# Identification and structural modeling of the chlamydial RNA polymerase omega subunit

**Running title:** Identification of *Chlamydia* RNA polymerase omega subunit

Andrew Cheng[1], Danny Wan[1,2], Arkaprabha Ghatak[1], Chengyuan Wang[3], Deyu Feng[3], Joseph D. Fondell[1], Richard H. Ebright[4,5], and Huizhou Fan[1*]

[1] Department of Pharmacology, Rutgers-Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA

[2] Graduate Program in Physiology and Integrative Biology, Rutgers School of Graduate Studies, Piscataway, NJ 08854, USA

[3] Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China

[4] Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA

[5] Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

* Correspondence: fanhu@rwjms.rutgers.edu

# ABSTRACT

Gene transcription in bacteria is carried out by the multisubunit RNA polymerase (RNAP), which is composed of a catalytic core enzyme and a promoter-recognizing σ factor. RNAP core enzyme comprises two α subunits, one β subunit, one β′ subunit, and one ω (omega) subunit. Across multiple bacterial taxa, the RNAP ω subunit plays critical roles in the assembly of RNAP core enzyme and in other cellular functions, including regulation of bacterial growth, stress response, and biofilm formation. However, for several intracellular bacterium, including the obligate intracellular bacterium *Chlamydia*, no RNAP ω subunit previously has been identified. Here, we report the identification of *Chlamydia trachomatis* hypothetical protein CTL0286 as the chlamydial RNAP ω ortholog, based on sequence, synteny, and AlphaFold and AlphaFold-Multimer three-dimensional-structure predictions. We conclude that CTL0286 functions as the previously missing chlamydial ω ortholog. Extensions of our analysis indicate that all obligate intracellular bacteria have ω orthologs.


# IMPORTANCE

Chlamydiae are common mammalian pathogens. Chlamydiae have a unique developmental cycle characterized with an infectious but nondividing elementary body (EB), which can temporarily survive outside host cells, and a noninfectious reticulate body (RB), which replicates only intracellularly. Chlamydial development inside host cells can be arrested during persistence in response to adverse environmental conditions. Transcription plays a central role in the progression of the chlamydial developmental cycle as well as entry into and recovery from persistence. The identification of the elusive ω subunit of chlamydial RNAP makes possible future study of its regulatory roles in gene expression during chlamydial growth, development,

and stress responses. This discovery also paves the way to prepare and study the intact chlamydial RNAP and its interactions with inhibitors *in vitro*.

# INTRODUCTION

RNA synthesis in bacteria is carried by a single RNA polymerase (RNAP). The bacterial RNAP is a multisubunit enzyme (1). In almost all bacteria, the catalytic core enzyme of the RNAP (RNAP core) is composed of two α subunits, one β subunit, one β' subunit, and one ω subunit (1, 2). Association of a σ factor to the core enzyme results in the formation of the RNAP holoenzyme (1). In the context of the holoenzyme, the σ factor is the primary determinant of promoter recognition and binding, and the RNAP core catalyzes the initiation and elongation of RNA synthesis using DNA as template (2-4).

The RNAP ω subunit, a protein of only about 10 kDa, initially was thought to be a contaminant in purified RNAP preparations (5-7). This view was prompted by the observation that ω-free RNAP preparations were active in transcription assays (8). However, the observation of increased transcription-initiation activity by an RNAP derivatives having ω fused to DNA-binding domains indicated ω was an integral component of RNAP (9). Further studies showed that ω is critical for the folding of the RNAP β' subunit and the for the assembly and stability of RNAP core enzyme (10-14). Studies using ω-deficient bacteria showed that ω is important for response to amino acid starvation, thermal and $CO_2$ acclimation, biofilm formation, and antibiotic production, and also affects growth under standard culture conditions (15-20). It was also shown that ω regulates the association of principal and alternative σ factors by the RNAP core enzyme and thus can affect promoter-recognition selectivity (21-23). Taken together, these and other studies suggest that ω serves as an important component of the bacterial RNAP holoenzyme and is required for numerous physiological functions [for review, see (24-26)].

ω has been found in all free-living bacteria and in some obligate intracellular bacteria (24, 25). ω also is present in some eukaryotic chloroplasts (27). An ortholog of ω, termed RpoK, is present in archaeal RNAP (28), and an ortholog of ω, termed RPB6, is present in eukaryotic RNAP I, II, and III (29).

Chlamydiae are intracellular bacteria that replicate only inside eukaryotic host cells (30, 31). Chlamydiae and *Chlamydia*-like organisms have been isolated from a wide range of hosts (32-46). Significantly, *Chlamydia trachomatis* is the number one sexually transmitted bacterial pathogen globally, and also is a major cause of preventable blindness in developing countries (47-49), and *C. pneumoniae* is a common respiratory pathogen (50-54). Several animal *Chlamydia* species are zoonotic pathogens (55-64). *Waddlia chondrophila* is one of several *Chlamydia*-like organisms, termed environmental chlamydiae, typically found in lower eukaryotes, such as amoebae, *but* can infect, and induce abortion in, vertebrates, including humans (65).

Chlamydiae are characterized by a unique developmental cycle consisting of two distinct cellular forms. The infectious but non-proliferative elementary body (EB) is capable of temporarily surviving in extracellular environments and invading host cells. Following invasion of host cells and entry into cytoplasmic vacuoles, EBs differentiate into proliferative reticulate bodies (RBs). Following multiple rounds of replication, RBs convert back into EBs, which then exit host cells (66-68). In addition to this "productive" chlamydial developmental cycle, under unfavorable environmental conditions (e.g., nutrient/mineral starvation, increased temperature, or exposure to inhibitory antibiotics, or cytokines), chlamydiae can enter into a "persistent" state characterized

by aberrant RBs inside infected cells, and, when environmental conditions improve, the aberrant RBs can exit the persistent state and resume production of EBs, (69-75).

Both the productive chlamydial developmental cycle and persistent infection are controlled by gene transcription (69, 71, 75-77). The chlamydial genome encodes three σ factors ($\sigma^{66}$, $\sigma^{28}$ and $\sigma^{54}$), as well as the α, β and β´ subunits of the core enzyme (78, 79). Surprisingly, it previously has not been possible to identify a candidate gene encoding the ω subunit in any chlamydial genome [e.g., (80-83)]. In principle, the chlamydial *rpoZ* gene may have been lost in the evolutionary process during which *Chlamydia* reduced its genome size to adapt to its unique developmental cycle. Alternatively, in principle, the chlamydial ω protein may have gone undetected due to low sequence homology with known bacterial and chloroplast ω factors.

Here, we report the identification of chlamydial ω, based on conserved amino-acid sequence, conserved synteny, and AlphaFold-predicted conserved three-dimensional structure and interactions. In addition, we also present an AlphaFold-Multimer model of the three-dimensional structure of a complex composed of the chlamydial RNAP β, β', and ω subunits. The identification of the previously elusive chlamydial ω sets the stage for investigation of its roles in regulation of gene expression during chlamydial growth, development, and stress responses. Our findings also set the stage to reconstitute the intact cRNAP from recombinant subunits *in vitro*, for future structural studies and for discovery and development of small-molecule inhibitors as possible anti-chlamydial drugs.

# METHODS

## BlastP analysis

Web-based BlastP was performed at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins using default settings (84). Multiple protein sequence alignment was performed with ClustalX2 on a Windows computer on PC (85) or Clustal Omega at https://www.ebi.ac.uk/Tools/msa/clustalo/ using default settings (86, 87).

## Three-dimensional-structure prediction

AlphaFold version 2.2.0 (88) was installed locally and run using the reduced database option with a maximum template date of November 1, 2021 and the multimer preset enabled for the cRNAP β'-CTL0286 and cRNAP β-β'-CTL0286 complex predictions. The multimer predictions were run with the default pre-trained AlphaFold-Multimer models (89), and the ranked 0 predictions (i.e., with lowest predicted local distance difference test [pLDDT] scores) were used for the figures for each complex. The CTL0286 monomer structure prediction was performed with the default monomer preset and the full database option and used the original CASP14 monomer models without ensembling. Model prediction and amber relaxation were performed for all predictions using a single NVIDIA Tesla V100 Volta GPU with 16GB of memory. Since the total sequence lengths significantly increase the space complexity, forced unified memory was enabled, and the XLA memory fraction environmental variable was set to 4.0 to avoid out of memory errors during runtime.

## Three-dimensional-structure similarity analysis

Structural homology search for AlphaFold model of CTL0286 was performed using the Dali server heuristic PDB Search option (90, 91) available at http://ekhidna2.biocenter.helsinki.fi/dali/. A PDB90 non-redundant subset at 90% sequence identity was used.

## Synteny analysis

CSBFinder-S (v0.6.3) (92) was used with the default settings to find the *gmk-rpoZ* synteny across 23,517 fully sequenced bacterial genomes downloaded from the NCBI genome database. DeepNOG (v1.2.3) (93) was run using the default setting to obtain the COG (Clusters of Orthologous Genes) ID for each gene. Strand information was obtained from the corresponding genomic.gff file for every genome downloaded. The *gmk-rpoZ* synteny was identified by finding COG0194 (*gmk*) and COG1758 (*rpoZ*) together within the CSBFinder-S output.

# RESULTS

## Identification of chlamydial ω: sequence similarity

Although ω has not been detected in Chlamydiae, ω had been detected in two other intracellular bacteria: *Rickettsia* and *Coxiella* (94, 95). Therefore, as a starting point to determine if chlamydiae encode an ω subunit, we performed BlastP analysis for chlamydial genomes using the amino-acid sequences of *R. rickettsii* ω and *C. burnettii* ω (94, 95) as queries. Using default parameters (84), the analysis did not detect sequence homolog to the *R. rickettsii* ω; in chlamydiae. However, the analysis did detect a possible sequence homolog of *C. burnettii* ω: Wcw_0707, a hypothetical protein encoded by the genome of the *Chlamydia*-like organism *W. chondrophila* (82) (Fig. 1A). Analysis of the sequence of Wcw_0707 revealed two features consistent with Wcw_0707 being an ω ortholog. First, Wcw_0707 is 107 amino acids long, similar in size to ω (~100 amino acids). Second, the Wcw_0707 *N*-terminal region (residues 7-62) exhibits strong sequence similarity to the *C. burnettii* ω (Fig. 1A) and *Escherichia coli* ω *N*-terminal regions (Fig. 1B), which are known to be responsible for binding to the RNAP β´ subunit and for facilitating the folding of β' (96). We hypothesized that Wcw_0707 may be the ω subunit in *W. chondrophila*.

