1H, 13C and 15N resonance assignments of σS activating protein Crl from Salmonella enterica serovar Typhimurium

Paola Cavaliere, Françoise Norel, Christina Sizun

^{1}H, ^{13}C and ^{15}N resonance assignments of σ^{S} activating protein Crl from *Salmonella enterica* serovar Typhimurium

Paola Cavaliere [1,2]

Françoise Norel [1,2]

Christina Sizun [3,*]

Phone +33 1 69 82 37 64
Email christina.sizun@cnrs.fr

[1] Département de Microbiologie, Laboratoire Systèmes Macromoléculaires et Signalisation, Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France

[2] CNRS ERL3526, rue du Docteur Roux, 75015 Paris, France

[3] Institut de Chimie des Substances Naturelles, CNRS UPR2301, 91190 Gif-sur-Yvette, France

Abstract

The general stress response in Enterobacteria, like *Escherichia coli* or *Salmonella*, is controlled by the transcription factor σ^{S}, encoded by the *rpoS* gene, which accumulates during stationary phase growth and associates with the core RNA polymerase enzyme (E) to promote transcription of genes involved in cell survival. Tight regulation of σ^{S} is essential to preserve the balance between self-preservation under stress conditions and nutritional competence in the absence of stress. Whereas σ factors are generally inactivated upon interaction with anti-sigma proteins, σ^{S} binding by the Crl protein facilitates the formation of the holoenzyme $E\sigma^{S}$, and therefore σ^{S}-controlled transcription. Previously, critical residues in both Crl and σ^{S} were identified and assigned to the binding interface in the Crl–σ^{S} complex. However, high-resolution structural data are missing to fully understand the molecular mechanisms underlying σ^{S} activation by Crl,

in particular the possible role of Crl in triggering domain rearrangements in the multi-domain protein σ^S. Here we provide the ^1H, ^{13}C and ^{15}N resonance assignments of *Salmonella enterica* serovar Typhimurium Crl, as a starting point for Crl_{STM} structure determination and further structural investigation of the Crl_{STM}–σ_{STM}^S complex.

AQ1

AQ2

Keywords

Salmonella Typhimurium
Crl
RpoS
Sigma factor
RNA polymerase
Stress response

Biological context

Bacteria adapt to environmental perturbations like extreme stress conditions by entering into a stationary growth phase, altering their gene expression pattern, cell morphology and physiology. In enterobacteria like *Escherichia coli* and *Salmonella*, protection against multiple stress is controlled at the molecular level by the master transcriptional regulator during stationary phase, σ^S, encoded by the *rpoS* gene (Battesti et al. 2011; Hengge 2010). During exponential growth, the housekeeping σ factor σ^{70} associates with the multi-subunit core enzyme E ($\alpha_2\beta\beta'\omega$) of the RNA polymerase to form the transcription initiation competent holoenzyme $E\sigma^S$. However, in the stationary growth phase, σ^S is expressed and competes with σ^{70}. It then down-regulates σ^{70}-controlled genes involved in nutritional competence and promotes transcription of more than 10 % of the bacterial genome that is essential for cell survival under starvation and multiple stress conditions. In *Salmonella enterica* serovar Typhimurium, which causes typhoid fever symptoms in mice and gastroenteritis in humans, σ^S is moreover involved in virulence (Hengge 2010).

To balance self-preservation and nutritional competence according to environmental cues, σ^S-dependent transcription is tightly regulated at multiple levels: transcription, translation, protein stability and activation (Battesti et

al. 2011; Hengge 2010). In *E. coli* and *Salmonella*, the Crl protein activates σ^S-dependent transcription and plays an important role in the biosynthesis of curli fibers, involved in adhesion to extracellular matrices and biofilm formation (Hengge 2010). Bacterial genomes containing the *crl* gene also contain the *rpoS* gene, underlining the close interplay between Crl and σ^S (Monteil et al. 2010a). The transcription factor Crl binds directly to σ^S, but not to DNA, and stimulates formation of the $E\sigma^S$ holoenzyme, increasing the competitiveness of σ^S, which is otherwise the σ factor with the weakest in vitro affinity for E (Hengge 2010).

