Peer-reviewed version available at Crystals 2018, 8, 54; doi:10.3390/cryst8020054

Type of the Paper (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] and Honghua Ge [1,2,*]

[1] School of Life Science, Anhui University, Hefei, Anhui 230601, China; 12055@ahu.edu.cn (N.Z.); leopard@mail.ustc.edu.cn (T.L.); 13365690310@163.com (Z.X.)
[2] Institute of Physical Science and Information Technology, Anhui University, Hefei, Anhui 230601, China; chenxf0714@163.com (X.C.); 1gonglingjian@163.com (X. G.); hameed90@163.com (M.F.H.)
* Correspondence: gehonghua@gmail.com; Tel.: +86-551-63861773
† These authors contributed equally to this work.

Abstract: *Legionella pneumophila* RpoS (*Lp*RpoS) is an alternative sigma factor of RNA polymerase (RNAP) 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 *Lp*RpoS (residues 95-194) containing a partial region 1.2 and region 2. The structure of *Lp*RpoS (residues 95-194) 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 *Lp*RpoS suggest that *Lp*RpoS 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

The bacteria are able to rapidly sense and adapt to varying environmental conditions for survival. Especially for pathogenic bacteria, it usually can resist various environmental stresses during transmission and infection. For adaptation to changing environments, bacteria require alternative sigma factors to initiate transcription and regulate the expression of a specific set of genes. RpoS is an alternative sigma factor of RNA polymerase (RNAP) essential for bacterial diverse stress resistance and adaptation. In many Gram-negative pathogens, RpoS play an important role in respond to diverse stresses, including starvation, oxidative stress, UV irradiation, and acidic condition [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 cell. It adopts a biphasic life cycle consisting of a replicative and transmissive phase in a host cell. *L. pneumophila* replication is necessary to leading to 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* (*Ec*RpoS), RpoS from *L. pneumophila* (*Lp*RpoS) was identified as necessary not only for osmotic stress resistance but also for virulence [4]. According to the previous studies, the characterized RpoS contain four structural motifs (regions 1.2, 2, 3, and 4) connected by flexible linkers. The prokaryotic promoters recognized by RpoS contain -35 element, extended -10 element, -10 element and discriminator

element [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 have 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 function to increase the affinity to the RNAP core for promoting transcription [10].

Although *Lp*RpoS 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 presence of different RpoS, there was no structure from facultative intracellular pathogens available. To gain insight into the structure and function of *Lp*RpoS, we try to determine the crystal structure of full-length *Lp*RpoS. However, full-length *Lp*RpoS was unstable and degraded into a more stable fragment (residues 95-194) containing a partial region 1.2 and region 2. Finally, we present the high-resolution crystal structure of *Lp*RpoS (residues 95-194). This structure confirms that *Lp*RpoS (residues 95-194) is more resistant to degradation than the other regions. Structural comparisons of *Lp*RpoS (residues 95-194) and the corresponding region from *Ec*RpoS reveal *Lp*RpoS have non-conserved residues in a helix α2 adjacent to N-terminal of region 2 and an unusual 3_{10}-helix in subregion 2.3, implying that *Lp*RpoS 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 pathogen.

2. Results

2.1. Crystallization of *Lp*RpoS (residues 95-194)

We have attempted to crystallize full length *Lp*RpoS. 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 *Lp*RpoS (residues 95-194).

2.2. Overall Structure of *Lp*RpoS (residues 95-194)

The crystal of *Lp*RpoS (residues 95-194) diffracts to 1.6Å resolution and belongs to the $P2_12_12_1$ space group, with two molecules in an asymmetric unit. Analysis using *PISA* [11] suggested that a monomer is likely to be the biologically relevant form of the protein. *Lp*RpoS (residues 95-194) adopts a helix-turn-helix fold consisting of five helices (α1, α2, α3, α4 and α5) [Fig. 1b]. Structure homology search by the *DALI* [12] server reveals that the *Lp*RpoS (residues 95-194) structure is very similar to the corresponding regions from the *E. coli* RpoS (*Ec*RpoS) structure (PDB code 5ip [13]). The superimposition of *Lp*RpoS (residues 95-194) monomer structure and its homologs showed that the architecture of *Lp*RpoS (residues 95-194) is remarkably conserved [Fig. 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 *Lp*RpoS (residues 95-194) and the corresponding region from *Ec*RpoS revealed that the conserved residues appear to be critical for promoter melting, DNA and core RNAP binding [Fig. 2a]. The *Lp*RpoS (residues 95-194) is similar to the structure of the corresponding region from *Ec*RpoS complexed with nucleic acids [Fig. 2b]. The Arg129 and Lys133 from subregions 2.2 and 2.3 of *Ec*RpoS related to DNA binding are also conserved in *Lp*RpoS (residues 95-194) (corresponding residues Lys160 and Lys164) [Fig. 2c] [14]. Subregion 2.3 from *Lp*RpoS (residues 95-194) has highly conserved aromatic residues (Phe171, Tyr176, Trp179, Trp180)