Given our primarily interest in transcriptional regulation by the human sexually transmitted pathogen *C. trachomatis*, we next used Wcw_0707 as the query to search for a putative ω gene in the *C. trachomatis* genome. The search revealed a strong sequence similarity between the *N*-terminal region of hypothetical protein CTL0286 of *C. trachomatis* serovar L2 and the N-terminal region of Wcw_0707 of *W. chondrophila* (Fig. 1C). CTL0286 is a small protein of 100

159 amino acids , similar in length to previously reported RNAP ω subunits and similar in length to
160 Wcw_0707, (81). Notably, although CTL0286 exhibits only low overall sequence similarity to
161 other reported bacterial ω subunits, it contains a key conserved set of amino acids found in ω
162 subunits of a broad range of bacterial taxa (Fig. 1D). Additional BlastP analysis of CTL0286
163 identified CTL0286 orthologs in all vertebrate chlamydiae (Fig. 2). These findings support the
164 hypothesis that Wcw_0707 is the ω subunit in *W. chondrophila* and enable the hypothesis that
165 CTL0286 and its orthologs are ω subunits in *C. trachomatis* and other vertebrate chlamydiae.

166 **Identification of chlamydial ω: synteny**

167 Upon manual examination of *rpoZ* in 10 bacterial genomes, we noted that the *rpoZ* gene always
168 is located immediately downstream of the *gmk* gene, which encodes guanylate kinase (Table 1).
169 An *in silico* analysis identified *gmk-rpoZ* synteny in 18302 of 23517 fully-sequenced bacterial
170 genomes. The conservation of *gmk-rpoZ* synteny across a majority of bacteria taxa suggests that
171 there likely is an adaptive advantage to *gmk-rpoZ* synteny, although the character of the adaptive
172 advantage is not readily clear. Interestingly, in *W. chondrophila*, the *wcw_0707* gene is located
173 immediately downstream of the *gmk* gene, and, in all vertebrate chlamydiae species, the *ctl0286*
174 gene and its orthologs are also located immediately downstream of *gmk* (Table 1). This
175 conserved gene order provides further support for the hypothesis that CTL0286 and its orthologs
176 are chlamydial ω subunits.

177 **Identification of chlamydial ω: predicted three-dimensional structural similarity**

178 AlphaFold has recently become an indispensable resource for predicting the three-dimensional
179 structures of proteins and protein complexes (88, 89). We first used AlphaFold to predict three-

dimensional structure of CTL0286. In the resulting predicted structure for CTL0286, the N-terminal region (residues 1-58) contains three α helices (α1, residues 9-15 residues; α2, residues 19-36; and α3, residues 44-55) that correspond to three α-helices present in all structurally characterized ω subunits (26, 97, 98), and the C-terminal region (residues 58-100) are mostly disordered, similar to in structurally characterized ω subunits having lengths greater ~60 amino acids (26, 97, 98). Three-dimensional-structure similarity searches of the AlphaFold prediction for full-length CTL0286, performed on the DALI server (90, 91), identified bacterial ω subunits as the three top hits, with Z-scores of 3.8, 3.4, and 3.1, for RNAP ω subunits of *Clostridium difficile* (99), *Mycobacterium tuberculosis* (100), and *Bacillus subtilis* (101), respectively (Table 2). Three-dimensional-structure similarity searches of the AlphaFold prediction for the N-terminal region of CTL0286 (residues 1-62), performed on the DALI server (90, 91), identified bacterial ω subunits as the three top hits, with Z-scores of 5.2, 5.1, and 4.9 for RNAP ω subunits of *Escherichia coli* (102), *Mycobacterium tuberculosis* (103), and *Bacillus subtilis* (104), respectively (Table 2).

We next used AlphaFold-Multimer (89) to predict the three-dimensional structure of a complex of CTL0286 and the *C. trachomatis* RNAP β' subunit (Fig. 4A). The resulting predicted three-dimensional structure of CTL0286-β' was superimposable, with an rmsd of 2.2 Å for CTL0286 and an rmsd of 4.0 Å for *C. trachomatis* RNAP β'on a crystal structure of the ω-β' subcomplex of *E. coli* RNAP holoenzyme (PDB 6ALH) (97) (Fig. 4B). Significantly, the predicted three-dimensional structure of CTL0286-β' includes interactions that bridge the RNAP β'-subunit N- and C-termini (Fig. 4C) as observed in experimental structures of ω-containing RNAP and RNAP complexes (97, 105, 106), where they are believed to reduce configurational entropy of

partly folded and folded states of the nearly 1400-residue RNAP β' subunits, and thereby to facilitate RNAP assembly and enhance RNAP stability (10, 12, 14). We further used AlphaFold-Multimer to predict the three-dimensional structure of a heterotrimeric protein complex comprising CTL0286, *C. trachomatis* RNAP β', and *C. trachomatis* RNAP β (Fig. 5A). The resulting predicted three-dimensional structure of CTL0286-β' was superimposable, with rmsd of 2.2 Å for CTL0286 and 2.7 Å for *C. trachomatis* RNAP β, and β, on a crystal structure of the ω-β'-β subcomplex of *E. coli* RNAP holoenzyme (PDB 6ALH) (97) (Fig. 5B) and includes interactions that bridge the N- and C-termini of β' (Fig. 5C).

Taken together, these findings provide further support for our hypothesis that CTL0286 and its orthologs are *bona fide* chlamydial ω subunits.

## Identification of ω in other obligate intracellular bacteria

After successful identification of RNAP ω subunit in chlamydiae, we next determined if ω is present in other obligate intracellular bacterial taxa beside rickettsiae. NCBI searches identified annotated ω orthologs in the proteomes of *Anaplasma, Ehrlichia, Orientia, Wolbachia* and *Candidatus Midichloria*. Pre-generated AlphaFold structural models of *Anaplasma, Ehrlichia, Orentia,* and *Wolbachia* ω orthologs at www.uniprot.org (107) show three-dimensional strctural similarity to experimentally determined structures of bacterial ω subunits, indicating that the annotations likely are correct. No pre-generated AlphaFold structural model of the annotated *Candidatus Midichloria* ω ortholog is available at www.uniprot.org (107). However, generation of an AlphaFold structural model for the annotated C*andidatus Midichloria* ω ortholog (Fig. 6), followed by three-dimensional-structure similarity searches on the DALI server (90, 91)

identified bacterial ω subunits as the three top hits, with Z-scores of 9.3, 8.6, and 8.6 for RNAP ω subunits of *Pseudomonas Aeruginosa* (108), *Mycobacterium tuberculosis* (100), and *Xanthomonos oryzae* (109), respectively, indicating that the annotation likely is correct (Table 3). We conclude that *Analplasma, Ehrlichia, Orientia, Wolbachia*, and *Candidatus Midichloria* all possess RNAP ω subunits.

**Absence of ω in *Mycoplasma* and *Ureaplasma***

We next extended our RNAP ω subunit search in the facultative intracellular bacterium *Mycoplasma genitalium*, whose 580-kb genome is the smallest known bacterial genome (110). NCBI search failed to identify an annotated *rpoZ* in *M. genitalium*. Interestingly, our search also failed to identify an annotated *rpoZ* in other *Mycoplasma* species, even though most have genome sizes comparable to that of *Chlamydia*. Our search also failed to identify an annotated *rpoZ* in *Ureaplasma* (111), which is phylogenetically closely related to *Mycoplasma.* To verify the absence of ω subunits in these organisms, we first checked the gene immediately downstream of the *gmk* gene in *Mycoplasma* for possible sequence similarity to *rpoZ*, and we found none (110, 111). We next performed AlphaFold modeling for all 68 hypothetical proteins of *Mycoplasma pneumoniae* having sizes comparable to bacterial ω subunits (i.e., sizes of 40-150 amino acids) (112). AlphaFold predicted multi-α-helix folds for 26 of the 68 proteins. Three-dimensional-structure similarity searches of these 26 AlphaFold predictions, performed on the DALI server (90, 91), failed to identify structures of experimentally determined bacterial ω subunits as possible matches. We infer that *Mycoplasma* and *Ureaplasma* are unlikely to have RNAP ω subunits.

# Discussion

In this report, we present multiple lines of evidence for the existence of an RNAP ω subunit in chlamydiae. Although a lack of strong, continuous sequence homology previously had precluded the identification of a chlamydial ω, a multi-step BlastP analysis led to the identification of CTL0286 as candidate (Fig. 1). Like *rpoZ* in the super majority of bacteria, *ctl0286* is located immediately downstream of *gmk* (Table 2 and data not shown). AlphaFold-predicted three-dimensional structures of CTL0286 exhibit strong similarities to experimental three-dimensional structures of ω subunits for a broad range of bacterial taxa (29, 101, 109, 113, 114). AlphaFold-Multimer predicted three-dimensional structures of complexes of CTL0286, with *C. trachomatis* RNAP β' subunit, and of CTL0286 with *C. trachomatis* RNAP β' and β subunits, exhibit strong similarity to experimental three-dimensional structures of ω-β' and β'-β complexes [(Fig. 4, 5); (97, 98)]. The identification of CTL0286 as the *C. trachomatis* ω demonstrates the power of use of combinations of sequence-similarity analysis, synteny analysis, and AlphaFold and AlphaFold-Multimer analysis for identifying proteins "missing" from proteomes and for annotating functions of hypothetical proteins in proteomes.

Our extended analysis further showed that like *Chlamydia*, other obligate intracellular bacteria (i.e., *Rickettsia*, *Anaplasma, Ehrlichia, Orientia, Wolbachia* and *Candidatus Midichloria*) also encode ω orthologs (Fig. 6 and data not shown), but facultative intracellular bacteria *Mycoplasma* and *Ureaplasma* do not. Together with previous findings demonstrating the existence of ω orthologs in archaea and eukaryotes (27-29), these findings suggest that all living organisms from bacteria to humans have omega orthologs, likely with *Mycoplasma* and *Ureaplasma* as only exceptions.