Whereas the functions of Crl have been clearly established, the molecular mechanisms underlying the activation of σ^S by Crl are still poorly understood. Like all σ factors of the σ^{70} family, σ^S is multidomain protein (Hengge 2010). In contrast to housekeeping σ, high-resolution three dimensional structures are not available for σ^S. It might be speculated that Crl triggers σ^S conformational rearrangements that favor efficient binding to E. It has been shown that the conserved domain 2 of σ^S is the only domain that binds Crl, displaying the same Crl binding properties as full-length σ^S (Monteil et al. 2010b; Banta et al. 2013; Cavaliere et al. 2014). In Crl, the conserved Arg51 was found to be essential for σ^S activation in *E. coli* (Banta et al. 2014), but in *Salmonella* additional critical aromatic and charged residues were identified (Cavaliere et al. 2014). In both cases, the σ^S binding interfaces on Crl were analyzed in light of X-ray structures of a Crl homolog in *Proteus mirabilis*, Crl_{PM} (Cavaliere et al. 2014; Banta et al. 2014). Crl is much less conserved than σ^S, and sequence identity is less than 50 % between Crl_{PM} and each *E. coli* Crl_{EC} and *S. enterica* serovar Typhimurium Crl_{STM} (Monteil et al. 2010a). Neither Crl_{EC} nor Crl_{STM} could be crystallized, but the small size of Crl made it amenable to NMR spectroscopy. Here we provide the nearly complete ^1H, ^{13}C and ^{15}N resonance assignments of Crl_{STM}, as a prerequisite for structure determination of Crl_{STM} and for further structural characterization of the Crl_{STM}–σ^S_{STM} complex.

Methods and experiments

Sample preparation

The pVFC681 (pET-MCN-EAVNH) plasmid for overexpression of *S. enterica* serovar Typhimurium Crl in *E. coli* BL21(DE3) strain was described in (Cavaliere et al. 2014). Crl_{STM} flanked by a 21-residue N-terminal His-tag was produced with ^{15}N-, ^{13}C ^{15}N- and ^{13}C ^{15}N-80 % ^2H-labeling. 1 L cultures

in minimal M9 medium, supplemented with 100 µg mL^{-1} ampicillin, 1 g L^{-1} $^{15}NH_4Cl$ (Eurisotop) and 3 g L^{-1} unlabeled or ^{13}C-labeled glucose (Cortecnet) and inoculated with 10 mL of saturated starter culture in Luria–Bertani broth, were grown at 37 °C until OD_{600} = 0.7. Induction was started with 1 mM IPTG at 30 °C for 4 h. For ^{13}C-^{15}N-80 % ^{2}H-labeling, cells grown in 1 L of unlabeled M9 medium were collected by centrifugation at OD_{600} = 0.7. They were resuspended in 100 mL of triple labeled M9 medium, prepared in 99.8 % deuterium oxide (Eurisotop) and supplemented with 1 g L^{-1} $^{15}NH_4Cl$ (Eurisotop) and 2 g L^{-1} ^{13}C-labeled glucose (Cortecnet), and incubated for 1 h at 37 °C. Cells were collected again, resuspended in 900 mL triple labeled M9 medium and incubated for 30 min. Protein expression was induced with IPTG at 28 °C for 16 h. Cell pellets were harvested by centrifugation and lyzed at 4 °C in a Cell disrupter (Constant System Ltd.) in buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl,10 mM imidazole, supplemented with anti-proteases and benzonase). After clarification by ultracentrifugation, the supernatant was loaded on to a Ni–NTA column (Protino Ni–NTA, 5 ml) and eluted with buffer B (buffer A added with 300 mM imidazole). After dialysis against buffer C (50 mM sodium phosphate pH 8.0, 300 mM NaCl), the protein was finally purified by size exclusion chromatography on a HiLoadTM 16/60 Superdex 75TM column (GE Healthcare) in buffer C. To accelerate back-protonation of amides, ^{13}C ^{15}N-80 % ^{2}H-labeled Crl was treated with 8 M urea for 12 h at 4 °C, then dialyzed into 2 M urea and finally into buffer C. Samples for NMR experiments with aliphatic and aromatic ^{1}H detection were exchanged on ~~m~~Micro Bio-Spin 6 columns (Biorad) into buffer C in 100 % deuterium oxide. All other samples contained 7 % D_2O. 2 mM dithiothreitol was added to all samples. Protein concentrations were 300, 250 and 350 µM for ^{15}N–Crl, ^{13}C ^{15}N–Crl and ^{13}C ^{15}N-80 % ^{2}H–Crl samples, respectively.