implicated in promoter melting [Fig. 2d] [14]. In *Ec*RpoS, conserved residues Gln152, Glu155 and Arg156 from subregion 2.4 are involved in recognizing the -10 promoter element [Fig. 2e] [14]. Meanwhile, these residues are also found in corresponding sites of *Lp*RpoS (residues 95-194). Similar to *Ec*RpoS structure, *Lp*RpoS (residues 95-194) contains the conserved residues from subregions 2.2 involving in core RNAP binding [Fig. 2a]. These structure comparison results indicated that *Lp*RpoS (residues 95-194) likely have a similar way to melt promoter, bind core RNAP, recognize and bind the -10 promoter element. Although *Ec*RpoS and *Lp*RpoS (residues 95-194) 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 *Ec*RpoS and *Lp*RpoS. Unlike *Ec*RpoS required for a general stress resistance, *Lp*RpoS is only involved in osmotic shock but not other stress resistance, and more important for the production of virulence factors [4].

2.3. *Lp*RpoS show a different regulatory factor binding site from *Ec*RpoS

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, *Ec*RpoS 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 *Ec*RpoS. 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[15]. 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 [16]. The Crl located the similar position as NCR in sigma factor 70, and may be function as NCR to provide extra interactions with the 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 *Ec*RpoS are important for *Ec*RpoS interacting with Crl [16]. Asp87 locate on α-helix near the N-terminal of *Ec*RpoS region 2. Asp135, Pro136, and Glu137 from *Ec*RpoS region 2 composed of the DPE motif. Especially the charge of Glu137 in the DPE motif is crucial for Crl recognition [17].

*Lp*RpoS (residues 95-194) and the corresponding region of *Ec*RpoS have very similar structure and highly conserved the core RNAP binding site. It deduced that *Lp*RpoS may have low affinity to core RNAP, and probably require regulatory factor to bind core RNAP. Unlike *E. coli* that utilizes Crl regulating *Ec*RpoS, *L. pneumophila* contains conserved RpoS but does not have Crl. Comparison of *Lp*RpoS (residues 95-194) and *Ec*RpoS structures reveals significant differences between Crl-binding site of *Ec*RpoS and the corresponding region from *Lp*RpoS. Although the conserved residue Asp118 found in helix α2 near the N-terminal of *Lp*RpoS (residues 95-194) region 2, *Lp*RpoS (residues 95-194) does not have the conserved DPE motif. Structural comparison between *Lp*RpoS (residues 95-194) and *Ec*RpoS show that Lys168 of *Lp*RpoS (residues 95-194) replaces the corresponding residue Glu137 from *Ec*RpoS [Fig. 3]. *Ec*RpoS Glu137Gln mutant did directly affect Crl binding, suggesting that the negative charge of Glu137 is importance for Crl binding. The Lys168 of *Lp*RpoS (residues 95-194) substituted Glu137 from *Ec*RpoS reveals that *Lp*RpoS (residues 95-194) might be utilize a different mode of regulatory factor recognition. Meanwhile, the residues Asp166, Pro167 and Lys168 of *Lp*RpoS located on a region between α4 and α5 adopt an unusual 3_{10}-helix architecture. It suggested that the special 3_{10}-helix fold and the oppositely charged residue may compose of 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 [18]. The *Lp*RpoS (residues 95-194) with the corresponding Leu113 substituted Arg82 from *St*RpoS indicated that *Lp*RpoS (residues 95-194) may not be suitable for Crl binding. It also found that the helix from region 1.2 adjacent to the N-terminal

of *Lp*RpoS region 2 contains non-conserved residues compared with the corresponding region in *Ec*RpoS. Altogether, these findings indicate that *Lp*RpoS may enhance its affinity to core RNAP through binding alternative regulatory factor.