ω plays roles in σ-RNAP core enzyme association (21-23) and thereby influences promoter-recognition selectivity (21-23). *Chlamydiae* possess a principal σ factor and two alternative σ factors (80, 81, 115). The principal σfactor, $\sigma^{66}$, is involved in transcription of most chlamydial genes throughout the developmental cycle; the alternative σ factors, $\sigma^{28}$ and $\sigma^{54}$, are required for expression of certain late genes (116-118). The different chlamydial σ factors also differentially affect response to stress conditions (71, 77). It would be equally interesting to investigate if and how the chlamydial ω regulates σ-RNAP core enzyme association in chlamydial developmental stages and in response to various stress condition.

In summary, we have identified the long-missing ω subunit of the cRNAP. As with most scientific studies, this discovery raises more questions than it answers. There is a need to determine whether the chlamydial ω plays solely a structural role in cRNAP assembly and stability, or whether it also functions in regulation of chlamydial growth, development, and stress response.

279 **ACKNOWLEDGEMENTS**

280 This work was supported by grants from the National Institutes of Health (AI071954 to HF and
281 GM041376 to RHE). We thank Yu Zhang and Liqiang Shen for bringing *gmk-rpoZ* synteny to
282 our attention and Jason Kaelber for helpful discussions.

283

## FIGURE LEGENDS

**Fig. 1. Identification of cRNAP ω candidate by BlastP and sequence alignment.** (A) BlastP-detected sequence homology between *Coxiella burnetii* RNAP ω subunit and wcw_0707, a hypothetical protein of the *Chlamydia*-like organism *Waddlia chondrophila*. (B) BlastP-detected sequence homology between *E. coli* RNAP ω and wcw_0707. (C) BlastP-detected sequence homology between wcw_0707 and CTL0286 of *Chlamydia trachomatis*. (D) ClustalX2-detected amino acids conserved in CTL0286 of *C. trachomatis*, wcw_0707 of *W. chondrophila*, and ωs of a variety of bacteria.

**Fig. 2. Sequence conservation among ω candidates in all vertebrate chlamydiae.** Alignment was performed using ClustalX2.

**Fig. 3. AlphaFold predictions for CTL0286.** (A) Superimposition of AlphaFold prediction for full-length CTL0286 (red) on experimental structures of *Clostridium difficile, Mycobacterium tuberculosis,* and *Bacillus subtilis* RNAP ω (blue, cyan, and gray, respectively). (B) Superimposition of AlphaFold prediction for N-terminal region (residues 1-62) of CTL0286 (red) on experimental structures of *Escherichia coli, Mycobacterium tuberculosis,* and *Bacillus subtilis* RNAP ω (blue, cyan, and gray, respectively).

**Fig. 4. AlphaFold-Multimer predictions for complex comprising CTL0286 and *C. trachomatis* RNAP β' subunit.** Superimposition of AlphaFold-Multimer prediction for CTL0286-β' (red for CTL0286; pink for β') on experimental structure of *E. coli* RNAP (PDB 6ALH; black for ω; light gray for β').

**Fig. 5. AlphaFold-Multimer predictions for complex comprising CTL0286 and *C. trachomatis* RNAP β' subunit, and β subunit.** Superimposition of AlphaFold-Multimer prediction for CTL0286-β'-β (red for CTL0286; pink for β'; cyan for β) on experimental structure of *E. coli* RNAP (PDB 6ALH; black for ω; light gray for β'; dark gray for β).

**Fig. 6. AlphaFold predictions for annotated ω of *Candidatus Midichloria* RNAP ω.** Superimposition of AlphaFold prediction for Candidatus Midichloria RNAP ω (red) on experimental structures of *Pseudomonas aeruginosa, M. tuberculosis,* and *Xanthomonas oryzae* RNAP ω (blue, cyan, and gray, respectively).

312

313 **Table 1. Conserved *gmk-rpoZ* linkage in bacterial genomes.**

| **Bacterium** | **Gram-stain** | **Upstream gene** | ***rpoZ* or equivalence** | **Downstream gene** |
|---|---|---|---|---|
| *Bacillus anthracis* | Positive | *gmk* | *rpoZ* | *coaBC* |
| *Clostridium difficile* | Positive | *gmk* | *rpoZ* | *coaBC* |
| *Lactobacillus acidophilus* | Positive | *gmk* | *rpoZ* | *priA* |
| *Staphylococcus epidermidis* | Positive | *gmk* | *rpoZ* | *SE0887* |
| *Coxiella burnetii* | Negative | *gmk* | *rpoZ* | *spoT* |
| *Escherichia coli* | Negative | *gmk* | *rpoZ* | *spoT* |
| *Haemohilus influenzae* | Negative | *gmk* | *rpoZ* | *spoT* |
| *Vibrio cholerae* | Negative | *gmk* | *rpoZ* | *spoT* |
| *Gardnerella vaginalis* | Variable | *gmk* | *rpoZ* | *dfp* |
| *Mycobacterium tuberculosis* | Variable | *gmk* | *rpoZ* | *metK* |
| ***Waddlia chondrophila*** | Negative | ***gmk*** | ***wcw_0707*** | *wcw_0708* |
| ***Chlamydia trachomatis*** | Negative | ***gmk*** | ***ctl0286*** | *metG* |

314
315

315 **Table 2. Proteins with structural homology to AlphaFold models of full-length CTL0286**

316 **(Fl-CTL0286) or N-terminus of CTL0286 (1-62).** Z-score is an optimized similarity score

317 defined as the sum of equivalent residue-wise C α -C α distances among two proteins.

318 Abbreviation: RMSD, Root-mean-square deviation of atomic positions.

| Model | Rank # | PDB structure (reference) | Protein | Bacterium | Z-value | RMSD |
|---|---|---|---|---|---|---|
| Fl-CTL0286 | 1 | 7L7B (99) | RNAP ω | *Clostridium difficile* | 3.8 | 3.9 |
| | 2 | 6BZO (100) | RNAP ω | *Mycobacterium tuberculosis* | 3.4 | 3.1 |
| | 3 | 7CKQ (101) | RNAP ω | *Bacillus subtilis* | 3.1 | 3.0 |
| | | | | | | |
| CTL0286 (1-62) | 1 | 5TJG (102) | RNAP ω | *Escherichia coli* | 5.2 | 2.1 |
| | 2 | 6KOP (103) | RNAP ω | *Mycobacterium tuberculosis* | 5.1 | 3.1 |
| | 3 | 7F75 (104) | RNAP ω | *Bacillus subtilis* | 4.9 | 2.8 |

319

320

320 **Table 3. Proteins with structural homology to AlphaFold model of annotated *Candidatus***

321 ***Midichloria* RNAP ω.**

322

| Rank # | PDB structure (reference) | Protein | Bacterium | Z-value | RMSD |
|---|---|---|---|---|---|
| 1 | 7XL3 (108) | RNAP ω | *Pseudomonas aeruginosa* | 9.3 | 2.7 |
| 2 | 7L7B (100) | RNAP ω | *Mycobacterium tuberculosis* | 8.6 | 3.0 |
| 3 | 6J9E (109) | RNAP ω | *Xanthomonos oryzae* | 8.6 | 3.7 |

323
324

## REFERENCES


1. Murakami KS, Darst SA. 2003. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol 13:31-9.
2. Darst SA. 2001. Bacterial RNA polymerase. Current opinion in structural biology 11:155-62.
3. Feklistov A, Sharon BD, Darst SA, Gross CA. 2014. Bacterial sigma factors: a historical, structural, and genomic perspective. Annu Rev Microbiol 68:357-76.
4. Mooney RA, Darst SA, Landick R. 2005. Sigma and RNA polymerase: an on-again, off-again relationship? Molecular cell 20:335-45.
5. Burgess RR. 1969. A new method for the large scale purification of Escherichia coli deoxyribonucleic acid-dependent ribonucleic acid polymerase. J Biol Chem 244:6160-7.
6. Berg D, Barrett K, Chamberlin M. 1971. [43] Purification of two forms of Escherichia coli RNA polymerase and of sigma component, p 506-519, Methods in Enzymology, vol 21. Academic Press.
7. Nüsslein C, Heyden B. 1972. Chromatography of RNA polymerase from escherichia coli on single stranded DNA-agarose columns. Biochemical and Biophysical Research Communications 47:282-289.
8. Heil A, Zillig W. 1970. Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. FEBS Letters 11:165-168.
9. Dove SL, Hochschild A. 1998. Conversion of the omega subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target. Genes Dev 12:745-54.
10. Ghosh P, Ramakrishnan C, Chatterji D. 2003. Inter-subunit recognition and manifestation of segmental mobility in Escherichia coli RNA polymerase: a case study with omega-beta' interaction. Biophys Chem 103:223-37.
11. Mukherjee K, Nagai H, Shimamoto N, Chatterji D. 1999. GroEL is involved in activation of Escherichia coli RNA polymerase devoid of the omega subunit in vivo. Eur J Biochem 266:228-35.
12. Bhowmik D, Bhardwaj N, Chatterji D. 2017. Influence of flexible "omega" on the activity of E. coli RNA polymerase: A thermodynamic analysis. Biophys J 112:901-910.
13. Mukherjee K, Chatterji D. 1997. Studies on the omega subunit of Escherichia coli RNA polymerase--its role in the recovery of denatured enzyme activity. Eur J Biochem 247:884-9.
14. Ghosh P, Ishihama A, Chatterji D. 2001. Escherichia coli RNA polymerase subunit omega and its N-terminal domain bind full-length beta' to facilitate incorporation into the alpha2beta subassembly. Eur J Biochem 268:4621-7.
15. Mathew R, Mukherjee R, Balachandar R, Chatterji D. 2006. Deletion of the rpoZ gene, encoding the omega subunit of RNA polymerase, results in pleiotropic surface-related phenotypes in Mycobacterium smegmatis. Microbiology 152:1741-50.
16. Chatterji D, Ogawa Y, Shimada T, Ishihama A. 2007. The role of the omega subunit of RNA polymerase in expression of the relA gene in Escherichia coli. FEMS Microbiol Lett 267:51-5.

