THE PROTEIN SOCIETY

Hypothetical protein CT398 (CdsZ) interacts with σ^{54} (RpoN)-holoenzyme and the type III secretion export apparatus in *Chlamydia trachomatis*

Michael L. Barta,[1] Kevin P. Battaile,[2] Scott Lovell,[3] and P. Scott Hefty[1]*

[1]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045
[2]IMCA-CAT, Hauptman-Woodward Medical Research Institute, Argonne, Illinois
[3]Protein Structure Laboratory, Del Shankel Structural Biology Center, University of Kansas, Lawrence, KS 66045

Received 27 April 2015; Accepted 6 July 2015
DOI: 10.1002/pro.2746
Published online 15 July 2015 proteinscience.org

Abstract: A significant challenge to bacteriology is the relatively large proportion of proteins that lack sufficient sequence similarity to support functional annotation (i.e. hypothetical proteins). The aim of this study was to apply protein structural homology to gain insights into a candidate protein of unknown function (CT398) within the medically important, obligate intracellular bacterium *Chlamydia trachomatis*. *C. trachomatis* is a major human pathogen responsible for numerous infections throughout the world that can lead to blindness and infertility. A 2.12 Å crystal structure of hypothetical protein CT398 was determined that was comprised of N-terminal coiled-coil and C-terminal Zn-ribbon domains. The structure of CT398 displayed a high degree of structural similarity to FlgZ (Flagellar-associated zinc-ribbon domain protein) from *Helicobacter pylori*. This observation directed analyses of candidate protein partners of CT398, revealing interactions with two paralogous type III secretion system (T3SS) ATPase-regulators (CdsL and FliH) and the alternative sigma factor RpoN (σ^{54}). Furthermore, genetic introduction of a conditional expression, affinity-tagged construct into *C. trachomatis* enabled the purification of a CT398-RpoN-holoenzyme complex, suggesting a potential role for CT398 in modulating transcriptional activity during infection. The interactions reported here, in tandem with previous FlgZ studies in *H. pylori*, indicate that CT398 functions as a regulator of several key areas of chlamydial biology throughout the developmental cycle. Accordingly, we propose that CT398 be named CdsZ (Contact-dependent secretion-associated zinc-ribbon domain protein).

Keywords: structural proteomics; type III secretion system; coiled-coil; *Chlamydia*; Zn-ribbon

BACTH, bacterial two-hybrid; *Cya*, adenylate cyclase; FliI, flagellar ATPase; GST, gluthione S-transferase; LGA, local-global alignment method; T3SS, type III secretion systems; TAP, tandem affinity pulldown

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: U.S. Department of Energy; Grant number: DE-AC02-06CH11357; Grant sponsor: National Institutes of Health; Grant number: AI079083, GM103420 and Grant sponsor: Hauptman-Woodward Medical Research Institute.

*Correspondence to: P. Scott Hefty, Ph.D., 1200 Sunnyside Avenue, University of Kansas, Lawrence, KS 66045.
E-mail: pshefty@ku.edu

Importance

Chlamydia trachomatis is a major human pathogen responsible for numerous infections throughout the world that can lead to blindness and infertility. The research strategy employed within was to determine the structure of a hypothetical protein (i.e. CT398) to identify proteins of similar structure and guide functional studies. This approach is likely to be widely applicable in studying proteins of unknown function within *Chlamydia*.

Introduction

Chlamydia trachomatis, an obligate intracellular bacterial pathogen, is the leading cause of preventable

blindness in underdeveloped countries as well as the most prevalent cause of sexually transmitted bacterial infection worldwide.[1–3] *Chlamydiae* are phylogenetically divergent organisms and characterized by a high percentage (~40%) of proteins of unknown function (hypothetical proteins). The limited functional information for such a large portion of proteins, along with the previous genetically intractable nature, has inhibited a comprehensive understanding of the basic biology and pathogenesis of *C. trachomatis*. Among these are the regulation and mechanistic function of systems that interface with and manipulate the eukaryotic host cell.