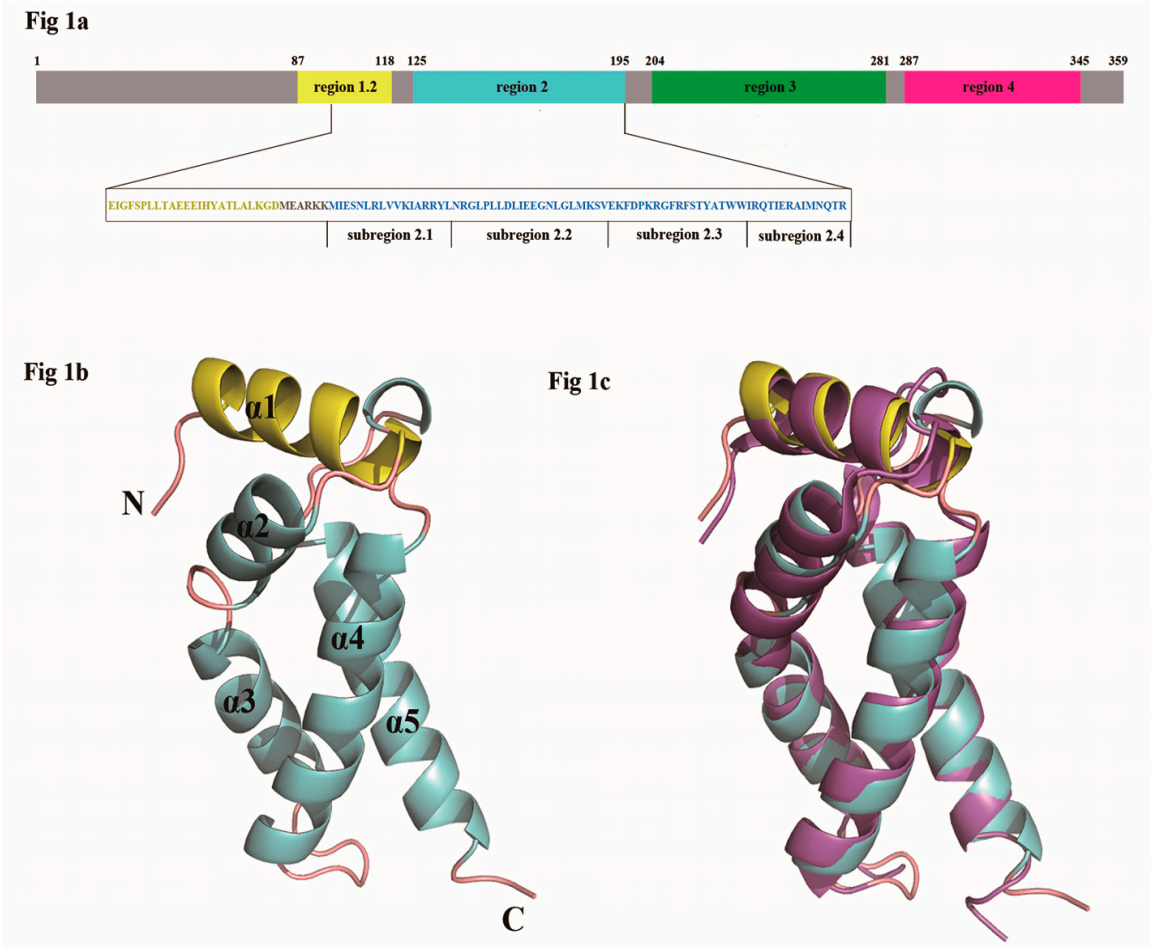

Figure 1. (a) A diagram to show the organization of *Lp*RpoS. (b) Cartoon representation of *Lp*RpoS (residues 95-194) tertiary structure. The secondary structure elements are labeled for monomer. The secondary structure elements from region 1.2 and region 2 are shown in yellow and cyan, respectively. (c) Superposition of crystal structure *Lp*RpoS (residues 95-194) and the corresponding region from *Ec*RpoS.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 November 2017 doi:10.20944/preprints201711.0182.v1

Peer-reviewed version available at Crystals 2018, 8, 54; doi:10.3390/cryst8020054