17. Kurkela J, Hakkila K, Antal T, Tyystjarvi T. 2017. Acclimation to High CO2 Requires the omega Subunit of the RNA Polymerase in Synechocystis. Plant Physiol 174:172-184.
18. Gunnelius L, Kurkela J, Hakkila K, Koskinen S, Parikainen M, Tyystjarvi T. 2014. The omega subunit of RNA polymerase is essential for thermal acclimation of the cyanobacterium Synechocystis sp. PCC 6803. PLoS One 9:e112599.
19. Santos-Beneit F, Barriuso-Iglesias M, Fernandez-Martinez LT, Martinez-Castro M, Sola-Landa A, Rodriguez-Garcia A, Martin JF. 2011. The RNA polymerase omega factor RpoZ is regulated by PhoP and has an important role in antibiotic biosynthesis and morphological differentiation in Streptomyces coelicolor. Appl Environ Microbiol 77:7586-94.
20. Kojima I, Kasuga K, Kobayashi M, Fukasawa A, Mizuno S, Arisawa A, Akagawa H. 2002. The rpoZ gene, encoding the RNA polymerase omega subunit, is required for antibiotic production and morphological differentiation in Streptomyces kasugaensis. J Bacteriol 184:6417-23.
21. Gunnelius L, Hakkila K, Kurkela J, Wada H, Tyystjarvi E, Tyystjarvi T. 2014. The omega subunit of the RNA polymerase core directs transcription efficiency in cyanobacteria. Nucleic Acids Res 42:4606-14.
22. Weiss A, Moore BD, Tremblay MH, Chaput D, Kremer A, Shaw LN. 2017. The omega subunit governs rna polymerase stability and transcriptional specificity in Staphylococcus aureus. J Bacteriol 199.
23. Igarashi K, Fujita N, Ishihama A. 1989. Promoter selectivity of Escherichia coli RNA polymerase: omega factor is responsible for the ppGpp sensitivity. Nucleic Acids Res 17:8755-65.
24. Patel UR, Gautam S, Chatterji D. 2020. Validation of Omega Subunit of RNA Polymerase as a Functional Entity. Biomolecules 10.
25. Mathew R, Chatterji D. 2006. The evolving story of the omega subunit of bacterial RNA polymerase. Trends Microbiol 14:450-5.
26. Kurkela J, Fredman J, Salminen TA, Tyystjarvi T. 2021. Revealing secrets of the enigmatic omega subunit of bacterial RNA polymerase. Mol Microbiol 115:1-11.
27. Douglas SE, Penny SL. 1999. The plastid genome of the cryptophyte alga, Guillardia theta: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236-44.
28. Magill CP, Jackson SP, Bell SD. 2001. Identification of a conserved archaeal RNA polymerase subunit contacted by the basal transcription factor TFB. J Biol Chem 276:46693-6.
29. Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K. 2001. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A 98:892-7.
30. Everett KD, Bush RM, Andersen AA. 1999. Emended description of the order Chlamydiales, proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards for the identification of organisms. Int J Syst Bacteriol 49 Pt 2:415-40.
31. Schachter J, Stephens RS, Timms P, Kuo C, Bavoil PM, Birkelund S, Boman J, Caldwell H, Campbell LA, Chernesky M, Christiansen G, Clarke IN, Gaydos C, Grayston JT,

412 Hackstadt T, Hsia R, Kaltenboeck B, Leinonnen M, Ocjius D, McClarty G, Orfila J,
413 Peeling R, Puolakkainen M, Quinn TC, Rank RG, Raulston J, Ridgeway GL, Saikku P,
414 Stamm WE, Taylor-Robinson DT, Wang SP, Wyrick PB. 2001. Radical changes to
415 chlamydial taxonomy are not necessary just yet. Int J Syst Evol Microbiol 51:249, 251-3.
416 32. Seth-Smith HMB, Buso LS, Livingstone M, Sait M, Harris SR, Aitchison KD, Vretou E,
417 Siarkou VI, Laroucau K, Sachse K, Longbottom D, Thomson NR. 2017. European
418 Chlamydia abortus livestock isolate genomes reveal unusual stability and limited
419 diversity, reflected in geographical signatures. BMC Genomics 18:344.
420 33. Salinas J, Ortega N, Borge C, Rangel MJ, Carbonero A, Perea A, Caro MR. 2012.
421 Abortion associated with Chlamydia abortus in extensively reared Iberian sows. Vet J
422 194:133-4.
423 34. Oseikria M, Pellerin JL, Rodolakis A, Vorimore F, Laroucau K, Bruyas JF, Roux C,
424 Michaud S, Larrat M, Fieni F. 2016. Can Chlamydia abortus be transmitted by embryo
425 transfer in goats? Theriogenology 86:1482-1488.
426 35. Li Z, Cao X, Fu B, Chao Y, Cai J, Zhou J. 2015. Identification and characterization of
427 Chlamydia abortus isolates from yaks in Qinghai, China. Biomed Res Int 2015:658519.
428 36. Li Z, Cai J, Cao X, Lou Z, Chao Y, Kan W, Zhou J. 2017. Whole-Genome Sequence of
429 Chlamydia abortus Strain GN6 Isolated from Aborted Yak Fetus. Genome Announc 5.
430 37. Appino S, Vincenti L, Rota A, Pellegrini S, Chieppa MN, Cadoni V, Pregel P. 2015.
431 Chlamydia abortus in Cows Oviducts, Occasional Event or Causal Connection? Reprod
432 Domest Anim 50:526-8.
433 38. Li J, Guo W, Kaltenboeck B, Sachse K, Yang Y, Lu G, Zhang J, Luan L, You J, Huang
434 K, Qiu H, Wang Y, Li M, Yang Z, Wang C. 2016. *Chlamydia pecorum* is the endemic
435 intestinal species in cattle while *C. gallinacea, C. psittaci* and *C. pneumoniae* associate
436 with sporadic systemic infection. Veterinary Microbiology 193:93-99.
437 39. Sait M, Livingstone M, Clark EM, Wheelhouse N, Spalding L, Markey B, Magnino S,
438 Lainson FA, Myers GS, Longbottom D. 2014. Genome sequencing and comparative
439 analysis of three Chlamydia pecorum strains associated with different pathogenic
440 outcomes. BMC Genomics 15:23.
441 40. Jelocnik M, Bachmann NL, Kaltenboeck B, Waugh C, Woolford L, Speight KN, Gillett
442 A, Higgins DP, Flanagan C, Myers GS, Timms P, Polkinghorne A. 2015. Genetic
443 diversity in the plasticity zone and the presence of the chlamydial plasmid differentiates
444 Chlamydia pecorum strains from pigs, sheep, cattle, and koalas. BMC Genomics 16:893.
445 41. Bachmann NL, Fraser TA, Bertelli C, Jelocnik M, Gillett A, Funnell O, Flanagan C,
446 Myers GS, Timms P, Polkinghorne A. 2014. Comparative genomics of koala, cattle and
447 sheep strains of Chlamydia pecorum. BMC Genomics 15:667.
448 42. Jelocnik M, Walker E, Pannekoek Y, Ellem J, Timms P, Polkinghorne A. 2014.
449 Evaluation of the relationship between Chlamydia pecorum sequence types and disease
450 using a species-specific multi-locus sequence typing scheme (MLST). Vet Microbiol
451 doi:10.1016/j.vetmic.2014.08.018.
452 43. Wu SM, Huang SY, Xu MJ, Zhou DH, Song HQ, Zhu XQ. 2013. Chlamydia felis
453 exposure in companion dogs and cats in Lanzhou, China: a public health concern. BMC
454 Vet Res 9:104.
455 44. Schulz C, Hartmann K, Mueller RS, Helps C, Schulz BS. 2015. Sampling sites for
456 detection of feline herpesvirus-1, feline calicivirus and Chlamydia felis in cats with feline
457 upper respiratory tract disease. J Feline Med Surg 17:1012-9.

458 45. Hartley JC, Stevenson S, Robinson AJ, Littlewood JD, Carder C, Cartledge J, Clark C,
459 Ridgway GL. 2001. Conjunctivitis due to Chlamydophila felis (Chlamydia psittaci feline
460 pneumonitis agent) acquired from a cat: case report with molecular characterization of
461 isolates from the patient and cat. J Infect 43:7-11.
462 46. Taylor-Brown A, Timms P, Polkinghorne A, Vaughan L, Greub G. 2014. Twenty years
463 of research into Chlamydia-like organisms: a revolution in our understanding of the
464 biology and pathogenicity of members of the phylum Chlamydiae. Pathogens and
465 Disease 73:1-15.
466 47. WHO. 2016. Global health sector strategy on sexually transmitted infections 2016–2021.
467 48. CDC. 2022. Sexually Transmitted Disease Surveillance 2020.
468 49. Taylor HR, Burton MJ, Haddad D, West S, Wright H. 2014. Trachoma. Lancet 384:2142-
469 52.
470 50. Wu JS, Lin JC, Chang FY. 2000. Chlamydia pneumoniae infection in community-
471 acquired pneumonia in Taiwan. J Microbiol Immunol Infect 33:34-8.
472 51. Clemmons NS, Jordan NN, Brown AD, Kough EM, Pacha LA, Varner SM, Hawksworth
473 AW, Myers CA, Gaydos JC. 2019. Outbreak of Chlamydia pneumoniae Infections and
474 X-ray-Confirmed Pneumonia in Army Trainees at Fort Leonard Wood, Missouri, 2014.
475 Mil Med 184:e196-e199.
476 52. Schmidt SM, Muller CE, Mahner B, Wiersbitzky SK. 2002. Prevalence, rate of
477 persistence and respiratory tract symptoms of *Chlamydia pneumoniae* infection in 1211
478 kindergarten and school age children. Pediatr Infect Dis J 21:758-62.
479 53. Yamamoto H, Watanabe T, Miyazaki A, Katagiri T, Idei T, Iguchi T, Mimura M,
480 Kamijima K. 2005. High prevalence of Chlamydia pneumoniae antibodies and increased
481 high-sensitive C-reactive protein in patients with vascular dementia. J Am Geriatr Soc
482 53:583-9.
483 54. Thom DH, Grayston JT, Campbell LA, Kuo CC, Diwan VK, Wang SP. 1994.
484 Respiratory infection with Chlamydia pneumoniae in middle-aged and older adult
485 outpatients. Eur J Clin Microbiol Infect Dis 13:785-92.
486 55. Joseph SJ, Marti H, Didelot X, Read TD, Dean D. 2016. Tetracycline Selective Pressure
487 and Homologous Recombination Shape the Evolution of Chlamydia suis: A Recently
488 Identified Zoonotic Pathogen. Genome Biol Evol 8:2613-23.
489 56. Hulin V, Oger S, Vorimore F, Aaziz R, de Barbeyrac B, Berruchon J, Sachse K,
490 Laroucau K. 2015. Host preference and zoonotic potential of Chlamydia psittaci and C.
491 gallinacea in poultry. Pathog Dis 73:1-11.
492 57. De Puysseleyr K, De Puysseleyr L, Dhondt H, Geens T, Braeckman L, Morre SA, Cox E,
493 Vanrompay D. 2014. Evaluation of the presence and zoonotic transmission of Chlamydia
494 suis in a pig slaughterhouse. BMC Infect Dis 14:560.
495 58. Branley J, Bachmann NL, Jelocnik M, Myers GS, Polkinghorne A. 2016. Australian
496 human and parrot Chlamydia psittaci strains cluster within the highly virulent 6BC clade
497 of this important zoonotic pathogen. Sci Rep 6:30019.
498 59. Zezekalo VK, Kulynych SM, Polishchuk Acapital A C, Kone MS, Avramenko Ncapital
499 O C, Vakulenko YV, Chyzhanska NV. 2020. Prevalence of chlamydia-related organisms
500 with zoonotic potential in farms of the poltava region. Wiad Lek 73:1169-1172.
501 60. Turin L, Surini S, Wheelhouse N, Rocchi MS. 2022. Recent advances and public health
502 implications for environmental exposure to Chlamydia abortus: from enzootic to zoonotic
503 disease. Vet Res 53:37.

