


Article

High-Resolution Crystal Structure of RpoS Fragment including a Partial Region 1.2 and Region 2 from the Intracellular Pathogen *Legionella pneumophila*

Nannan Zhang ^{1,2,†}, Xiaofang Chen ^{2,†}, Xiaojian Gong ², Tao Li ¹, Zhiyuan Xie ¹, Muhammad Fazal Hameed ², Mingzhu Wang ^{1,2} and Honghua Ge ^{1,2,*} 

¹ School of Life Science, Anhui University, Hefei 230601, Anhui, China; 12055@ahu.edu.cn (N.Z.); leopard@mail.ustc.edu.cn (T.L.); 13365690310@163.com (Z.X.); wangmzh@ibp.ac.cn (M.W.)

² Institute of Physical Science and Information Technology, Anhui University, Hefei 230601, Anhui, China; chenxf0714@163.com (X.C.); 1gonglingjian@163.com (X.G.); hameed90@163.com (M.F.H.)

* Correspondence: gehonghua@gmail.com; Tel.: +86-551-6386-1773

† These authors contributed equally to this work.

Received: 25 November 2017; Accepted: 20 January 2018; Published: 23 January 2018

Abstract: *Legionella pneumophila* RpoS (*LpRpoS*), an alternative sigma factor of RNA polymerase (RNAP), is essential for virulence and stress resistance. To investigate the mechanism of RpoS in the intracellular pathogen *L. pneumophila*, we determined the high-resolution crystal structure of the *LpRpoS*^{95–195} containing a partial region 1.2 and region 2. The structure of *LpRpoS*^{95–195} reveals that the conserved residues are critical for promoter melting, DNA and core RNAP binding. The differences in regulatory factor binding site between *Escherichia coli* RpoS and *LpRpoS* suggest that *LpRpoS* may employ a distinct mechanism to recruit alternative regulatory factors controlling transcription initiation.

Keywords: RpoS; crystal structure; *Legionella pneumophila*; intracellular pathogen; regulatory factor

1. Introduction

Bacteria are able to rapidly sense and adapt to varying environmental conditions in order to survive. In particular, pathogenic bacteria can often resist various environmental stresses during transmission and infection. Many bacteria require alternative sigma factors to initiate transcription and regulate the expression of a specific set of genes in order to adapt to changing environments. RpoS is an alternative sigma factor of RNA polymerase (RNAP) that is essential for diverse bacterial stress resistances and adaptations. In many Gram-negative pathogens, RpoS plays an important role in response to diverse stresses, including starvation, oxidative stress, UV irradiation, and acidic conditions [1].

Legionella pneumophila, the causative agent of Legionnaires' disease, is a Gram-negative facultative intracellular pathogen capable of multiplying in a wide spectrum of eukaryotic cells. It adopts a biphasic life cycle consisting of a replicative and transmissive phase in a host cell. *L. pneumophila* replication is necessary to initiate lung destruction, and this process requires massive virulence factors expression [2]. In order to investigate the intracellular pathogens mechanism of improved survival in host eukaryotic cells, *L. pneumophila* has been exploited as a model intracellular pathogen for a variety of molecular and biochemical studies. To establish a replicative niche inside host cells and escape from host cell vacuoles, *L. pneumophila* has been shown to utilize RpoS to initiate the expression distinct groups of virulence factors for infection of host cells and intracellular growth [3]. In contrast to a general stress response regulator RpoS in *Escherichia coli* (*EcRpoS*), RpoS from *L. pneumophila* (*LpRpoS*) was identified as necessary not only for osmotic stress resistance but also for

virulence [4]. According to the previous studies, the characterized RpoS contains four structural motifs (regions 1.2, 2, 3, and 4) connected by flexible linkers. The prokaryotic promoters recognized by RpoS contain -35 elements, extended -10 elements, -10 elements and discriminator elements [5]. Region 1.2 plays an important role in RNAP activity, promoter DNA melting, and stabilization of the open promoter complex [6]. Region 2 is the most highly conserved in RpoS family. Several key residues in RpoS region 2 are critical for binding RNAP, recognizing -10 and discriminator element, and melting dsDNA [7]. Extended -10 and -35 elements are recognized by RpoS regions 3 and 4, respectively [8]. Compared to the housekeeping sigma factors, RpoS has relatively weak affinity to the RNAP core [9]. The RpoS binding to the RNAP core can be modulated by regulatory factors. Crl as one of regulatory factors can directly bind to RpoS, and functions to increase the affinity to the RNAP core, thus promoting transcription [10].