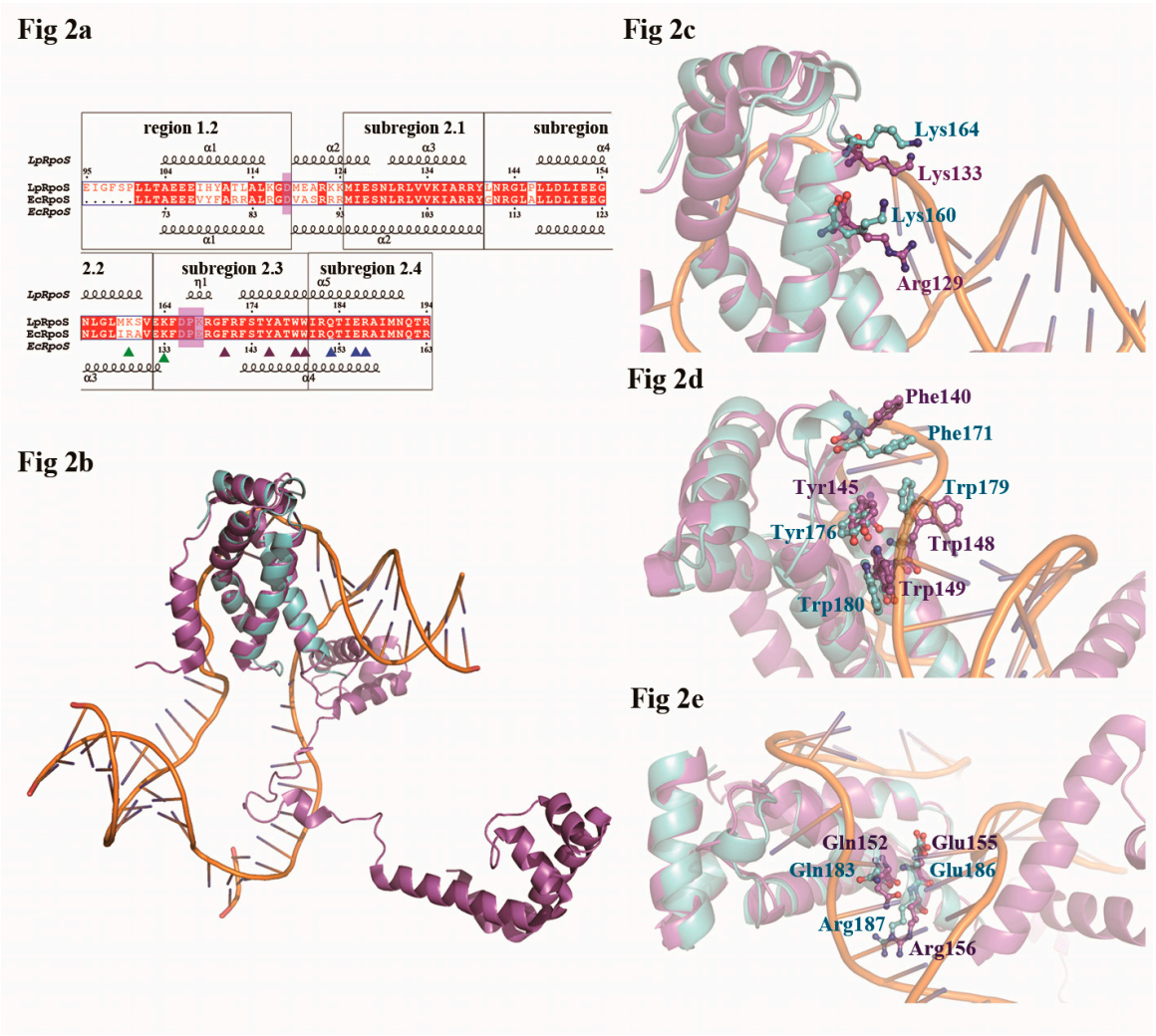

Figure 2. (a) Multiple alignment of *Lp*RpoS (residues 95-194) and the corresponding region from *Ec*RpoS (PDB accession code 5ipl). The alignment is performed using MultAlin(Corpet, 1988) and ESPript(Gouet *et al.*, 1999). The secondary structural elements of *Lp*HGPRT and human HGPRT are separately displayed at the top and bottom of the alignment. The α-helices, η-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 *Ec*RpoS (Asp87, Asp135, Pro136, and Glu137) and the corresponding residues from *Lp*RpoS (residues 95-194) (Asp118, Asp166, Pro167 and Lys168) are marked in pink. (b) Superposition of *Lp*RpoS (residues 95-194) (cyan) and full-length *Ec*RpoS complexed with DNA (purple) (PDB accession code 5ipl). DNA is marked in orange. (c) Comparison of *Lp*RpoS (residues 95-194) (cyan) and *Ec*RpoS (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 *Lp*RpoS (residues 95-194) (cyan) and *Ec*RpoS (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 *Lp*RpoS (residues 95-194) (cyan) and *Ec*RpoS (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.