Type III secretion systems (T3SS) are a primary mechanism employed by *Chlamydia* to interact with the eukaryotic host cell, enabling access to the intracellular environment, modification of the early endosome and to establish and maintain an intracellular environment. T3SSs are a critical component of two bacterial systems: the flagellum, an extracellular motor essential to motility,[4,5] and the non-flagellar (NF)-T3SS, an energy-dependent, molecular syringe that facilitates the transport of host-altering effector proteins into the host cytosol during infection.[6] NF-T3SSs are a common virulence determinant of pathogenic gram-negative bacteria.[7] While the majority of T3SS-possesing pathogens encode T3S components in a single operon, the chlamydial T3SS is scattered throughout the genome albeit in ten σ^{70}-like regulated operons.[8] Numerous aspects at both the structural and functional level of the T3SS are conserved between the flagellar and NF-T3SS export systems,[9] with strong evidence that the NF-T3SS likely evolved from an exaptation of the flagellar T3SS.[10]

The complete NF-T3SS encoded by *C. trachomatis* is critical for host cell invasion (components denoted with contact-dependent secretion or Cds[11,12]). Proteins secreted by this system are termed effectors, and play a major role in determining the intracellular fate of the inclusion through interactions with host pathways.[13–15] Regulation of secretion hierarchy within *Chlamydia* is poorly understood, especially in regard to the relative paucity of canonical NF-T3SS chaperones and the predicted number of effectors (10% of genome) within this obligate intracellular pathogen.[16] Intriguingly, *Chlamydia* also encode several flagellar T3SS orthologs.[17] As *Chlamydiae* are thought to be non-motile and lack flagella, the presence of these genes is confounding. The continued presence of these flagellar orthologs could merely reflect evolutionary progression of the NF-T3SS. Alternatively, they could provide redundant or flexible functionality in the export of temporally-regulated effectors during the developmental cycle, as interactions between the flagellar ATPase (FliI) and the NF-T3SS ATPase regulator (CdsL) were observed.[18]

Structural biology has been effective in assisting the discovery and subsequent elucidation of protein function. This strategy was leveraged to better understand the function and potential role of a protein of unknown function encoded by *C. trachomatis* CT398. A crystal structure of CT398 was determined and revealed similarity to *Helicobacter pylori* FlgZ (flagellar-associated zinc-ribbon domain protein). Functional traits of FlgZ in *H. pylori* directed subsequent studies to further understand the biological role of CT398 within *C. trachomatis*. All together, these observations suggest that CT398 functions as a regulatory protein partner in several key areas of chlamydial biology, including as a posttranslational chaperone of RpoN and as a modulator of T3S-dependent events.

Results

Crystal structure of CT398 from Chlamydia trachomatis

Protein sequence-based search analysis of CT398 from *C. trachomatis* resulted in numerous weakly scored protein matches (e-values between $1e^{-10}$ and $1e^{-22}$) from across a diverse range of predominantly gram-negative bacterial species. All targets were annotated as proteins of unknown function and encoded a C-terminal DUF164 domain. The DUF164 domain encodes a putative Zn-ribbon fold (C_4-type) with the following motif C-X_2-C-X_{20}-C-X_2-C responsible for metal coordination. This sequence motif is conserved within CT398. DUF164 domains are found throughout bacteria, both alone and along with other domain architectures.[19]

To better understand the function and role of CT398 within *Chlamydia*, purified protein was screened for crystallization. CT398 crystals diffracted synchrotron X-rays to 2.12 Å resolution (Table I) and were solved by SAD, using endogenous Zn^{2+} bound within the C_4-type Zn-ribbon domain, by collecting at the Zn^{2+} κ-edge (1.283 Å or 9.686 keV). The final model (two polypeptides in the asymmetric unit) consists of residues 1-233 of CT398 [Fig. 1(A)], a three-residue (SNA) N-terminal remnant from affinity tag cleavage, 197 water molecules, two Zn^{2+} atoms, and one ethanediol molecule. Interpretable electron density was absent for the last 21 amino acids (residues 234-254), which are presumed to be intrinsically disordered (SDS-PAGE analysis of crystallized CT398 migrated at a similar size as the purified protein; data not shown).

