NH experiments. All methyl carbon and protons were assigned, as were the aromatic carbon and protons, W120 excepted for which we could assign only aromatic protons from the ^1H - ^1H NOESY spectrum in D_2O . ^1H , ^{15}N , and ^{13}C backbone chemical shifts were used as input for the TALOS-N software to predict secondary structures and backbone internal mobility (Shen and Bax 2013). According to TALOS-N, the CAT and PRD1 domains possesses 4 beta strands and 5 helices respectively, in accordance with the crystal structures of the individual fragments (Fig. 2) (van Tilbeurgh et al. 2001; Yang et al. 2002; Graille et al. 2005). The Random Coil Index derived S^2 parameter (RCI- S^2) suggests that the linker region between CAT and PRD1 has a higher level of disorder than the CAT and PRD1 modules (Fig.2), and that the latter domains may to some extent tumble independently.

This work represents the first report of chemical shifts of a fragment of a BlgG-like antitermination protein containing simultaneously a RNA-binding effector module and regulatory module in an activated state. It should enable the elucidation of the structure of this fragment, and later of the full length protein under its active form, isolated or in interaction with its cognate RNA target.

The chemical shift assignments have been deposited in the BioMagResBank (<http://www.brmb.wisc.edu>) with the 27893 accession number.

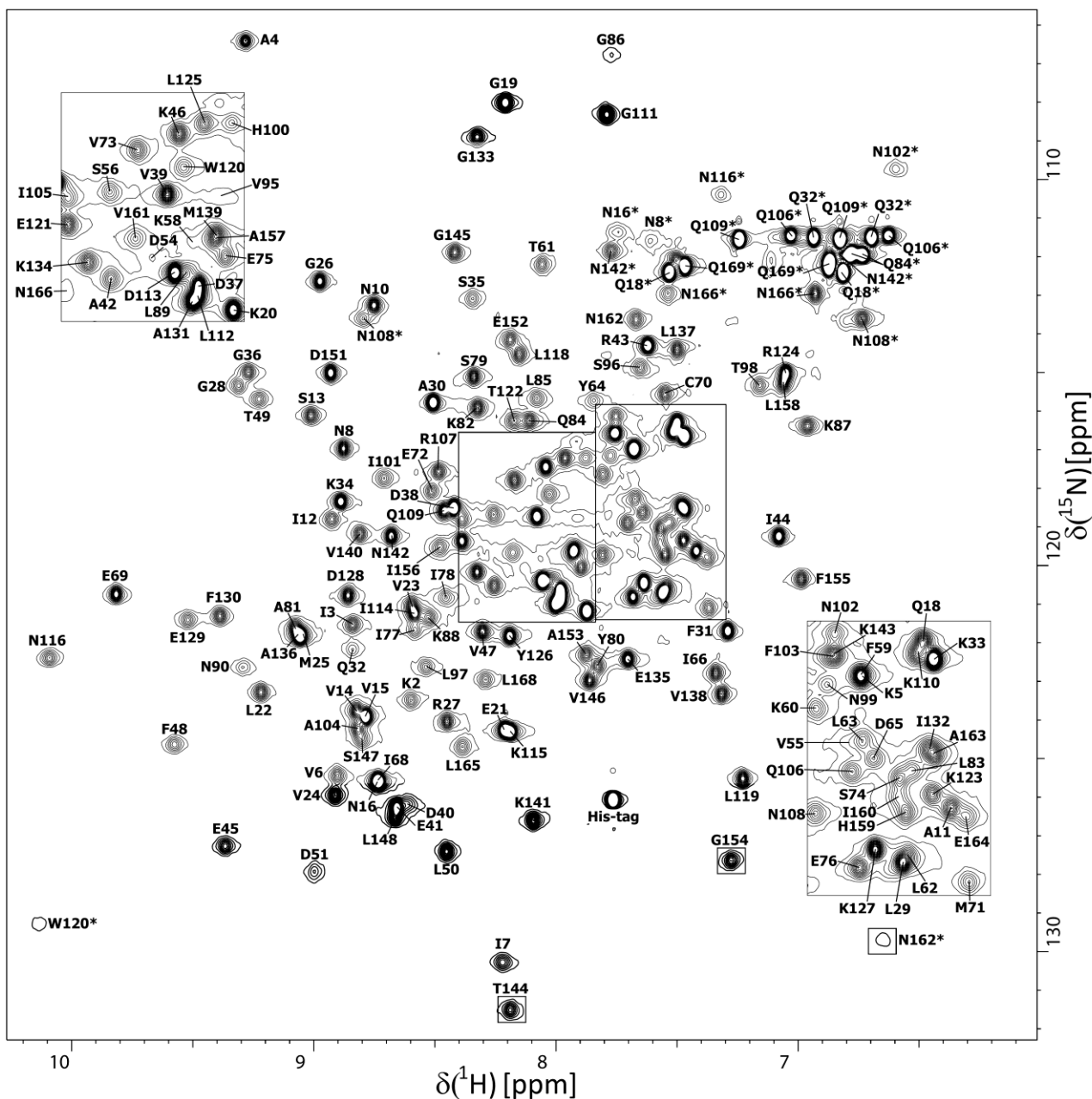


Fig. 1: 2D ^1H - ^{15}N HSQC spectrum of a 800 μM sample of ^{15}N -labeled D99 LicT-CAT-PRD1 in 10 mM Tris pH 8, 200 mM NaCl, at 308K. Each backbone amide resonance is labeled with the amino acid type (one-letter code) and the number in the sequence. Aliased correlation peaks are framed. $\text{H}\epsilon 1$ - $\text{N}\epsilon 1$ amino groups from tryptophanes are labeled with an asterisk (*). Also labelled with an asterisk are the assigned amide groups from glutamine and asparagine side chains, which are in addition outlined with horizontal lines.