NMR experiments

NMR measurements were carried out on a Bruker Avance III spectrometer equipped with a cryogenic TCI probe at a magnetic field of 18.8 T and a temperature of 293 K. Assignment of Crl_{STM} backbone resonance frequencies was achieved by analyzing standard triple resonance experiments recorded on ^{13}C ^{15}N–Crl_{STM} (3D HNCA, 3D HN(CO)CA, 3D HNCO, 3D CBCA(CO)NH) and ^{13}C ^{15}N-80 % ^{2}H–Crl_{STM} (3D HNCA, 3D HN(CO)CA, 3D HNCO, 3D HN(CA)CO). To retrieve signals in solvent exposed regions like the N-terminal His-tag, a 3D HNCACB experiment was recorded on a ^{15}N ^{13}C–Crl_{STM} sample at pH 6.5. $^{1}H_{\alpha}$ assignments were obtained from a 3D ^{1}H–^{15}N

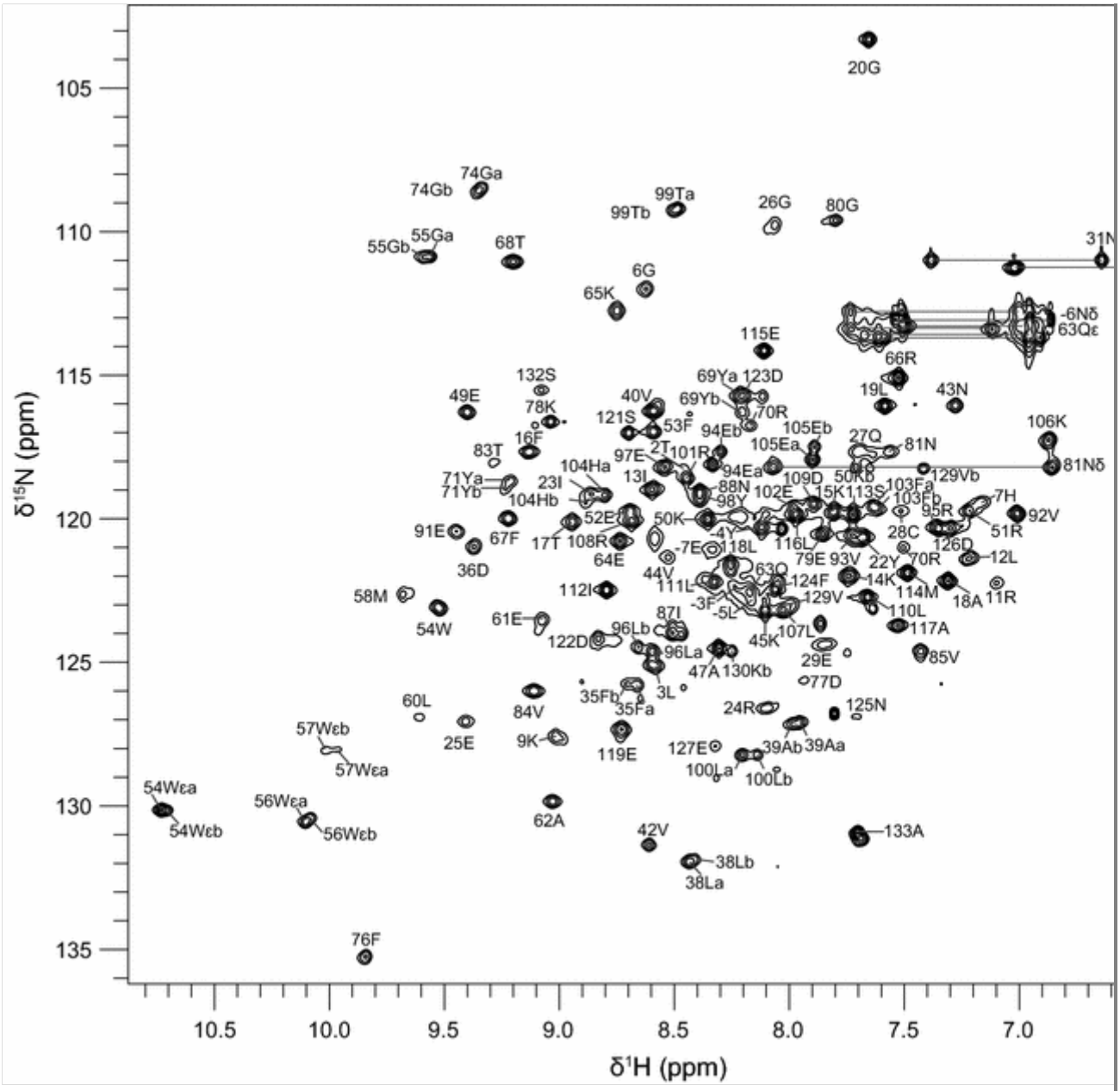
NOESY-HSQC, all other proton side chain assignments from 3D hCCH-TOCSY (15.6 ms spin lock) and 2D ^{1}H–^{1}H NOESY spectra. The NOESY mixing time was 80 ms. ^{1}H chemical shifts were referenced to DSS. Spectra were processed with Topspin 3.1 (Bruker Biospin) or NMRPipe (Delaglio et al. 1995) and analyzed with CCPNMR 2.2 software (Vranken et al. 2005).