The protein structure revealed that CT398 is comprised of two domains: an ~120 Å long coiled-coil hairpin at the N-terminus and a C_4-type Zn-ribbon at the C-terminus. Several interdomain contacts occur between the coiled-coil and Zn-ribbon domains [Fig. 1(B)], including hydrogen bonding between Asp15 (α1) and Tyr179 (α3) and a salt bridge between Glu13 (α1) and Arg189 (loop between α3 and β1). The modular nature of the CT398 structure suggests each domain could independently fulfill separate

Table I. *Diffraction Data and Refinement Statistics*

	CT398 native	CT398 Zn
PDB ID	4ILO	
Data Collection		
Unit-cell parameters (Å)	$a = 88.35$, $b = 92.63$, $c = 82.65$, $\beta = 95.62$	$a = 87.93$, $b = 92.70$, $c = 82.50$, $\beta = 95.95$
Space group	$C2$	$C2$
Resolution (Å)[1]	48.75–2.12 (2.18–2.12)	48.54–2.20 (2.27–2.20)
Wavelength (Å)	1.000	1.283
Temperature (K)	100	100
Observed reflections	126,564 (10,245)	105,721 (8,352)
Unique reflections	37,012 (2,973)	32,605 (2,720)
$<I/\sigma(I)>$[a]	9.8 (1.9)	9.1 (1.5)
Completeness (%)[a]	98.7 (98.0)	97.7 (95.3)
Multiplicity[a]	3.4 (3.4)	3.2 (3.1)
R_{merge} (%)[a,b]	6.9 (59.5)	7.1 (37.6)
R_{meas} (%)[a,c]	9.5 (81.4)	9.7 (51.5)
R_{pim} (%)[a,c]	6.5 (55.3)	6.6 (35.0)
$CC_{1/2}$[d]	0.998 (0.753)	0.995 (0.864)
Phasing		
Anom. completeness (%)[a]		89.3 (82.8)
Anom. multiplicity[a]		1.5 (1.4)
FOM		0.634
Refinement		
Resolution (Å)	41.13–2.12 (2.15–2.12)	
Reflections (working/test)	36,746/1,829 (2,684/127)	
R_{factor}/R_{free} (%)[e]	20.3/24.0 (31.3/35.4)	
No. of atoms (protein/ligand/solvent)	3761/2/201	
Model quality		
R.m.s deviations		
Bond lengths (Å)	0.012	
Bond angles (°)	1.113	
Average B-factor (Å^2)		
All atoms	30.54	
Protein	31.58	
Ligand	11.87	
Solvent	29.77	
Coordinate error, maximum likelihood (Å)	0.27	
Ramachandran plot		
Most favored (%)	98.7	
Additionally allowed (%)	1.3	
Outliers (%)	0.0	

[a] Values in parenthesis are for the highest resolution shell.
[b] $R_{merge} = \Sigma_{hkl}\Sigma_i \; |I_i(hkl) - <I(hkl)>| / \Sigma_{hkl}\Sigma_i \; I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the ith reflection and $<I(hkl)>$ is the average intensity of all reflections with indices hkl.
[c] R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}.[20,21] R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}.[22,23].
[d] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data.[24,25].
[e] $R_{factor} = \Sigma_{hkl} \; ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} \; |F_{obs}(hkl)|$; R_{free} is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.

functional roles. Consistent with inclusion in the DUF164 superfamily, the Zn-ribbon domain of CT398 contains a single Zn^{2+} atom coordinated by four cysteine residues [amino acids 199, 202, 223, and 226, Fig. 1(C)], each with bond coordination lengths (~2.3 Å), in agreement with canonical values.[26]