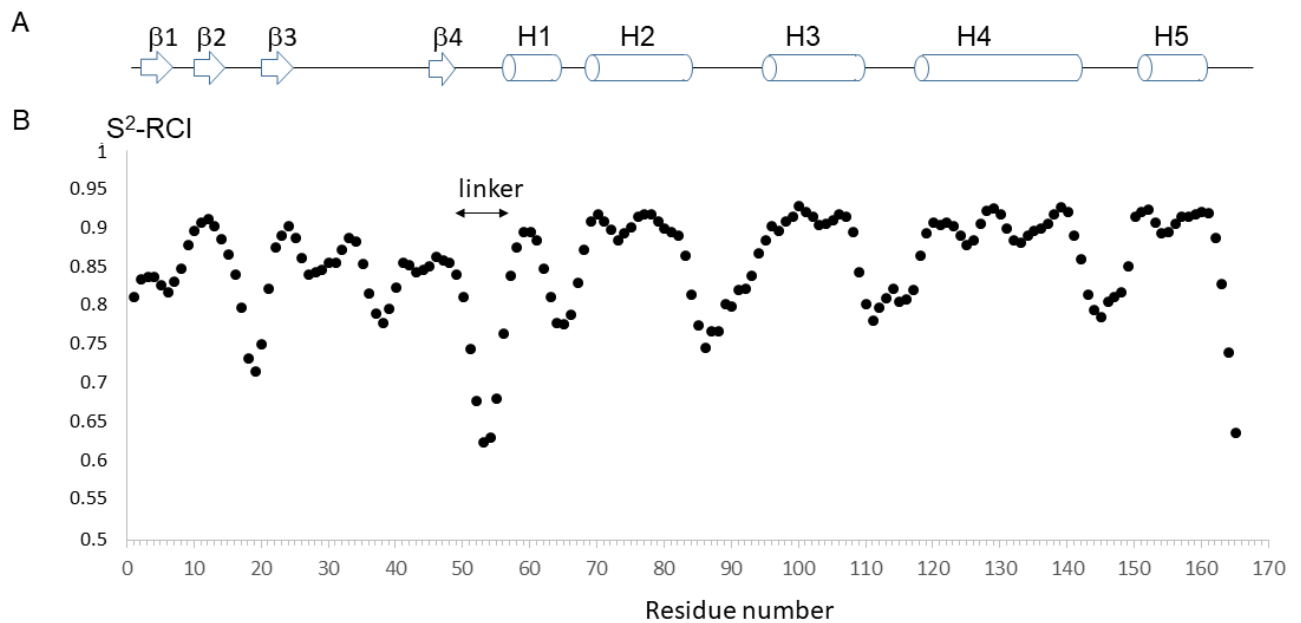


Fig. 2: Prediction of secondary structures (A) and RCI-S² order parameter (B) for D99 LicT-CAT-PRD1 by the Talos-N software (Shen and Bax 2013).

References

- Déméné H, Ducat T, De Guillen K, et al (2008) Structural Mechanism of Signal Transduction between the RNA-binding Domain and the Phosphotransferase System Regulation Domain of the LicT Antiterminator. *J Biol Chem* 283:30838–30849. doi: 10.1074/jbc.M805955200
- Ducat T, Declerck N, Kochoyan M, Déméné H (2002) 1H, 15N and 13C backbone resonance assignments of the 40 kDa LicT-CAT-PRD1 protein. *J Biomol NMR* 23:325–6
- Graille M, Zhou C-Z, Receveur-Bréchet V, et al (2005) Activation of the LicT Transcriptional Antiterminator Involves a Domain Swing/Lock Mechanism Provoking Massive Structural Changes. *J Biol Chem* 280:14780–14789. doi: 10.1074/jbc.m414642200
- Krüger S, Hecker M (1995) Regulation of the putative bglPH operon for aryl-beta-glucoside utilization in *Bacillus subtilis*. *J Bacteriol* 177:5590–7
- Lindner C, Hecker M, Le Coq D, Deutscher J (2002) *Bacillus subtilis* mutant LicT antiterminators exhibiting enzyme I- and HPr-independent antitermination affect catabolite repression of the bglPH operon. *J Bacteriol* 184:4819–4828. doi: 10.1128/JB.184.17.4819-4828.2002

- Pons J-L, Malliavin T, Delsuc M (1996) Gifa V. 4: A complete package for NMR data set processing. *J Biomol NMR* 8:. doi: 10.1007/BF00228146
- Schnetz K, Stülke J, Gertz S, et al (1996) LicT, a *Bacillus subtilis* transcriptional antiterminator protein of the BglG family. *J Bacteriol* 178:1971–9
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241. doi: 10.1007/s10858-013-9741-y
- van Tilbeurgh H, Le Coq D, Declerck N (2001) Crystal structure of an activated form of the PTS regulation domain from the LicT transcriptional antiterminator. *EMBO J* 20:3789–3799. doi: 10.1093/emboj/20.14.3789
- Yang Y, Declerck N, Manival X, et al (2002) Solution structure of the LicT-RNA antitermination complex: CAT clamping RAT. *EMBO J* 21:1987–1997. doi: 10.1093/emboj/21.8.1987