Although *LpRpoS* plays a critical role in the expression of secreted virulence factors and the *L. pneumophila* pathogenesis, the mechanism of its regulating expression of virulence factors is still elusive. Despite the presence of different RpoS, there are no structures available from the facultative intracellular pathogens. To gain insight into the structure and function of *LpRpoS*, we try to determine the crystal structure of full-length *LpRpoS*. However, full-length *LpRpoS* was unstable and degraded into a more stable fragment (residues 95–195) containing a partial region 1.2 and region 2. Finally, we present the high-resolution crystal structure of *LpRpoS*^{95–195}. This structure confirms that *LpRpoS*^{95–195} is more resistant to degradation than the other regions. Structural comparisons of *LpRpoS*^{95–195} and the corresponding region from *EcRpoS* reveal *LpRpoS* have non-conserved residues in a helix $\alpha 2$ adjacent to N-terminal of region 2 and an unusual 3_{10} -helix in subregion 2.3, implying that *LpRpoS* has different regulatory factor binding site. We speculated that *L. pneumophila* may employ a distinct mechanism to recruit alternative regulatory factors controlling transcription initiation. This study will lead to further explorations of molecular mechanisms used by RpoS to control gene expression in intracellular pathogens.

2. Materials and Methods

2.1. Cloning, Expression and Purification of *LpRpoS*

The coding sequence for *LpRpoS* was amplified by PCR from *Legionella pneumophila* genomic DNA by using sense (5' GGAATTCATATGATGTTAAGAAGTAAAAACTATTTCAAGG 3') and antisense (5' CCGCTCGAGTCAAACAATCCTCTTGTGTTAGTC 3') primers. The amplified fragments were then cloned to a modified pET28a vector (Novagen, Madison, MI, USA) with an additional 6× His coding sequence following. The recombinant *LpHGPR*T expression plasmid was confirmed by restriction-endonuclease digestion and further verified using DNA sequencing (Sangon Biotech, Shanghai, China). 20 μ L of a glycerol stock of the transformed *E. coli* BL21 (DE3) (Novagen) was inoculated into 4 mL Luria Broth (LB) media with supplement of 0.1 mg mL⁻¹ kanamycin. This culture was grown overnight at 310 K and then transferred into 400 mL LB with supplement of 0.01 mg mL⁻¹ kanamycin. The culture was grown at 310 K to an A_{600nm} of 0.76 and induced at 290 K with 0.25 mM isopropyl β -D-thiogalactoside (IPTG) for 16 h. After harvesting, cells were resuspended in 50 mL buffer of 200 mM NaCl, 20 mM Tris-HCl, pH 8.0. After three cycles of freeze-thaw followed by 3 min sonication, the lysed cells were centrifuged at 15,000× *g* for 40 min. The supernatant was loaded onto a Ni²⁺ NTA (GE Healthcare, Chicago, IL, USA) equilibrated with binding buffer (200 mM NaCl, 20 mM Tris-HCl, pH 8.0) and the target protein was eluted with elution buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 250 mM imidazole). The eluate was subsequently loaded onto a Superdex 200 column (Amersham Biosciences, Piscataway, NJ, USA) equilibrated with 200 mM NaCl, 20 mM Tris-HCl, pH 8.0. Protein concentration was determined by the Bradford method (Bio-Rad protein assay, Hercules, CA, USA) using bovine serum albumin as a standard. Overall yield for this protein was 53.9 mg per liter of culture. The purity of the fractions was checked on the SDS-PAGE. The purified protein

was concentrated to 20 mg mL⁻¹ using Amicon Ultracel 10 K cut-off concentrator (Merck Millipore, Darmstadt, Germany).

2.2. Crystallization and Data Collection

LpRpoS was crystallized at 289 K using sitting-drop vapor diffusion. The crystals of *LpRpoS* were grown within 3 weeks in a drop of 20 mg mL⁻¹ protein in 200 mM NaCl, 20 mM Tris-HCl, pH 8.0 with an equal volume of reservoir solution containing 0.2 M NaCl, 0.1 M Hepes pH 7.5, 12% Polyethylene glycol monomethyl ether 8000. The crystals were harvested using cryoloops and immersed briefly in a cryoprotectant solution containing the reservoir solution with 15% glycerol. The crystals were subsequently flash-frozen and stored in liquid nitrogen. The X-ray diffraction data were collected at beamline BL17U [11] at the Shanghai Synchrotron Radiation Facility (SSRF). The data set from crystal of *LpRpoS*⁹⁵⁻¹⁹⁵ were processed with Mosflm [12].

2.3. Structure Determination and Refinement

The structure of *LpRpoS*⁹⁵⁻¹⁹⁵ was solved by molecular replacement. The sigma 70 (chain Y 375–454) coordinates from *E. coli* sigma70 holoenzyme structure (PDB code 4yg2 [13]), which has 65% identity to the target structure, was used as the search model. A single solution obtained using Phaser [14] showed a log-likelihood gain (LLG) of 86 and Z scores for the rotation function (RFZ) of 3.9 and translation function (TFZ) of 12.5. This initial model was subjected to 20 cycles of rigid-body refinement and 10 cycles of restrained refinement using REFMAC5 [15], which resulted in R_{factor} and R_{free} values of 45.7% and 49.5%, respectively. The model was completed by iterative manual building in Coot [16] and refined with REFMAC [15] and PHENIX [17]. The final refined model contains two *LpRpoS*⁹⁵⁻¹⁹⁵ molecules in the asymmetric unit and was refined to an $R_{\text{factor}}(R_{\text{free}})$ of 13.7% (17.7%). The final crystallographic model was evaluated using MolProbity [18]. The final coordinates and structure factors were deposited in the Protein Data Bank (<http://www.rcsb.org/pdb>) under the accession code 5H6X for *LpRpoS*⁹⁵⁻¹⁹⁵. The refinement statistics are summarized in Table 1. Amino acid sequences were aligned by Multalin [19], and the figure of structure-based sequence alignment was generated using ESPript [20]. All illustrations were prepared with PyMOL (<http://www.pymol.org>).