Assignment and data deposition

The ^{15}N–HSQC spectrum and amide assignments of Crl_{STM} at pH 8.0 and 293 K are shown in Fig. 1. Assignment completeness for Crl_{STM} was 82.0 % for amide resonances, 90.9 % for backbone resonances ($^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$, $^{1}H_{\alpha}$) and 91.1 % for side chain protons. Taking the His-tag into account, completeness amounted to 81.0, 83.2 and 82.6 %. Exchangeable arginine and lysine side chain protons were not assigned. Assignments are completely missing for Asp30 and Glu89.

Fig. 1

^{1}H–^{15}N HSQC spectrum at 800 MHz of 0.3 mM ^{15}N-labeled *Salmonella enterica* serovar Typhimurium Crl at pH 8.0 and 293 K. Peak assignments are indicated according to numbering in the native sequence. N-terminal His-tag residues have negative numbers. In the case of duplicated signals related to partial cleavage of the C-terminus, major and minor species are indicated by *letters a* and *b*, respectively. Asn, Gln and Trp side chain resonances are denoted with δ and ε, and NH_2 signals connected by *horizontal lines*



Crl_{STM} is a 133 residue protein containing 20 aromatic residues (4 tryptophans, 4 tyrosines, 10 phenyalanines and 2 histidines) that contribute to the core structure of the protein by stacking the two main helices α1 and α3 onto the central five stranded β-sheet (the topology of Crl is indicated in Fig. 2). A high proportion of ^{1}H chemical shift outliers was found in the vicinity of these aromatic moieties, ranging from backbone atoms (Gly55-$H_{\alpha 3}$, Thr68-H_{α}, Gly80-$H_{\alpha 3}$, His104-H_{α}, Leu38-$H_{\beta 3}$, L107-$H_{\beta 3}$) to methyl (Val42-$H_{\gamma 2*}$, Leu60-$H_{\delta 1*}$, L100-$H_{\delta 1*\delta 2*}$) and aromatic protons (Phe76-$H_{\delta *}$). Conversely the high amount of aromatic residues may explain the difficulty to assist backbone chemical shift assignment of Crl_{STM} by prediction from homology models based on the X-ray structures of *P. mirabilis* Crl_{PM} (47 % sequence identity, PDB IDs 4Q11 and 3RPJ). Nevertheless the secondary structure prediction of Crl_{STM} by TALOS-N (Shen and Bax 2013) matches

well with the X-ray structure of Crl_{PM}, as shown in Fig. 2.

Fig. 2

Sequence alignment of the Crl_{STM} construct used for NMR measurements with *Proteus mirabilis* Crl. Secondary structure elements determined from the X-ray structure of Crl_{PM} with PDB accession number 4Q11 (Cavaliere et al. 2014) are depicted on top for comparison with the secondary structure prediction of Crl_{STM} by TALOS-N (Shen and Bax 2013). Crl_{STM} residues with unassigned amide signals are shown in *green*; those located in the central β-sheet and for which line broadening by exchange precludes amide assignment are in *green bold* type. Residues with broad but assigned amide signals are in *sky blue bold* type. The C-terminal stretch that is partly cleaved by proteolysis is *boxed* in *magenta*. Residues that are spatially close to this stretch in the Crl_{PM} structure, or can be relayed to it via side chain contacts, and display two amide signals, corresponding to the full-length and truncated forms, are shown in *magenta bold* type

Severe broadening of amide signals constituted a major issue for assignment, even when lowering the pH from 8.0 to 6.5, at which Crl_{STM} was considerably less stable. In the case of Glu89-Thr90, the lack of N-terminal capping of helix α3 explains line broadening. However most of the broadening occurred in the central β-sheet, in addition to the flexible loop 1 between helix α1 and strand β1 (Fig. 2, residues in bold green and sky blue type). Amide resonances are missing downstream of and inside strand β1 (Asp30-Phe34, Cys37), in strand β3 (Phe73 and Leu75), but also in the central strand β2 (Trp56, Trp57 and Glu59), suggesting that conformational fluctuations at the millisecond time scale inside the β-sheet contribute to line broadening.