CT398 is structurally similar to FlgZ, a regulator of flagellar synthesis in Helicobacter pylori

To provide insight into the function of CT398, structurally related motifs within experimentally characterized proteins were identified via the DALI server.[27] The top five statistically significant hits are listed in Table II. The highest scoring structural homolog was HP0958 (FlgZ) from *H. pylori*.[28] FlgZ and CT398 share 22% sequence identity across the entirety of the coding sequence; however a large majority of this conservation is clustered within the Zn-ribbon domain (32% identity). Unsurprisingly, a majority of the remaining hits are proteins with extended length coiled-coils that mediate protein–protein interactions: including the *Shigella* T3SS first translocator IpaB[29] and the regulatory subunit of phosphoinositide-3-kinase (p85α),[30] among others.

Structural superposition of CT398 and FlgZ [Fig. 2(A)] reveals that the Zn-ribbons from both proteins are extremely similar (RMSD of 0.63 Å over 47/47 Cα atoms within 5.0 Å). There is a significant cluster of

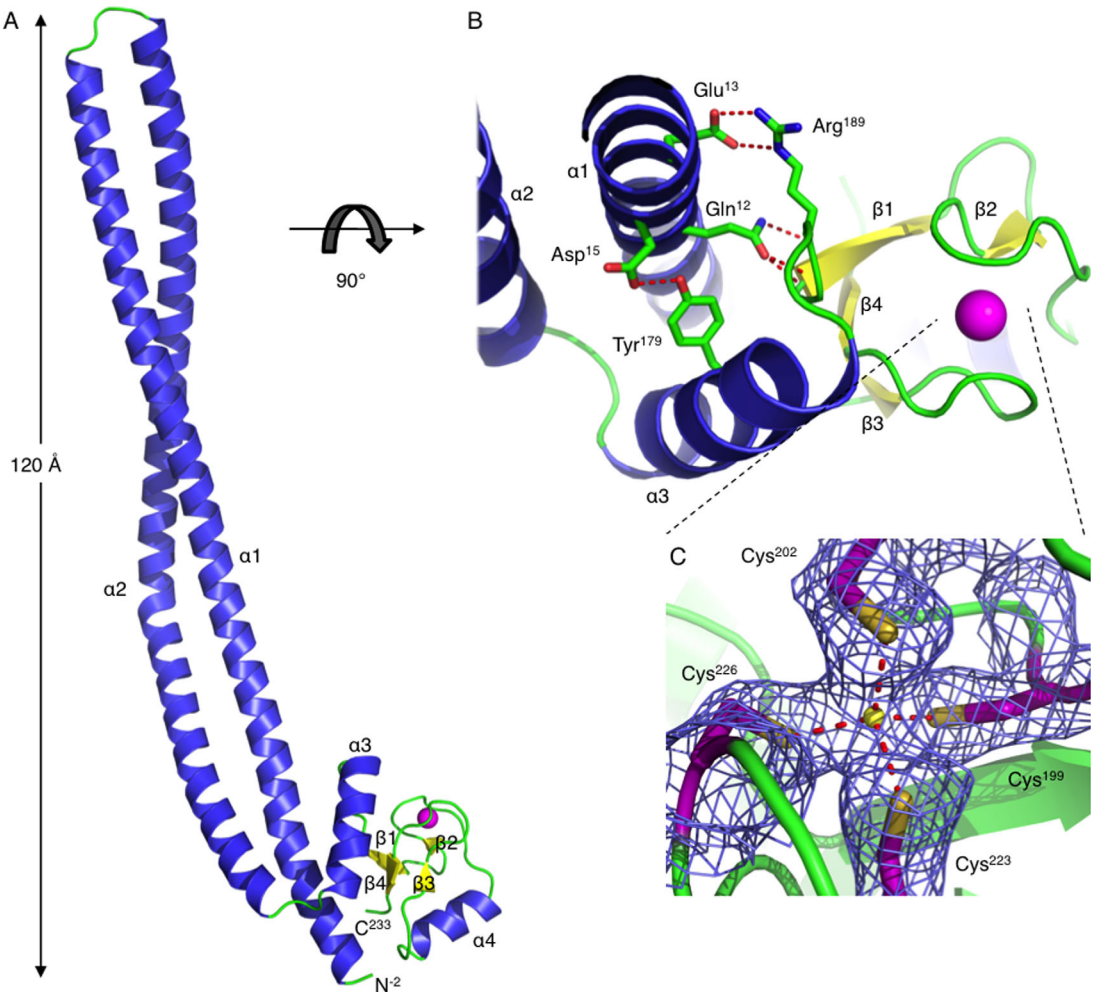

Figure 1. 2.12 Å Crystal structure of CT398 from *Chlamydia trachomatis*. A: crystal structure of *C. trachomatis* CT398 shown in cartoon ribbon format. Two copies of CT398 are found within the asymmetric unit (one shown for clarity) and are colored according to secondary structure (α-helix, blue; β-sheet, yellow; random coil, green). Tetracysteine-coordinated Zn^{2+} atom colored magenta, shown as sphere representation. B: Side chains within hydrogen bonding distance from each domain (coiled-coil and C_4-type Zn-ribbon) are depicted in ball and stick format (green). Backbone color scheme is the same as (A), rotated 90° about horizontal axis. C: $2F_o$-F_c map (blue mesh at 1.5σ contour) of the refined structure with one Zn atom (yellow) modeled per C_4-type Zn-ribbon. Coordinating cysteine residues are depicted in ball and stick format (purple). Bonding distance is 2.3 Å for each Zn-S coordination bond.