Table 1. Data collection and refinement statistics.

| | <i>LpRpoS</i> ⁹⁵⁻¹⁹⁵ |
|-----------------------------------|---|
| SSRF beamline | BL17U |
| Wavelength (Å) | 0.97915 |
| Space group | <i>P</i> 2 ₁ 2 ₁ 2 ₁ |
| Molecules/ASU | 2 |
| <i>Cell parameters</i> | |
| a/b/c (Å) | 37.93/68.84/70.38 |
| α/β/γ (°) | 90.00/90.00/90.00 |
| Resolution range (Å) | 33.4–1.45 (1.53–1.45) |
| No. of unique reflections | 30,122 (3739) |
| Corresponding % solvent | 39.8 |
| R_{pim} ^a (%) | 3.8 (24.4) |
| $CC_{1/2}$ | 0.99 (0.86) |
| Average I/σ(I) | 12 (3.1) |
| Redundancy | 13.8 (14) |
| Completeness (%) | 90.7 (78.5) |

Table 1. Cont.

| <i>LpRpoS</i> ^{95–195} | |
|--|---------------------|
| <i>Refinement statistics</i> | |
| Resolution range (Å) | 20–1.45 (1.50–1.45) |
| R-factor ^b (%) | 13.7 (20.8) |
| R-free ^c (%) | 17.7 (22.9) |
| RMSD ^d bond length (Å) | 0.007 |
| RMSD bond angles (°) | 1.08 |
| No. of non-H atoms | |
| Protein | 1616 |
| Solvent | 192 |
| Average of B factors (Å ²) | |
| Protein | 21.5 |
| Solvent | 36.0 |
| Ramachandran plot ^f (%) | |
| Ramachandran favored | 99 |
| Ramachandran Outliers | 0 |
| PDB ID code | 5H6X |

Values in parentheses are for the highest resolution shell. ^a $R_{p.i.m} = \frac{\sum_{hkl} [1/(n_{hkl} - 1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where n_{hkl} is the number of observations of reflection hkl . ^b R-factor = $\frac{\sum_h ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively. Summation includes all reflections used in the refinement. ^c Free R factor = $\frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|}$, evaluated for a randomly chosen subset of 5% of the diffraction data not included in the refinement. ^d Root-mean square-deviation from ideal values. ^f Categories were defined by Molprobit.

3. Results

3.1. Crystallization of *LpRpoS*^{95–195}

We have attempted to crystallize full length *LpRpoS*. However, unit cell parameters of crystals were in obvious discrepancy with the molecular weight of the expressed protein. SDS-PAGE analysis of a dissolved crystal revealed a 10 kDa fragment, much smaller than the predicted size of the cloned protein (41 kDa). N-terminal sequencing identified the residues EIGFS. Coupled with mass spectrometry analysis, we determined that the crystallized fragment contains a partial region 1.2 and a region 2 of *LpRpoS*^{95–195} (Figure 1a).