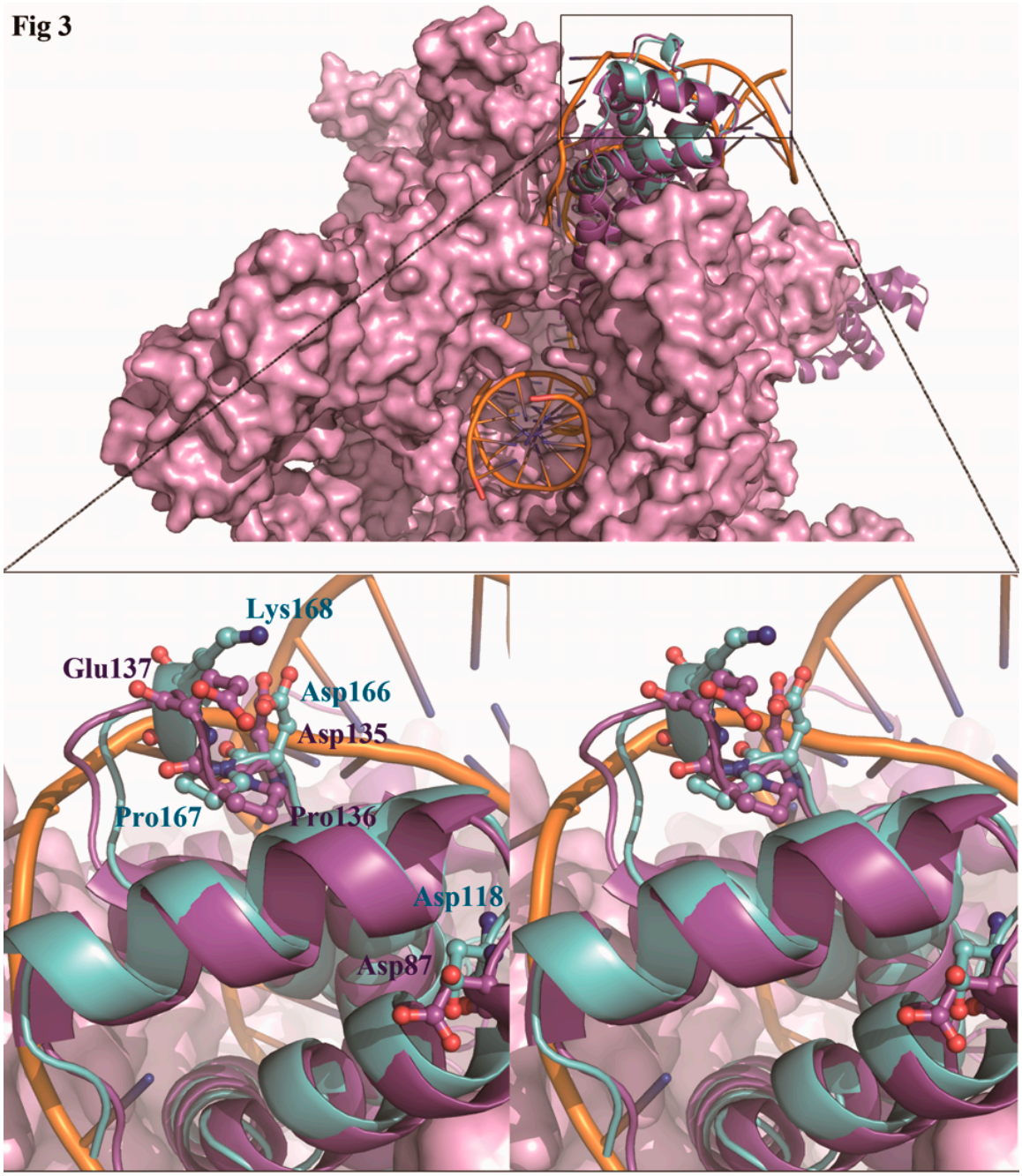

Figure 3. Superposition of Crl binding site from *Ec*RpoS (purple) (PDB accession code 5ipl) with the corresponding site from *Lp*RpoS (residues 95-194) (cyan). *Lp*RpoS (residues 95-194) and *Ec*RpoS from *E. coli* RNAP holoenzyme are shown as cartoon in addition to the molecular surfaces of other *E. coli* RNAP subunits. DNA is marked in orange. The right picture is magnified view of the boxed region in superposition of *Lp*RpoS (residues 95-194) and *Ec*RpoS. The key residues from *Ec*RpoS (Asp87, Asp135, Pro136, and Glu137) and the corresponding residues from *Lp*RpoS (residues 95-194) (Asp118, Asp166, Pro167 and Lys168) are shown in stick form.

3. Discussion

RpoS plays a role in stress response gene expression and production of virulence factors in *L. pneumophila*. The analysis showed that *Lp*RpoS displays a highly conserved residues and a similar fold as the corresponding regions from *E. coli* RpoS. Most interestingly, Sequence and conformation of the region 2 in *Ec*RpoS, which is involved in regulatory factor Crl binding, is different from that observed in *Lp*RpoS (residues 95-194). The observations suggest that *L. pneumophila* may apply a distinct mechanism to employ alternative regulatory factors controlling *Lp*RpoS transcription

initiation. It is probably important for *L. pneumophila* survival and virulence. The structure presents here further shed light into understanding the mechanism of the intracellular pathogen pathogenesis.

4. Materials and Methods

4.1. Cloning, expression and purification of *Lp*RpoS.

The coding sequence for *Lp*RpoS was amplified by PCR from *Legionella pneumophila* genomic DNA by using sense (5' GGAATTCCATATGATGTTAAGAAGTAAAAAACTATTTCAAGG 3') and antisense (5' CCGCTCGAGTCAAAACAAATCCTCTTGTGTTAGTC 3') primers. The amplified fragments were then cloned to a modified pET28a vector (Novagen, USA) with an additional 6×His coding sequence following. The recombinant *Lp*HGPRT expression plasmid was confirmed by restriction-endonuclease digestion and further verified using DNA sequencing (Sangon Biotech, Shanghai, China). 