positively charged residues surrounding the coiled-coil and Zn-ribbon interface within CT398 [Supporting Information Fig. S1(A)]. Potentially reflecting differences in nucleic acid recognition motifs, positively charged residues are clustered near the coordinated metal binding site and absent from the coiled-coil interface within FlgZ [Supporting Information Fig. S1(B)]. Despite the strong structural similarity observed within each Zn-ribbon domain, only the four essential Cys residues are conserved across DUF164-possessing homologs [Fig. 2(B)].

Intriguingly, the coiled-coil domains of CT398 and FlgZ appear to be quite similar in length (~120 Å), yet CT398 lacks the kink found within α2 of FlgZ resulting in a shift of ~45 Å between the hairpin region of each coiled-coil [Fig. 2(A)]. The coiled-coil domains of CT398 and FlgZ align with an RMSD of 3.66 Å over 170/170 Cα atoms within 5.0 Å (**Fig. SX**). The FlgZ kink was structurally linked to a heptad stutter involving side chain interactions between Arg29 and Asp33.[28] The side chains at equivalent positions within CT398 (Gln and Ala, respectively; blue boxes in Supporting Information Fig. S2) are incapable of interacting, providing a structural rationale for this difference between CT398 and FlgZ. Given the prevalence for coiled-coil domains to mediate protein–protein interactions,[31] it seems plausible that these structural differences could impart, or be a reflection of, organism-specific protein binding properties.

FlgZ was identified as a gene essential for proper flagellum formation in *H. pylori*, as a strain with a disruption in the *flgZ* coding region was non-motile and failed to produce flagella.[32,33] A genome wide yeast two-hybrid screen within *H. pylori* identified several potential protein–protein interactions

Table II. *CT398 DALI Search Statistics*

Protein name	PDB code	*Z*-score[a]	RMSD	*Cα* range[b]	.% ID[c]
FlgZ (HP0958), *H. pylori*	3NA7	17.8	5.7	233/237	22
IpaB, *S. flexneri*	3U0C	13.1	5.1	136/151	13
niSH2 p85α, *H. sapiens*	3HHM	11.8	2.3	127/247	8
Prefoldin, *M. thermautotrophicus*	1FXK	11.2	2.0	97/109	11
Lysine-specific histone demethylase, *H. sapiens*	2X01	10.4	14.4	154/670	9