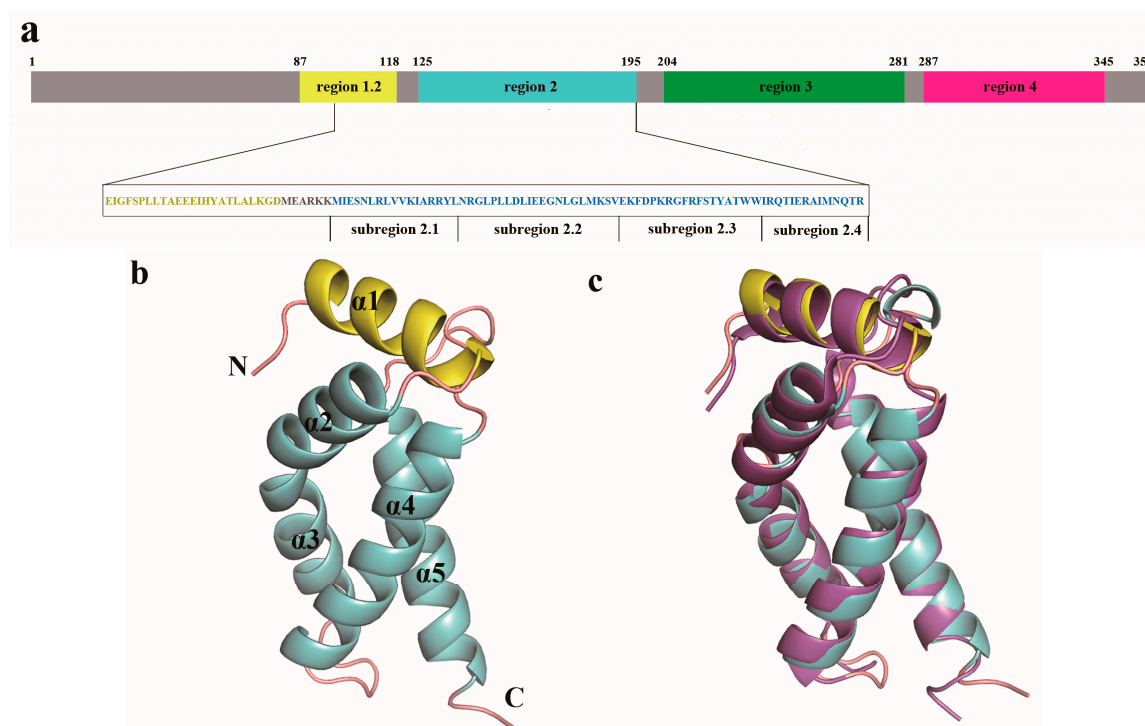


Figure 1. (a) A diagram to show the organization of *LpRpoS*. (b) Cartoon representation of *LpRpoS*^{95–195} tertiary structure. The secondary structure elements are labeled to show the monomer. The secondary structure elements from region 1.2 and region 2 are shown in yellow and cyan, respectively. (c) Superposition of crystal structure *LpRpoS*^{95–195} and the corresponding region from *EcRpoS*.

3.2. Overall Structure of *LpRpoS*^{95–195}

The crystal of *LpRpoS*^{95–195} diffracts to 1.45 Å resolution and belongs to the $P2_12_12_1$ space group, with two molecules in an asymmetric unit. Analysis using *PISA* [21] suggested that a monomer is likely to be the biologically relevant form of the protein. *LpRpoS*^{95–195} adopts a helix-turn-helix fold consisting of five helices ($\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 4$ and $\alpha 5$) (Figure 1b). Structure homology search by the *DALI* [22] server reveals that the *LpRpoS*^{95–195} structure is very similar to the corresponding regions from the *E. coli* RpoS (*EcRpoS*) structure (PDB code 5ipl [23]). *LpRpoS*^{95–195} shares 61% sequence identity and 76% sequence similarity to the corresponding regions of *EcRpoS*. The superimposition of *LpRpoS*^{95–195} monomer structure and its homologs showed that the architecture of *LpRpoS*^{95–195} is remarkably conserved (Figure 1c). Core RNAP binding, recognition and binding of the –10 promoter element, and promoter melting are all controlled by the well-conserved residues from RpoS region 2 [7]. RpoS region 2 is divided into four parts, subregions 2.1, 2.2, 2.3 and 2.4.

Structure-based alignment of *LpRpoS*^{95–195} and the corresponding region from *EcRpoS* revealed that the conserved residues appear to be critical for promoter melting, DNA and core RNAP binding (Figure 2a). The *LpRpoS*^{95–195} is similar to the structure of the corresponding region from *EcRpoS* complexed with nucleic acids (Figure 2b). The Arg129 and Lys133 from subregions 2.2 and 2.3 of *EcRpoS* related to DNA binding are also conserved in *LpRpoS*^{95–195} (corresponding residues Lys160 and Lys164) (Figure 2c) [24]. Subregion 2.3 from *LpRpoS*^{95–195} has highly conserved aromatic residues (Phe171, Tyr176, Trp179, Trp180) implicated in promoter melting (Figure 2d) [24]. In *EcRpoS*, conserved residues Gln152, Glu155 and Arg156 from subregion 2.4 are involved in recognizing the –10 promoter element (Figure 2e) [24]. Meanwhile, these residues are also found in corresponding sites of *LpRpoS*^{95–195}. Similar to *EcRpoS* structure, *LpRpoS*^{95–195} contains the conserved residues from subregions 2.2 involving in core RNAP binding (Figure 2a). These structure comparison results indicated that *LpRpoS*^{95–195} likely have a similar way to melt promoter, bind core RNAP, recognize and bind the

–10 promoter element. Although *EcRpoS* and *LpRpoS*^{95–195} have the high similarity in recognition and binding of the –10 promoter element, promoter melting sites and core RNAP binding site, there are also some a significant distinction in biological function between *EcRpoS* and *LpRpoS*. Unlike *EcRpoS*, which is required for a general stress resistance, *LpRpoS* is only involved in osmotic shock but in no other stress resistance, and is also more important for the production of virulence factors [4].