61. Ramakers BP, Heijne M, Lie N, Le TN, van Vliet M, Claessen VPJ, Tolsma PJP, De Rosa M, Roest HIJ, Vanrompay D, Heddema ER, Schneeberger P, Hermans MHA. 2017. Zoonotic Chlamydia caviae Presenting as Community-Acquired Pneumonia. N Engl J Med 377:992-994.
62. Li L, Luther M, Macklin K, Pugh D, Li J, Zhang J, Roberts J, Kaltenboeck B, Wang C. 2017. Chlamydia gallinacea: a widespread emerging Chlamydia agent with zoonotic potential in backyard poultry. Epidemiol Infect 145:2701-2703.
63. Dickx V, Vanrompay D. 2011. Zoonotic transmission of Chlamydia psittaci in a chicken and turkey hatchery. J Med Microbiol 60:775-779.
64. Ravichandran K, Anbazhagan S, Karthik K, Angappan M, Dhayananth B. 2021. A comprehensive review on avian chlamydiosis: a neglected zoonotic disease. Tropical animal health and production 53:414-414.
65. Baud D, Goy G, Osterheld MC, Borel N, Vial Y, Pospischil A, Greub G. 2011. Waddlia chondrophila: From Bovine Abortion to Human Miscarriage. Clin Infect Dis 52:1469-71.
66. Belland RJ, Zhong G, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL, Caldwell HD. 2003. Genomic transcriptional profiling of the developmental cycle of Chlamydia trachomatis. Proc Natl Acad Sci U S A 100:8478-83.
67. Hybiske K, Stephens RS. 2007. Mechanisms of *Chlamydia trachomatis* entry into nonphagocytic cells. Infect Immun 75:3925-3934.
68. Hybiske K, Stephens RS. 2007. Mechanisms of host cell exit by the intracellular bacterium *Chlamydia*. Proc Natl Acad Sci USA 104:11430-11435.
69. Belland RJ, Nelson DE, Virok D, Crane DD, Hogan D, Sturdevant D, Beatty WL, Caldwell HD. 2003. Transcriptome analysis of chlamydial growth during IFN-γ-mediated persistence and reactivation. Proceedings of the National Academy of Sciences 100:15971-15976.
70. Panzetta ME, Valdivia RH, Saka HA. 2018. Chlamydia Persistence: A Survival Strategy to Evade Antimicrobial Effects in-vitro and in-vivo. Frontiers in Microbiology 9.
71. Shima K, Kaufhold I, Eder T, Kading N, Schmidt N, Ogunsulire IM, Deenen R, Kohrer K, Friedrich D, Isay SE, Grebien F, Klinger M, Richer BC, Gunther UL, Deepe GS, Jr., Rattei T, Rupp J. 2021. Regulation of the mitochondrion-fatty acid axis for the metabolic reprogramming of *Chlamydia trachomatis* during treatment with beta-lactam antimicrobials. mBio 12.
72. Huston WM, Theodoropoulos C, Mathews SA, Timms P. 2008. Chlamydia trachomatis responds to heat shock, penicillin induced persistence, and IFN-gamma persistence by altering levels of the extracytoplasmic stress response protease HtrA. BMC Microbiol 8:190.
73. Slade JA, Brockett M, Singh R, Liechti GW, Maurelli AT. 2019. Fosmidomycin, an inhibitor of isoprenoid synthesis, induces persistence in Chlamydia by inhibiting peptidoglycan assembly. PLOS Pathogens 15:e1008078.
74. Wyrick PB. 2010. Chlamydia trachomatis Persistence In Vitro: An Overview. Journal of Infectious Diseases 201:S88-S95.
75. Brinkworth AJ, Wildung MR, Carabeo RA. 2018. Genomewide Transcriptional Responses of Iron-Starved Chlamydia trachomatis Reveal Prioritization of Metabolic Precursor Synthesis over Protein Translation. mSystems 3:e00184-17.

548 76. Belland RJ, Zhong G, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL,
549 Caldwell HD. 2003. Genomic transcriptional profiling of the developmental cycle of
550 *Chlamydia trachomatis*. Proc Natl Acad Sci USA 100:8478-83.
551 77. Huang Y, Wurihan W, Lu B, Zou Y, Wang Y, Weldon K, Fondell JD, Lai Z, Wu X, Fan
552 H. 2021. Robust heat shock response in Chlamydia lacking a typical heat shock sigma
553 factor. Front Microbiol 12:812448.
554 78. Engel JN, Pollack J, Malik F, Ganem D. 1990. Cloning and characterization of RNA
555 polymerase core subunits of Chlamydia trachomatis by using the polymerase chain
556 reaction. J Bacteriol 172:5732-41.
557 79. Tan M, Klein R, Grant R, Ganem D, Engel J. 1993. Cloning and characterization of the
558 RNA polymerase alpha-subunit operon of Chlamydia trachomatis. J Bacteriol 175:7150-
559 9.
560 80. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, Mitchell W, Olinger
561 L, Tatusov RL, Zhao Q, Koonin EV, Davis RW. 1998. Genome sequence of an obligate
562 intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282:754-9.
563 81. Thomson NR, Holden MTG, Carder C, Lennard N, Lockey SJ, Marsh P, Skipp P,
564 O'Connor CD, Goodhead I, Norbertzcak H, Harris B, Ormond D, Rance R, Quail MA,
565 Parkhill J, Stephens RS, Clarke IN. 2008. Chlamydia trachomatis: Genome sequence
566 analysis of lymphogranuloma venereum isolates. Genome Res 18:161-171.
567 82. Bertelli C, Collyn F, Croxatto A, Rückert C, Polkinghorne A, Kebbi-Beghdadi C,
568 Goesmann A, Vaughan L, Greub G. 2010. The *Waddlia* Genome: A Window into
569 Chlamydial Biology. PLoS ONE 5:e10890.
570 83. Greub G, Kebbi-Beghdadi C, Bertelli C, Collyn F, Riederer BM, Yersin C, Croxatto A,
571 Raoult D. 2009. High Throughput Sequencing and Proteomics to Identify Immunogenic
572 Proteins of a New Pathogen: The Dirty Genome Approach. PLoS ONE 4:e8423.
573 84. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu YK.
574 2005. Protein database searches using compositionally adjusted substitution matrices.
575 FEBS J 272:5101-9.
576 85. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,
577 Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.
578 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-8.
579 86. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003.
580 Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res
581 31:3497-500.
582 87. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H,
583 Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of
584 high-quality protein multiple sequence alignments using Clustal Omega. Molecular
585 systems biology. 7(539. doi:10.1038/msb.2011.75.
586 88. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool
587 K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ,
588 Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D,
589 Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver
590 D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate
591 protein structure prediction with AlphaFold. Nature 596:583-589.
592 89. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R,
593 Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko

A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D. 2022. Protein complex prediction with AlphaFold-Multimer. bioRxiv doi:10.1101/2021.10.04.463034:2021.10.04.463034.