This table gives the CT398 structural homology and top 5 (unique) DALI scores.
[a] Similarity score representing a function that evaluates the overall level of similarity between two structures. *Z*-scores higher than 8.0 indicate that the two structures are most likely homologous.[27].
[b] Denotes the number of residues from the query structure that superimpose within an explicit distance cutoff of an equivalent position in the aligned structure.
[c] Denotes the percent sequence identity across the region of structural homology.

involving FlgZ, most notably the alternative sigma factor σ^{54} (RpoN) and FliH, a regulator of ATP-dependent translocation within the flagellar-T3SS.[34] Subsequent studies further implicated FlgZ as both a posttranslational chaperone of RpoN and in the proper secretion of flagellin via the binding of *flaA* mRNA transcripts and subsequent interaction with FliH at the T3SS export apparatus. Douillard *et al*. proposed that the FlgZ-FliH interaction was expected to create a concentration gradient of *flaA* mRNA transcripts at flagellar-T3SS export sites that could be rapidly secreted upon translation.[35] Given the structural similarity with FlgZ, CT398 is hypothesized to maintain similar protein interactions within *Chlamydia*, albeit with components of the functionally related NF-T3SS.

CT398 interacted with RpoN and T3SS ATPase-regulators in vitro

Physical associations with chlamydial homologs to FlgZ interacting proteins were systematically tested using the adenylate cyclase (*cya*) bacterial two-hybrid (BACTH) method.[36] *Chlamydia* encode a complete set of genes for a NF-T3SS that is critical to infection and, despite the absence of a functional flagellum, several homologs to flagellar-T3SS export proteins.