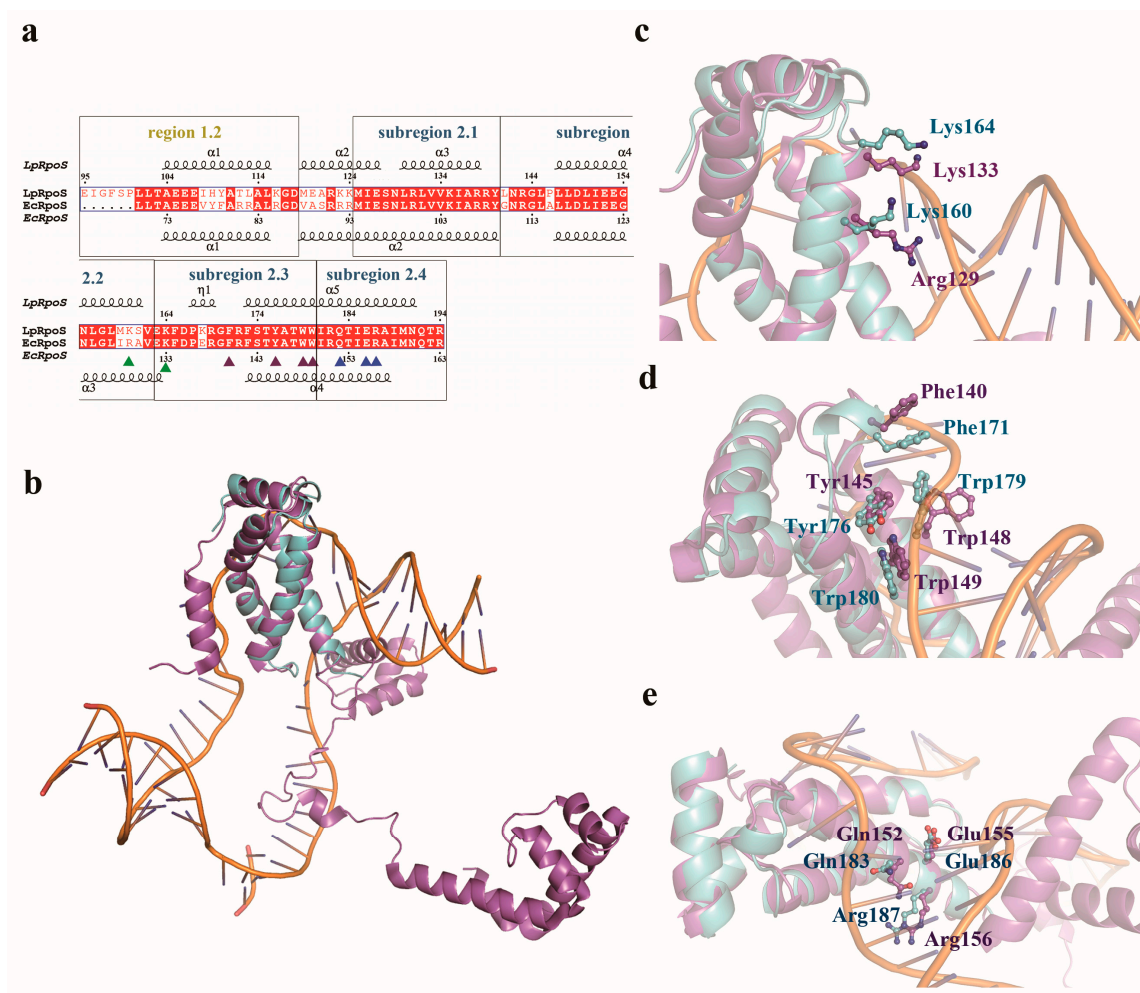


Figure 2. (a) Multiple alignment of *LpRpoS*^{95–195} and the corresponding region from *EcRpoS* (PDB accession code 5ipl). The alignment is performed using MultAlin [11] and ESPrnt [12]. The secondary structural elements of *LpHGPRT* and human *HGPRT* are separately displayed at the top and bottom of the alignment. The α -helices, 3_{10} -helices, β -sheets and strict β -turns are denoted as α , η , β and TT, correspondingly. The DNA binding site, promoter melting site and the residues involved in recognizing the –10 promoter element are marked by filled green, purple and blue triangles, respectively. The key residues from *EcRpoS* (Asp87, Asp135, Pro136, and Glu137) and the corresponding residues from *LpRpoS*^{95–195} (Asp118, Asp166, Pro167 and Lys168) are marked in pink. (b) Superposition of *LpRpoS*^{95–195} (cyan) and full-length *EcRpoS* complexed with DNA (purple) (PDB accession code 5ipl). DNA is marked in orange. (c) Comparison of *LpRpoS*^{95–195} (cyan) and *EcRpoS* (purple) (PDB accession code 5ipl) structures in DNA binding site. The residues, which are involved in DNA binding site, are shown in stick form. (d) Comparison of *LpRpoS*^{95–195} (cyan) and *EcRpoS* (purple) (PDB accession code 5ipl) structures in promoter melting site. The residues, which are involved in promoter melting site, are shown in stick form. (e) Comparison of *LpRpoS*^{95–195} (cyan) and *EcRpoS* (purple) (PDB accession code 5ipl) structures in recognizing the –10 promoter element site. The residues, which are involved in recognizing the –10 promoter element, are shown in stick form.