20 µl of a glycerol stock of the transformed *E. coli* Rosetta was inoculated into 4 ml Luria Broth (LB) media with supplement of 0.01 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 $A_{600\,nm}$ 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 Ni^{2+} NTA (GE Healthcare, 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, USA) equilibrated with 200 mM NaCl, 20 mM Tris-HCl, pH 8.0. The purity of the fractions was checked on the SDS-PAGE.

4.2. Crystallization and data collection

*Lp*RpoS was crystallized at 289 K using sitting-drop vapor diffusion. The crystals of *Lp*RpoS were grown 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 8,000 within 3 weeks. 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 [19] at the Shanghai Synchrotron Radiation Facility (SSRF). The data set from crystal of *Lp*RpoS (residues 95-194) were processed with HKL2000 [20].

4.3. Structure determination and refinement

The structure of *Lp*RpoS (residues 95-194) was solved by molecular replacement. The sigma 70 (chain Y 375-454) coordinates from *E. coli* sigma70 holoenzyme structure (PDB code 4yg2 [21]), which has 65% identity to the target structure, was used as the search model. A single solution obtained using Phaser [22] 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[23], 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 [24] and refined with REFMAC [23] and PHENIX [25]. The final refined model contains two *Lp*RpoS (residues 95-194) molecules in the asymmetric unit and was refined to an R_{factor}(R_{free}) of 16.4% (19.2%). The final crystallographic model was evaluated using MolProbity [26]. The final coordinate and structure factor were deposited in the Protein Data Bank (http://www.rcsb.org/pdb) under the accession code 5H6X for *Lp*RpoS (residues 95-194). The refinement statistics are summarized in Table 1. Amino acid sequences were aligned by Multalin [27], and the figure of structure-based sequence alignment was generated using ESPript [28]. All illustrations were prepared with PyMOL (http://www.pymol.org).