Several Crl_{STM} signals, mostly amides, were duplicated with a 60:40 intensity ratio. In the ^{15}N–HSQC spectrum (Fig. 1) the corresponding peaks are marked “a” and “b”. Although Crl_{PM} crystallized as a dimer with a small intermolecular β-sheet formed by the β4′ extension, Crl_{STM} dimerization in solution was ruled out by analytic ultracentrifugation (Cavaliere et al. 2014) as well as by ^{15}N relaxation measurements. ^{15}N R_1 and R_2 relaxation rates yielded an average effective rotational correlation time of 11.3 ± 1.0 ns and a hydrodynamic radius of 23 Å, compatible with a monomeric state of Crl_{STM}. No evidence could be found either for significant proline cis–trans isomerization. Duplicated signals were prominently clustered in helix α3 (Glu94-Leu96, Tyr98-L100, Phe103-Glu105) and proximal to those displaying line broadening, i.e. in strands β1 (Phe35, Leu38-Ala39), β2 (Gly55) and β3 (Tyr69, Tyr71, Gly74). The equivalent regions in Crl_{PM} are in close contact in the Crl_{PM} X-ray structure (Cavaliere et al. 2014), indicating that a common perturbation might be responsible for the minor signals (Fig. 2, residues in magenta). Subsequent NMR spectra of Crl_{STM} samples purified only by affinity tag displayed a set of additional intense peaks corresponding to the C-terminal sequence Asp122–Ala133 (magenta box in Fig. 2), revealing a proteolytic site in the C-terminus. Proteolytic activity was mostly removed from the samples during purification, but full-length and truncated Crl_{STM} co-eluted during size exclusion chromatography, accounting for signal duplication in differently labeled Crl_{STM} samples.

The chemical shift assignments for Crl_{STM} have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 25476.

Acknowledgments

This work was supported by the French National Research Agency (Grant ANR-11-BSV3-009) and the IR-RMN-THC (CNRS FR3050)

Conflict of interest The authors declare that they have no conflict of interest.

References

Banta AB, Chumanov RS, Yuan AH, Lin H, Campbell EA, Burgess RR, Gourse RL (2013) Key features of sigmaS required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. Proc Natl Acad Sci USA 110(40):15955–15960.

doi:10.1073/pnas.1311642110

Banta AB, Cuff ME, Lin H, Myers AR, Ross W, Joachimiak A, Gourse RL (2014) Structure of the RNA polymerase assembly factor Crl and identification of its interaction surface with sigma S. J Bacteriol 196(18):3279–3288. doi:10.1128/JB.01910-14

Battesti A, Majdalani N, Gottesman S (2011) The RpoS-mediated general stress response in *Escherichia coli*. Annu Rev Microbiol 65:189–213. doi:10.1146/annurev-micro-090110-102946

Cavaliere P, Levi-Acobas F, Mayer C, Saul FA, England P, Weber P, Raynal B, Monteil V, Bellalou J, Haouz A, Norel F (2014) Structural and functional features of Crl proteins and identification of conserved surface residues required for interaction with the RpoS/sigmaS subunit of RNA polymerase. Biochem J 463(2):215–224. doi:10.1042/BJ20140578

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293

Hengge R (2010) The general stress response in gram-negative bacteria. In: Hengge GSaR (ed) Bacterial stress responses. ASM Press, Washington, pp 251–289

Monteil V, Kolb A, D'Alayer J, Beguin P, Norel F (2010a) Identification of conserved amino acid residues of the *Salmonella* sigmaS chaperone Crl involved in Crl-sigmaS interactions. J Bacteriol 192(4):1075–1087. doi:10.1128/JB.01197-09

Monteil V, Kolb A, Mayer C, Hoos S, England P, Norel F (2010b) Crl binds to domain 2 of sigma(S) and confers a competitive advantage on a natural *rpoS* mutant of *Salmonella enterica* serovar Typhi. J Bacteriol 192(24):6401–6410. doi:10.1128/JB.00801-10

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56(3):227–241. doi:10.1007/s10858-013-9741-y

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich

EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59(4):687–696. doi:10.1002/prot.20449