3.3. *LpRpoS* Show a Different Regulatory Factor Binding Site from *EcRpoS*

Sigma factors compete for binding to a limited amount of the core RNAP. In contrast to housekeeping *E. coli* sigma factor 70 with strong affinity for the RNAP core enzyme, *EcRpoS* binds relatively weakly to the core RNAP in the absence of nucleic acids [9]. It's predicted that *E. coli* sigma factor 70 with an additional non-conserved region (NCR) probably increase the affinity to interact with the core RNAP compare to *EcRpoS*. Many housekeeping sigma factors contain a NCR located between subregions 1.2 and 2.1. NCR is not present in alternative sigma factors like *RpoS*. In response to a variety of environmental conditions, different sigma factors binding the core RNAP can be modulated by several regulatory factors [25]. The *Crl*, as the only known regulatory factor, can directly interact with region 2 of *RpoS* and enhance its affinity to the RNAP core for promoting transcription [26]. The *Crl* located the similar position as NCR in sigma factor 70, and possibly functions similarly to NCR to provide extra interactions with core RNAP. Despite of non-conserved and less widespread in bacterial species, *Crl* homologs share a similar mechanism and function. Previously studies showed that *RpoS* interacting with *Crl* involves two regions including α -helix adjacent to N-terminal of *RpoS* region 2 and loop in *RpoS* subregion 2.3. It is noteworthy that the two regions from *RpoS* are exposed to the outside of the RNAP main channel for *Crl* recognition. Meanwhile, it also found that the key residues of Asp87, Asp135, Pro136, and Glu137 from *EcRpoS* are important for *EcRpoS* interacting with *Crl* [26]. Asp87 locate on α -helix near the N-terminal of *EcRpoS* region 2. Asp135, Pro136, and Glu137 from *EcRpoS* region 2 with the DPE motif. In particular, the charge of Glu137 in the DPE motif is crucial for *Crl* recognition [27].

*LpRpoS*⁹⁵⁻¹⁹⁵ and the corresponding region of *EcRpoS* have very similar structure and highly conserved the core RNAP binding site. It deduced that *LpRpoS* may have low affinity to core RNAP, and probably require regulatory factor to bind core RNAP. Unlike *E. coli* that utilizes *Crl* regulating *EcRpoS*, *L. pneumophila* contains conserved *RpoS* but does not have *Crl*. Comparison of *LpRpoS*⁹⁵⁻¹⁹⁵ and *EcRpoS* structures reveals significant differences between *Crl*-binding site of *EcRpoS* and the corresponding region from *LpRpoS*. Although the conserved residue Asp118 found in helix α 2 near the N-terminal of *LpRpoS*⁹⁵⁻¹⁹⁵ region 2, *LpRpoS*⁹⁵⁻¹⁹⁵ does not have the conserved DPE motif. Structural comparison between *LpRpoS*⁹⁵⁻¹⁹⁵ and *EcRpoS* show that Lys168 of *LpRpoS*⁹⁵⁻¹⁹⁵ replaces the corresponding residue Glu137 from *EcRpoS* (Figure 3). *EcRpoS* Glu137Gln mutant did directly affect *Crl* binding, suggesting that the negative charge of Glu137 is importance for *Crl* binding. The Lys168 of *LpRpoS*⁹⁵⁻¹⁹⁵ substituted Glu137 from *EcRpoS* reveals that *LpRpoS*⁹⁵⁻¹⁹⁵ might utilize a different mechanism to regulate factor recognition. Meanwhile, the residues Asp166, Pro167 and Lys168 of *LpRpoS* located on a region between α 4 and α 5 adopt an unusual 3_{10} -helix architecture. It suggests that the special 3_{10} -helix fold and the oppositely charged residue may together compose a distinct regulatory factor binding site. The region 1.2 adjacent to region 2 was proposed as a recognition site of the regulatory factor *Crl*. *Salmonella enterica* serovar Typhimurium *RpoS* R82L mutant is defective in *Crl* binding and activation [28]. The fact that *LpRpoS*⁹⁵⁻¹⁹⁵ uses the corresponding Leu113 as a substitute for Arg82 in *StRpoS* indicates that *LpRpoS*⁹⁵⁻¹⁹⁵ may not be suitable for *Crl* binding. It also found that the helix from region 1.2 adjacent to the N-terminal of *LpRpoS* region 2 contains non-conserved residues compared with the corresponding region in *EcRpoS*. Altogether, these findings indicate that *LpRpoS* may enhance its affinity to core RNAP through binding alternative regulatory factor.