Table1. Data collection and refinement statistics.

	*Lp*RpoS (residues 95-194)
SSRF beamline	BL17U
Wavelength (Å)	0.97915
Space group	$P2_12_12_1$
Molecules/ASU	1
Cell parameters	
a/b/c (Å)	37.91/68.91/70.36
α/β/γ (°)	90.00/90.00/90.00
Resolution range (Å)	20-1.6 (1.640-1.599)
No. of unique reflections	23727
Corresponding % solvent	39.8
*R*merge [a] (%)	8.8 (49.6)
Average I/σ(I)	31.6 (6.1)
Redundancy	12.7 (13.3)
Completeness (%)	99.9 (100)
Refinement statistics	
R-factor[b] (%)	16.4
R-free[c] (%)	19.2
RMSD[d] bond length (Å)	0.011
RMSD bond angles (°)	1.303
No. of non-H atoms	
Protein	1597
Ion	2
Ligand	6
Solvent	185
Average of B factors (Å^2)	
Protein	24.1
Ion	19.8
Ligand	23.4
Solvent	35.1
Ramachandran plot[f] (%)	
Ramachandran favored	99.5
Ramachandran Outliers	0.5
PDB ID code	5H6X

Values in parentheses are for the highest resolution shell.

[a]Rmerge = $\Sigma|I_{hkl} - \langle I_{hkl} \rangle| / \langle I_{hkl} \rangle$, where I_{hkl} is a single value of the measured intensity of the *hkl* reflection and $\langle I_{hkl} \rangle$ is the mean of all measured values of the intensity of the *hkl* reflections.

[b]R-factor = $\Sigma h \, | \, |F_{obs}| - |F_{calc}| \, | / \Sigma |F_{obs}|$, 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 = $\Sigma \, | \, |F_{obs}| - |F_{calc}| \, | / \Sigma |F_{obs}|$, 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 Molprobity.

Acknowledgments: The authors thank Dr. 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: NZ and HG designed experiments and analyzed the data; XC performed experiments; TL and ZX contributed reagents and materials; XG and HG solved the structure; NZ, MFH and HG 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.
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.
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.
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 (vol 72, pg 2468, 2004). Infection and Immunity 2004, 72, 6190-6190.
5. Gruber, T.M.; Gross, C.A. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol 2003, 57, 441-466.
6. Haugen, S.P.; Ross, W.; Manrique, M.; Gourse, R.L. Fine structure of the promoter-sigma region 1.2 interaction. Proceedings of the National Academy of Sciences of the United States of America 2008, 105, 3292-3297.
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.
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.
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.
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. The Journal of biological chemistry 2008, 283, 33455-33464.
11. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 2007, 372, 774-797.
12. Holm, L.; Sander, C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993, 233, 123-138.
13. Liu, B.; Zuo, Y.H.; Steitz, T.A. Structures of e-coli sigma(s)-transcription initiation complexes provide new insights into polymerase mechanism. Proceedings of the National Academy of Sciences of the United States of America 2016, 113, 4051-4056.
14. Tomsic, M.; Tsujikawa, L.; Panaghie, G.; Wang, Y.; Azok, J.; deHaseth, 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. The Journal of biological chemistry 2001, 276, 31891-31896.
15. Osterberg, S.; del Peso-Santos, T.; Shingler, V. Regulation of alternative sigma factor use. Annual Review of Microbiology, Vol 65 2011, 65, 37-55.
16. 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. Proceedings of the National Academy of Sciences of the United States of America 2013, 110, 15955-15960.
17. 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.
18. 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.

19. 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.
20. Otwinowski, Z.; Minor, W.; et al. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol 1997, 276, 307-326.
21. Murakami, K.S. X-ray crystal structure of escherichia coli rna polymerase sigma(70) holoenzyme. Journal of Biological Chemistry 2013, 288, 9126-9134.
22. 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.
23. 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.
24. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004, 60, 2126-2132.
25. 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.
26. 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 Research 2007, 35, W375-W383.
27. Corpet, F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 1988, 16, 10881-10890.
28. Gouet, P.; Courcelle, E.; Stuart, D.I.; Metoz, F. Espript: Analysis of multiple sequence alignments in postscript. Bioinformatics 1999, 15, 305-308.