90. Holm L. 2020. Using Dali for Protein Structure Comparison. Methods Mol Biol 2112:29-42.
91. Holm L. 2020. DALI and the persistence of protein shape. Protein Science 29:128-140.
92. Svetlitsky D, Dagan T, Chalifa-Caspi V, Ziv-Ukelson M. 2018. CSBFinder: discovery of colinear syntenic blocks across thousands of prokaryotic genomes. Bioinformatics 35:1634-1643.
93. Feldbauer R, Gosch L, Lüftinger L, Hyden P, Flexer A, Rattei T. 2020. DeepNOG: fast and accurate protein orthologous group assignment. Bioinformatics 36:5304-5312.
94. Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, Ward NL, Tettelin H, Davidsen TM, Beanan MJ, Deboy RT, Daugherty SC, Brinkac LM, Madupu R, Dodson RJ, Khouri HM, Lee KH, Carty HA, Scanlan D, Heinzen RA, Thompson HA, Samuel JE, Fraser CM, Heidelberg JF. 2003. Complete genome sequence of the Q-fever pathogen Coxiella burnetii. Proceedings of the National Academy of Sciences 100:5455-5460.
95. McLeod MP, Qin X, Karpathy SE, Gioia J, Highlander SK, Fox GE, McNeill TZ, Jiang H, Muzny D, Jacob LS, Hawes AC, Sodergren E, Gill R, Hume J, Morgan M, Fan G, Amin AG, Gibbs RA, Hong C, Yu X-j, Walker DH, Weinstock GM. 2004. Complete Genome Sequence of Rickettsia typhi and Comparison with Sequences of Other Rickettsiae. J Bacteriol 186:5842-5855.
96. Ghosh P, Ishihama A, Chatterji D. 2001. Escherichia coli RNA polymerase subunit ω and its N-terminal domain bind full-length β′ to facilitate incorporation into the α2β subassembly. Eur J Biochem 268:4621-4627.
97. Kang JY, Olinares PDB, Chen J, Campbell EA, Mustaev A, Chait BT, Gottesman ME, Darst SA. 2017. Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex. eLife 6:e25478.
98. Qayyum MZ, Molodtsov V, Renda A, Murakami KS. 2021. Structural basis of RNA polymerase recycling by the Swi2/Snf2 family of ATPase RapA in Escherichia coli. J Biol Chem 297:101404.
99. Anonymous. 2020. 7L7B: Clostridioides difficile RNAP with fidaxomicin, Online.
100. Boyaci H, Chen J, Lilic M, Palka M, Mooney RA, Landick R, Darst SA, Campbell EA. 2018. Fidaxomicin jams Mycobacterium tuberculosis RNA polymerase motions needed for initiation via RbpA contacts. Elife 7.
101. Fang C, Li L, Zhao Y, Wu X, Philips SJ, You L, Zhong M, Shi X, O'Halloran TV, Li Q, Zhang Y. 2020. The bacterial multidrug resistance regulator BmrR distorts promoter DNA to activate transcription. Nat Commun 11:6284.
102. Feklistov A, Bae B, Hauver J, Lass-Napiorkowska A, Kalesse M, Glaus F, Altmann KH, Heyduk T, Landick R, Darst SA. 2017. RNA polymerase motions during promoter melting. Science 356:863-866.
103. Li L, Molodtsov V, Lin W, Ebright RH, Zhang Y. 2020. RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. Proc Natl Acad Sci U S A 117:5801-5809.

104. Shi J, Li F, Wen A, Yu L, Wang L, Wang F, Jin Y, Jin S, Feng Y, Lin W. 2021. Structural basis of transcription activation by the global regulator Spx. Nucleic Acids Res 49:10756-10769.
105. Lin W, Mandal S, Degen D, Liu Y, Ebright YW, Li S, Feng Y, Zhang Y, Mandal S, Jiang Y, Liu S, Gigliotti M, Talaue M, Connell N, Das K, Arnold E, Ebright RH. 2017. Structural Basis of Mycobacterium tuberculosis Transcription and Transcription Inhibition. Mol Cell 66:169-179 e8.
106. Wang C, Molodtsov V, Firlar E, Kaelber JT, Blaha G, Su M, Ebright RH. 2020. Structural basis of transcription-translation coupling. Science 369:1359-1365.
107. Consortium TU. 2020. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research 49:D480-D489.
108. He D, You L, Wu X, Shi J, Wen A, Yan Z, Mu W, Fang C, Feng Y, Zhang Y. 2022. Pseudomonas aeruginosa SutA wedges RNAP lobe domain open to facilitate promoter DNA unwinding. Nat Commun 13:4204.
109. You L, Shi J, Shen L, Li L, Fang C, Yu C, Cheng W, Feng Y, Zhang Y. 2019. Structural basis for transcription antitermination at bacterial intrinsic terminator. Nat Commun 10:3048.
110. Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, 3rd, Smith HO, Venter JC. 2006. Essential genes of a minimal bacterium. Proc Natl Acad Sci U S A 103:425-30.
111. Marques LM, Guimarães AM, Martins HB, Rezende IS, Barbosa MS, Campos GB, do Nascimento NC, Dos Santos AP, Amorim AT, Santos VM, Messick JB, Timenetsky J. 2015. Genome Sequence of Ureaplasma diversum Strain ATCC 49782. Genome Announc 3.
112. Anonymous. 2021. Mycoplasma pneumoniae strain NCTC10119 chromosome 1, December 16, 2021 ed. NCBI.
113. Abdelkareem M, Saint-André C, Takacs M, Papai G, Crucifix C, Guo X, Ortiz J, Weixlbaumer A. 2019. Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. Mol Cell 75:298-309.e4.
114. Cao X, Boyaci H, Chen J, Bao Y, Landick R, Campbell EA. 2022. Basis of narrow-spectrum activity of fidaxomicin on Clostridioides difficile. Nature 604:541-545.
115. Read TD, Brunham RC, Shen C, Gill SR, Heidelberg JF, White O, Hickey EK, Peterson J, Utterback T, Berry K, Bass S, Linher K, Weidman J, Khouri H, Craven B, Bowman C, Dodson R, Gwinn M, Nelson W, DeBoy R, Kolonay J, McClarty G, Salzberg SL, Eisen J, Fraser CM. 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucl Acids Res 28:1397-1406.
116. Soules KR, LaBrie SD, May BH, Hefty PS. 2020. Sigma 54-regulated transcription is associated with membrane reorganization and type III secretion effectors during conversion to infectious forms of *Chlamydia trachomatis*. mBio 11.
117. Hiu Yin Yu H, Tan M. 2003. Sigma28 RNA polymerase regulates hctB, a late developmental gene in *Chlamydia*. Mol Microbiol 50:577-84.
118. Shen L, Feng X, Yuan Y, Luo X, Hatch TP, Hughes KT, Liu JS, Zhang YX. 2006. Selective promoter recognition by chlamydial sigma28 holoenzyme. J Bacteriol 188:7364-77.

**A**

```
                                4                                                   59
Coxiella burnetii RNAPω         VTVEDCLEHVENRFDLVLKAAKRAHILELGGAEPMVPRDNDKPAVLALREIAAGYD
                                +T E   ++  N+FDLV  A K A  +   G EP V  + + PA+L L EI  G D
Waddlia chon. wcw_0707          LTNEKISKNFNNQFDLVNYAIKLAANMIQTGREPRVKMNTENPALLILEEIIEGKD
                                7                                                   62
```

**B**

```
                                3                                                   57
E. coli RNAPω                   RVTVQDAVEKIGNRFDLVLVAARRARQMQVGGKDPLVPEENDKTTVIALREIEEG
                                 +T +   +   N+FDLV  A + A  M   G++P V    +   ++ L EI EG
Waddlia chon. wcw_0707          HLTNEKISKNFNNQFDLVNYAIKLAANMIQTGREPRVKMNTENPALLILEEIIEG
                                6                                                   60
```

**C**

```
                                5                                  37
Waddlia chon. wcw_0707          DHLTNEKISKNFNNQFDLVNYAIKLAANMIQTG
                                D LTNE+++K F++ F LVNY IK A N I  G
C. trachomatis CTL0286          DRLTNERLNKLFDSPFSLVNYVIKQAKNKIARG
                                5                                  37
```

**D**

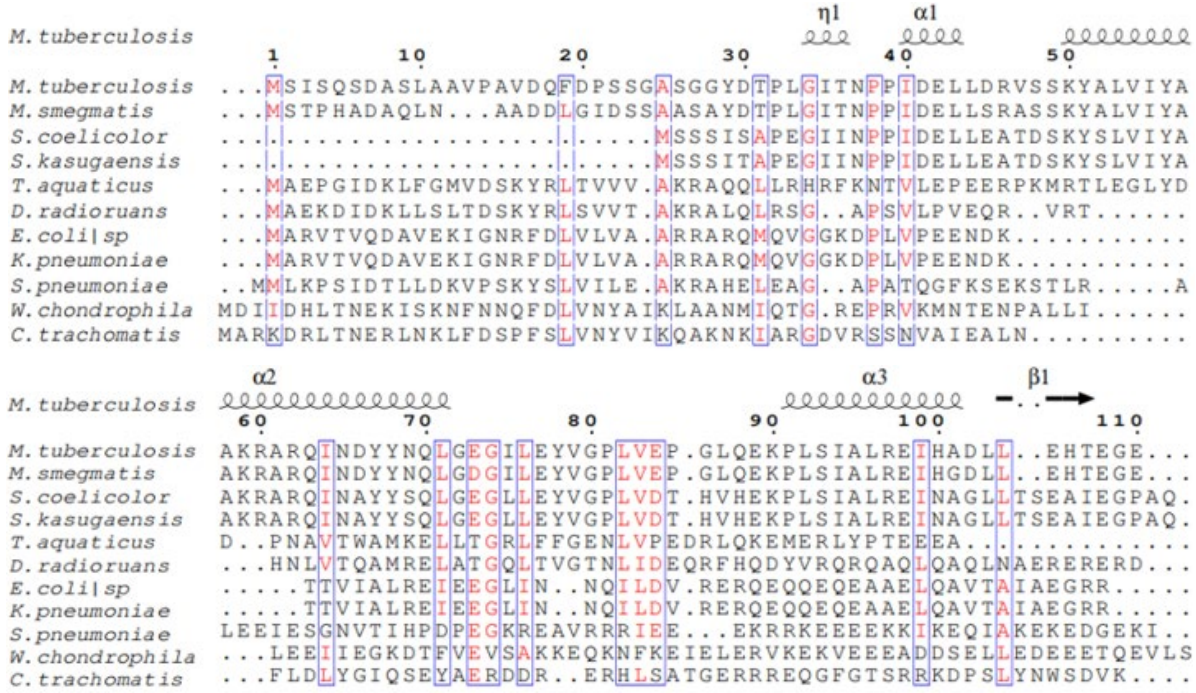


**Fig. 1. Identification of cRNAP ω subunit candidate by BlastP and sequence alignments.** (A) BlastP-detected sequence homology between *Coxiella burnetii* RNAP ω subunit and wcw_0707, a hypothetical protein of the *Chlamydia*-like organism *Waddlia chondrophila*. (B) BlastP-detected sequence homology between *E. coli* RNAP ω and wcw_0707. (C) BlastP-detected sequence homology between wcw_0707 and CTL0286 of *Chlamydia trachomatis*. (D) ClustalX2-detected amino acids conserved in CTL0286 of *C. trachomatis*, wcw_0707 of *W. chondrophila*, and ωs of a variety of bacteria.