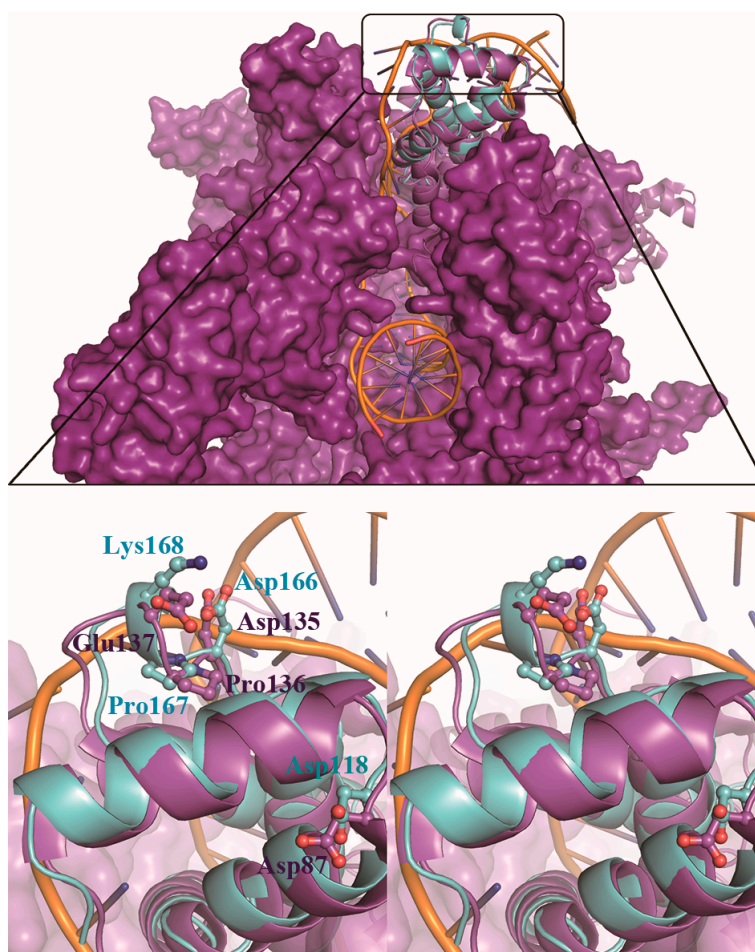


Figure 3. Superposition of Crl binding site from *EcRpoS* (purple) (PDB accession code 5ipl) with the corresponding site from *LpRpoS*^{95–195} (cyan). *LpRpoS*^{95–195} and *EcRpoS* from *E. coli* RNAP holoenzyme are shown as cartoon in addition to the molecular surfaces of other *E. coli* RNAP subunits. DNA is marked in orange. The right picture is magnified view of the boxed region in superposition of *LpRpoS*^{95–195} and *EcRpoS*. The key residues from *EcRpoS* (Asp87, Asp135, Pro136, and Glu137) and the corresponding residues from *LpRpoS*^{95–195} (Asp118, Asp166, Pro167 and Lys168) are shown in stick form.

4. Discussion

RpoS play a role in stress response gene expression and production of virulence factors in *L. pneumophila*. Our analysis showed that *LpRpoS* displays highly conserved residues and has folds that are similar to the corresponding regions in *E. coli* RpoS. Most interestingly, Sequence and conformation of the region 2 in *EcRpoS*, which is involved in regulatory factor Crl binding, is different from that observed in *LpRpoS*^{95–195}. It found that the Crl binding site at loop of *EcRpoS* (Asp87, Asp135, Pro136, and Glu137) is different from the corresponding site of *LpRpoS* with the residues Asp166, Pro167 and Lys168, which located on a region between $\alpha 4$ and $\alpha 5$ adopt an unusual 3_{10} -helix architecture. Importantly, we determined that the function of Glu137 in *E. coli* is substituted by Lys168 in *LpRpoS*. Meanwhile, there is no Crl homologs in *L. pneumophila*. These observations suggest that *L. pneumophila* may apply a distinct mechanism to employ alternative regulatory factors controlling *LpRpoS* transcription initiation. This mechanism is probably important for *L. pneumophila* survival and virulence. The structures present here further shed light into understanding the mechanism of the intracellular pathogen pathogenesis.

Acknowledgments: The authors thank Zhaoqing Luo (Purdue University) for *L. pneumophila* genomic DNA. We also thank the staff at SSRF beamline BL17U for assistance with synchrotron data collection. This work was supported by grants from the National Natural Science Foundation of China (31400641, 31270770) and the Science Foundation of Anhui University (J18520219, KYXL2016105).

Author Contributions: Nannan Zhang and Honghua Ge designed experiments and analyzed the data; Xiaofang Chen performed experiments; Tao Li and Zhiyuan Xie contributed reagents and materials; Xiaojian Gong and Mingzhu Wang solved the structure; Nannan Zhang, Muhammad Fazal Hameed and Honghua Ge wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dong, T.; Schellhorn, H.E. Role of rpos in virulence of pathogens. *Infect. Immun.* **2010**, *78*, 887–897. [[CrossRef](#)] [[PubMed](#)]
2. Isberg, R.R.; O'Connor, T.J.; Heidtman, M. The legionella pneumophila replication vacuole: Making a cosy niche inside host cells. *Nat. Rev. Microbiol.* **2009**, *7*, 13–24. [[CrossRef](#)] [[PubMed](#)]
3. Abu-Zant, A.; Asare, R.; Graham, J.E.; Abu Kwaik, Y. Role for rpos but not rela of legionella pneumophila in modulation of phagosome biogenesis and adaptation to the phagosomal microenvironment. *Infect. Immun.* **2006**, *74*, 3021–3026. [[CrossRef](#)] [[PubMed](#)]
4. Bachman, M.A.; Swanson, M.S. Genetic evidence that legionella pneumophila rpos modulates expression of the transmission phenotype in both the exponential phase and the stationary phase. *Infect. Immun.* **2004**, *72*, 2468–2476. [[CrossRef](#)] [[PubMed](#)]
5. Gruber, T.M.; Gross, C.A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **2003**, *57*, 441–466. [[CrossRef](#)] [[PubMed](#)]
6. Haugen, S.P.; Ross, W.; Manrique, M.; Gourse, R.L. Fine structure of the promoter-sigma region 1.2 interaction. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3292–3297. [[CrossRef](#)] [[PubMed](#)]
7. Feklistov, A.; Darst, S.A. Structural basis for promoter-10 element recognition by the bacterial rna polymerase sigma subunit. *Cell* **2011**, *147*, 1257–1269. [[CrossRef](#)] [[PubMed](#)]
8. Campbell, E.A.; Muzzin, O.; Chlenov, M.; Sun, J.L.; Olson, C.A.; Weinman, O.; Trester-Zedlitz, M.L.; Darst, S.A. Structure of the bacterial rna polymerase promoter specificity sigma subunit. *Mol. Cell* **2002**, *9*, 527–539. [[CrossRef](#)]
9. Maeda, H.; Fujita, N.; Ishihama, A. Competition among seven escherichia coli sigma subunits: Relative binding affinities to the core rna polymerase. *Nucleic Acids Res.* **2000**, *28*, 3497–3503. [[CrossRef](#)] [[PubMed](#)]
10. England, P.; Westblade, L.F.; Karimova, G.; Robbe-Saule, V.; Norel, F.; Kolb, A. Binding of the unorthodox transcription activator, crl, to the components of the transcription machinery. *J. Biol. Chem.* **2008**, *283*, 33455–33464. [[CrossRef](#)] [[PubMed](#)]
11. Wang, Z.; Pan, Q.; Yang, L.; Zhou, H.; Xu, C.; Yu, F.; Wang, Q.; Huang, S.; He, J. Automatic crystal centring procedure at the ssrf macromolecular crystallography beamline. *J. Synchrotron Radiat.* **2016**, *23*, 1323–1332. [[CrossRef](#)] [[PubMed](#)]
12. Battye, T.G.; Kontogiannis, L.; Johnson, O.; Powell, H.R.; Leslie, A.G. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 271–281. [[CrossRef](#)] [[PubMed](#)]
13. Murakami, K.S. X-ray crystal structure of escherichia coli rna polymerase sigma(70) holoenzyme. *J. Biol. Chem.* **2013**, *288*, 9126–9134. [[CrossRef](#)] [[PubMed](#)]
14. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [[CrossRef](#)] [[PubMed](#)]
15. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. Refmac5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [[CrossRef](#)] [[PubMed](#)]
16. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]

17. Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. Phenix: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221. [[CrossRef](#)] [[PubMed](#)]
18. Davis, I.W.; Leaver-Fay, A.; Chen, V.B.; Block, J.N.; Kapral, G.J.; Wang, X.; Murray, L.W.; Arendall, W.B.; Snoeyink, J.; Richardson, J.S.; et al. Molprobity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375–W383. [[CrossRef](#)] [[PubMed](#)]
19. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
20. Gouet, P.; Courcelle, E.; Stuart, D.I.; Metz, F. Espript: Analysis of multiple sequence alignments in postscript. *Bioinformatics* **1999**, *15*, 305–308. [[CrossRef](#)] [[PubMed](#)]
21. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)] [[PubMed](#)]
22. Holm, L.; Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **1993**, *233*, 123–138. [[CrossRef](#)] [[PubMed](#)]
23. Liu, B.; Zuo, Y.H.; Steitz, T.A. Structures of e-coli sigma(s)-transcription initiation complexes provide new insights into polymerase mechanism. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4051–4056. [[CrossRef](#)] [[PubMed](#)]
24. Tomsic, M.; Tsujikawa, L.; Panaghie, G.; Wang, Y.; Azok, J.; de Haseth, P.L. Different roles for basic and aromatic amino acids in conserved region 2 of escherichia coli sigma(70) in the nucleation and maintenance of the single-stranded DNA bubble in open rna polymerase-promoter complexes. *J. Biol. Chem.* **2001**, *276*, 31891–31896. [[CrossRef](#)] [[PubMed](#)]
25. Osterberg, S.; del Peso-Santos, T.; Shingler, V. Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.* **2011**, *65*, 37–55. [[CrossRef](#)] [[PubMed](#)]
26. Banta, A.B.; Chumanov, R.S.; Yuan, A.H.; Lin, H.L.; Campbell, E.A.; Burgess, R.R.; Gourse, R.L. Key features of sigma(s) required for specific recognition by crl, a transcription factor promoting assembly of rna polymerase holoenzyme. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15955–15960. [[CrossRef](#)] [[PubMed](#)]
27. Cavaliere, P.; Levi-Acobas, F.; Mayer, C.; Saul, F.A.; England, P.; Weber, P.; Raynal, B.; Monteil, V.; Bellalou, J.; Haouz, A.; et al. 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.* **2014**, *463*, 215–224. [[CrossRef](#)] [[PubMed](#)]
28. Cavaliere, P.; Sizun, C.; Levi-Acobas, F.; Nowakowski, M.; Monteil, V.; Bontems, F.; Bellalou, J.; Mayer, C.; Norel, F. Binding interface between the salmonella sigma(s)/rpos subunit of rna polymerase and crl: Hints from bacterial species lacking crl. *Sci. Rep.* **2015**, *5*, 13564. [[CrossRef](#)] [[PubMed](#)]