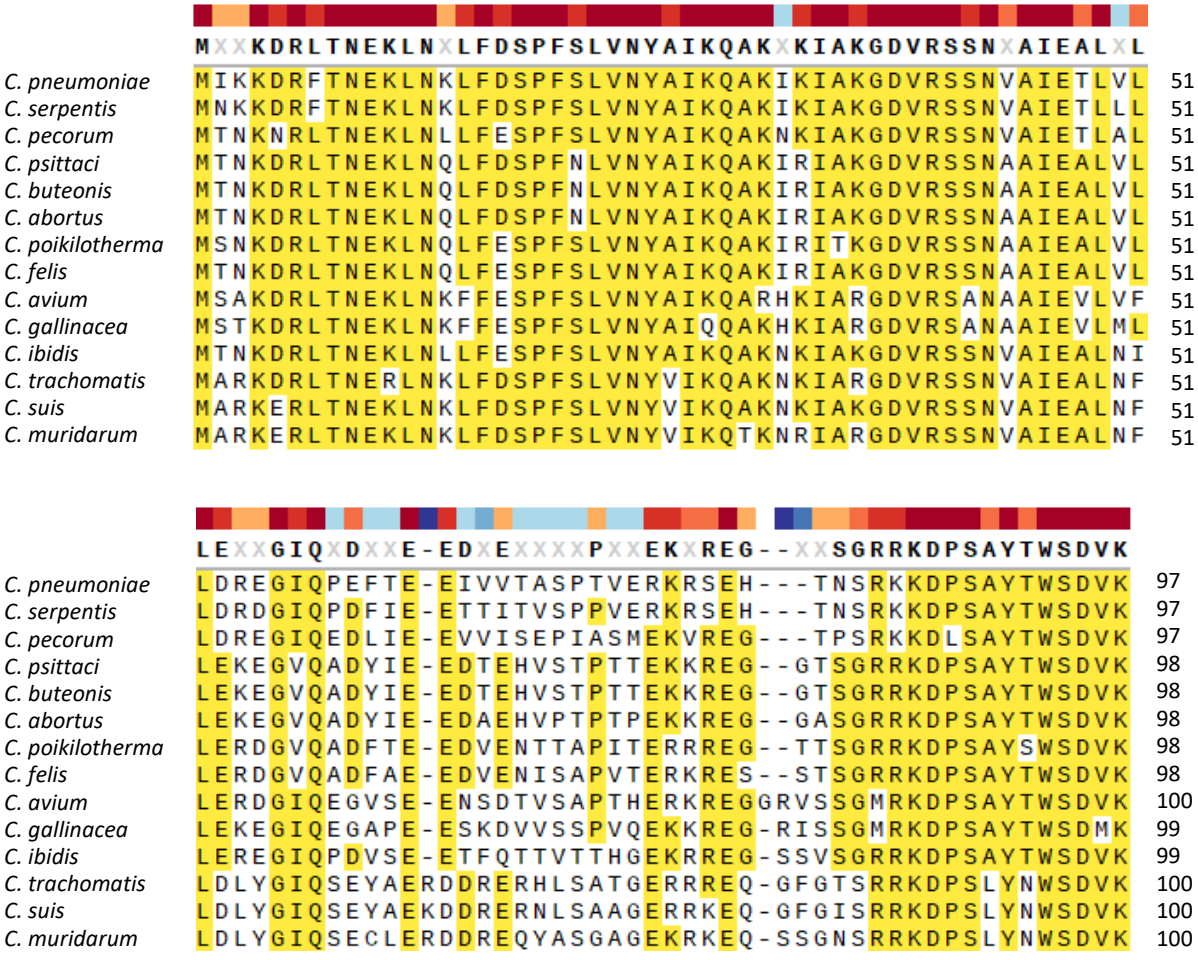


**Fig. 2. A high degree of sequence conservation among candidate ω subunits in all vertebrate chlamydiae.** Alignment was performed using ClustalX2.

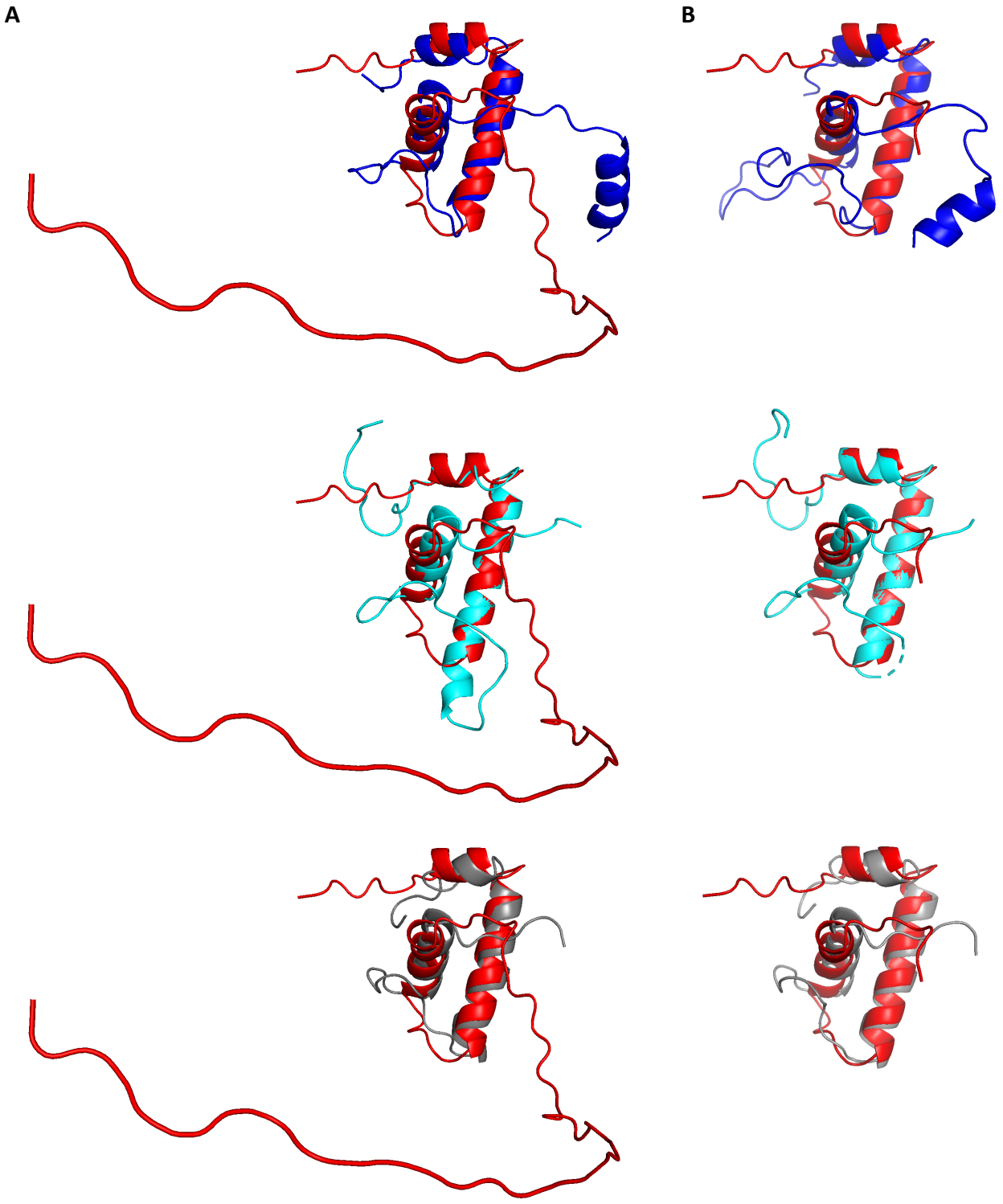


**Fig. 3. AlphaFold predictions for CTL0286.** (A) Superimposition of AlphaFold prediction for full-length CTL0286 (red) on experimental structures of *Clostridium difficile, Mycobacterium tuberculosis,* and *Bacillus subtilis* RNAP ω (blue, cyan, and gray, respectively). (B) Superimposition of AlphaFold prediction for N-terminal region (residues 1-62) of CTL0286 (red) on experimental structures of *Escherichia coli, Mycobacterium tuberculosis,* and *Bacillus subtilis* RNAP ω (blue, cyan, and gray, respectively).

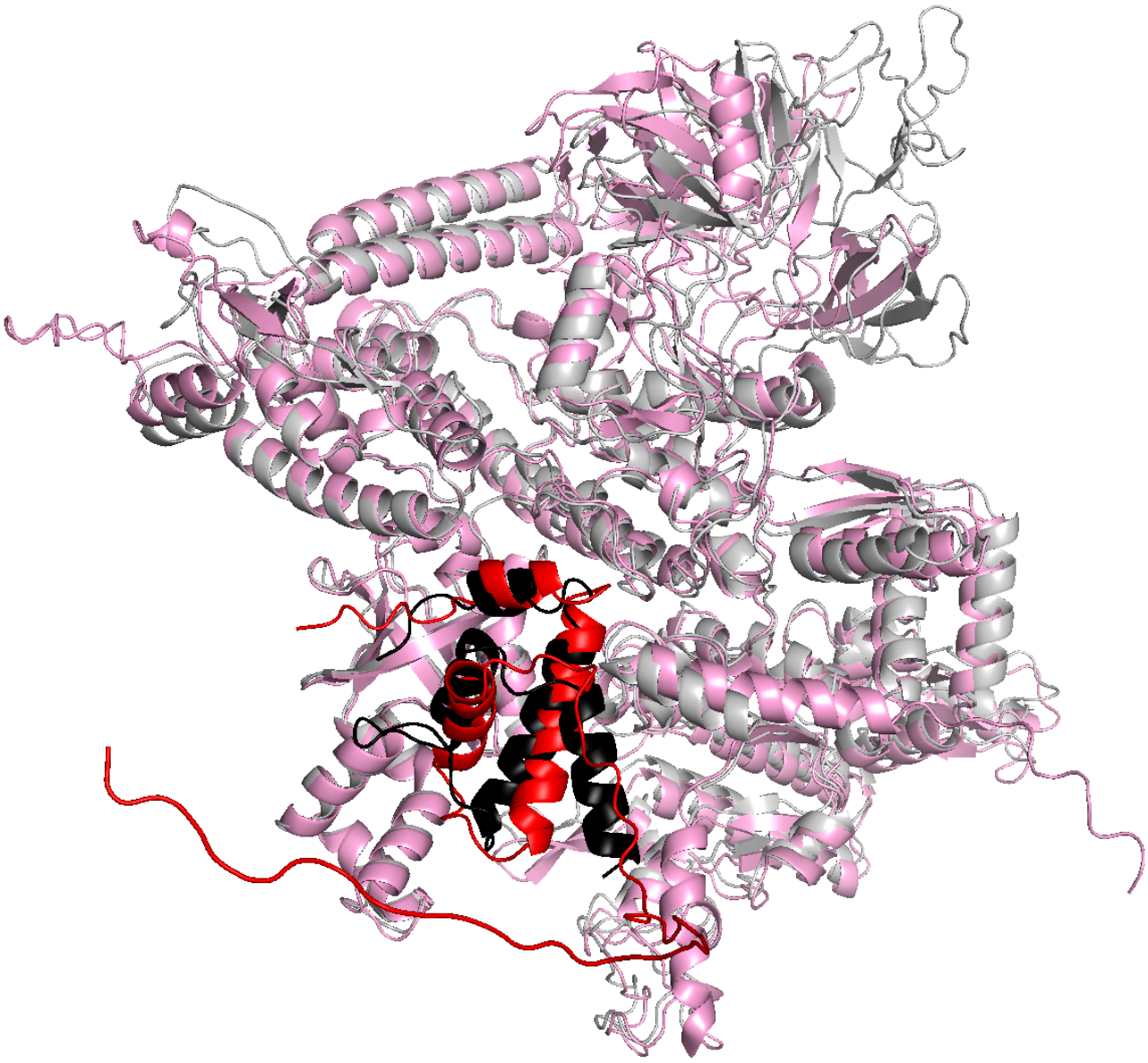

**Fig. 4. AlphaFold-Multimer predictions for complex comprising CTL0286 and *C. trachomatis* RNAP β' subunit.** Superimposition of AlphaFold-Multimer prediction for CTL0286-β' (red for CTL0286; pink for β') on experimental structure of *E. coli* RNAP (PDB 6ALH; black for ω; light gray for β').

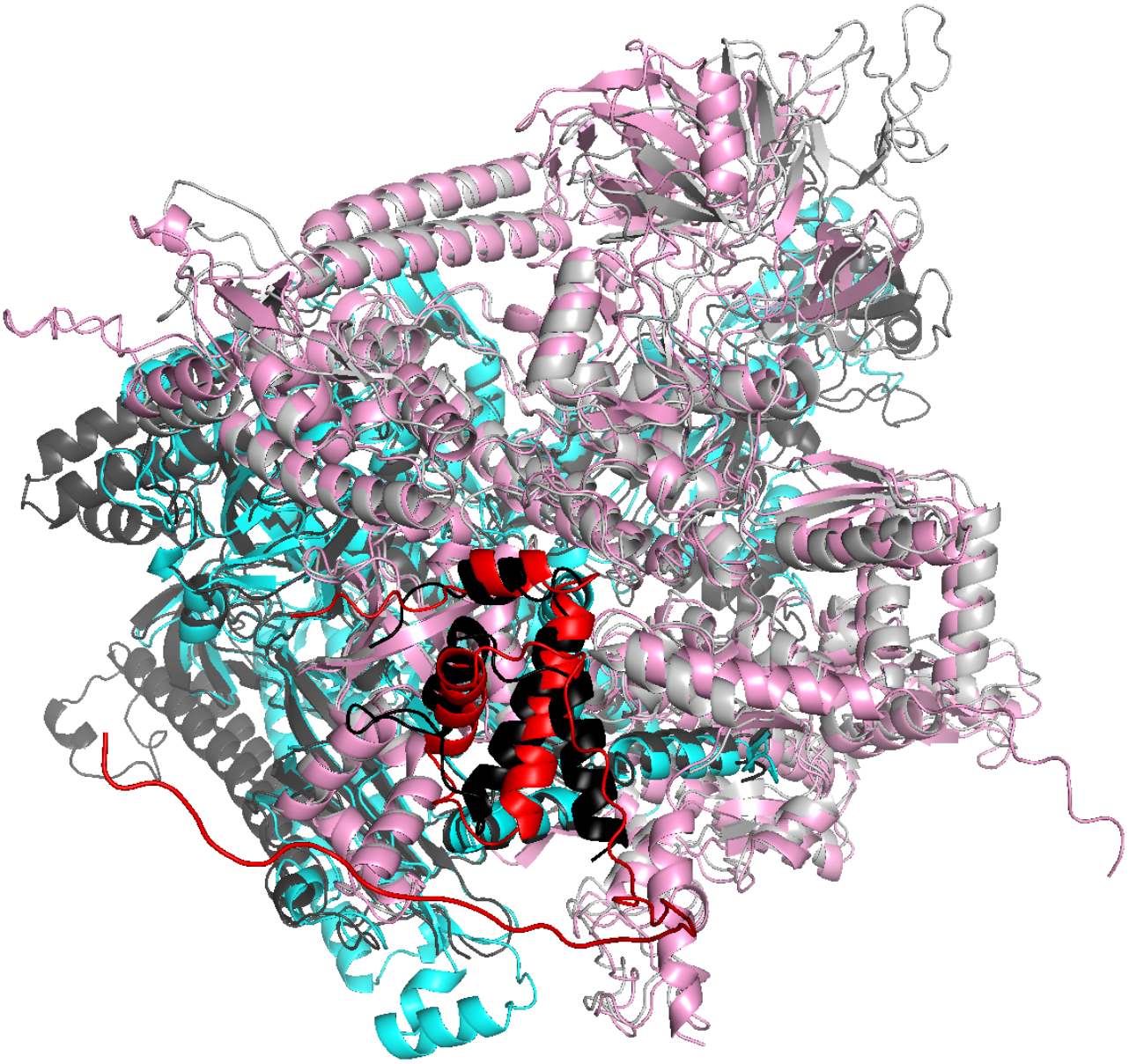

**Fig. 5. AlphaFold-Multimer predictions for complex comprising CTL0286 and *C. trachomatis* RNAP β' subunit, and β subunit.** Superimposition of AlphaFold-Multimer prediction for CTL0286-β'-β (red for CTL0286; pink for β'; cyan for β) on experimental structure of *E. coli* RNAP (PDB 6ALH; black for ω; light gray for β'; dark gray for β).

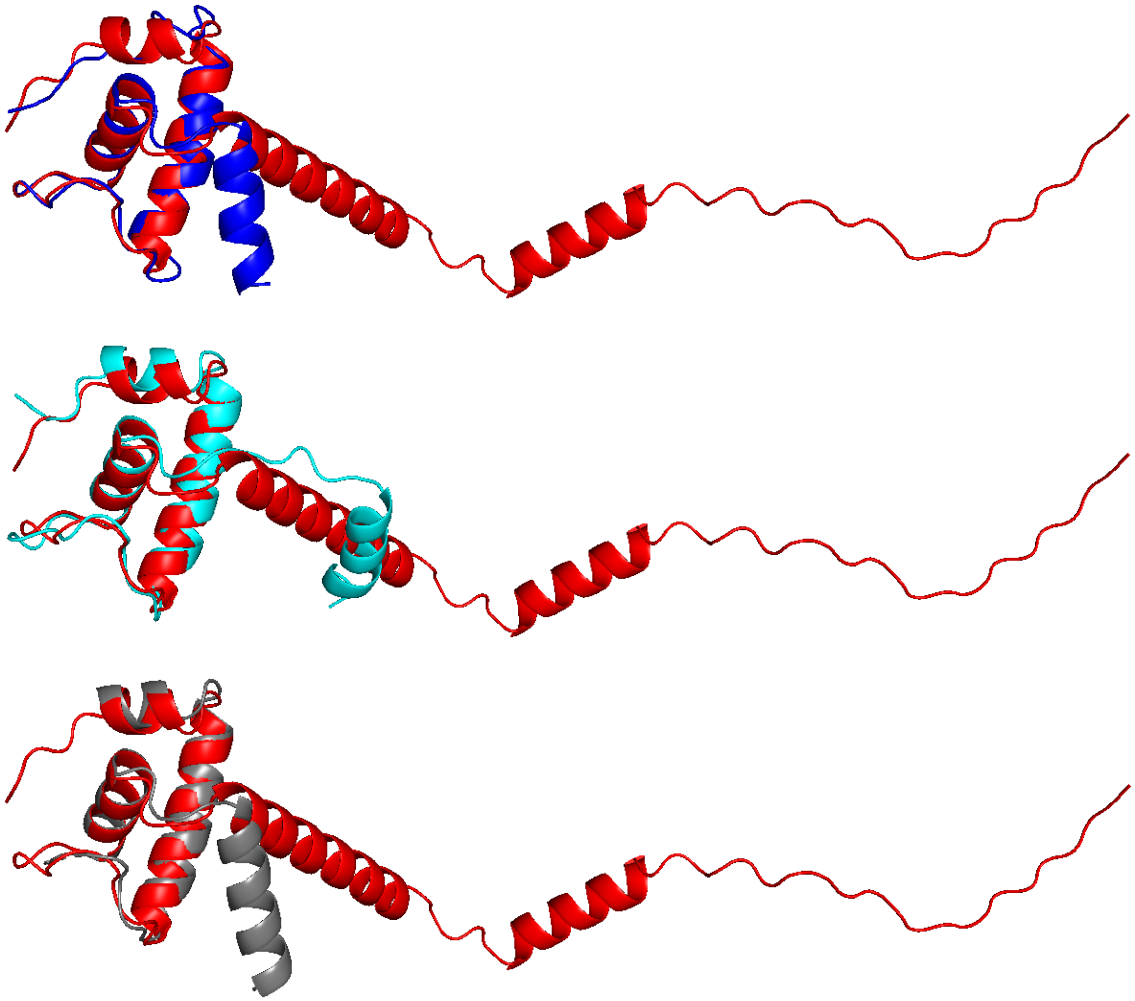

**Fig. 6. AlphaFold predictions for annotated ω of *Candidatus Midichloria* RNAP ω.** Superimposition of AlphaFold prediction for Candidatus Midichloria RNAP ω (red) on experimental structures of *Pseudomonas aeruginosa, M. tuberculosis,* and *Xanthomonas oryzae* RNAP ω (blue, cyan, and gray, respectively).