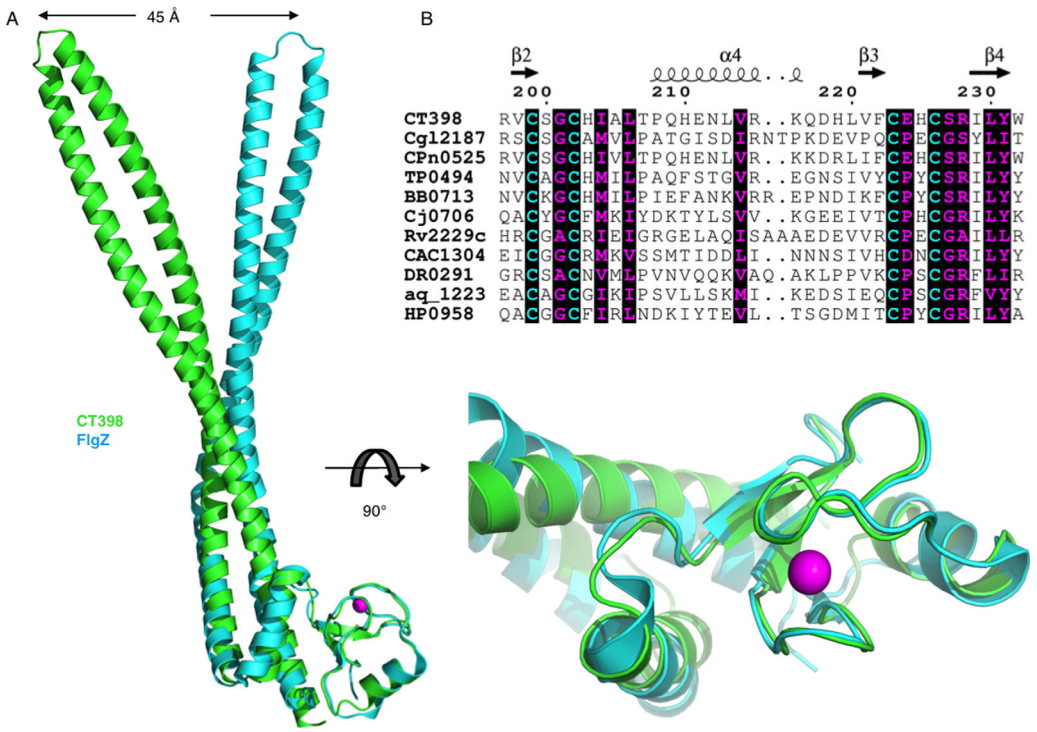

Figure 2. Structural superposition of C_4-type Zn-ribbons from CT398 and FlgZ. A: Cartoon ribbon diagram of a structural alignment of the C_4-type Zn-ribbons from CT398 (residues 186–233, green) and FlgZ (residues 188–235, cyan) with an RMSD of 0.63 Å over 47/47 Cα atoms within 5.0 Å, rotated 90° about horizontal axis. B: Limited structure-based sequence alignment of DUF164 domain possessing proteins colored according to residue conservation (cyan = absolute and purple = similar) as judged by the BLOSUM62 matrix. Numbers above the sequences correspond to *C. trachomatis* CT398. Secondary structure elements of CT398 are shown above the alignment.

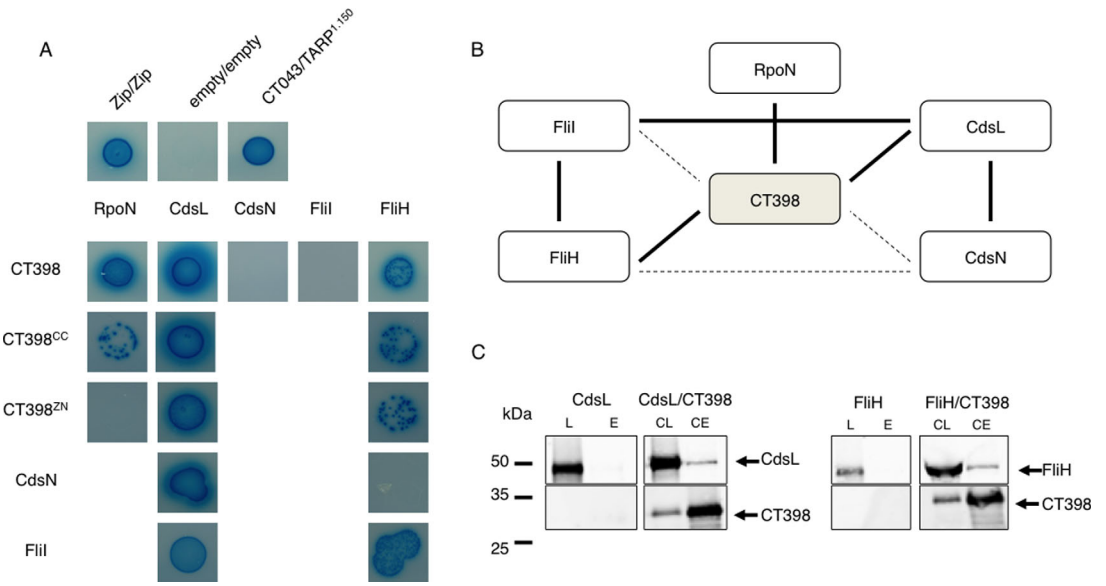

Figure 3. Identification of CT398 Protein-Protein Interactions. A: Detection of protein–protein interactions involving positive (Zip/Zip and CT043/TARP$^{1.150}$) and negative (empty/empty) controls and CT398, RpoN, and both flagellar and non-flagellar T3SS ATPase and ATPase regulators. Growth and blue color signal positive interactions between cotransformed fusion proteins; methodology can be found in Materials and Methods. Data are representative of multiple independent cotransformation trials. Further information regarding fusion partner orientation can be found in Supporting Information Table I. B: Interactome map based upon panel A. Solid black lines indicate positive interaction was detected (i.e. blue colonies) while dashed black lines indicative no interaction was detected (i.e. white colonies). C: Immunoblot of CT398-His$_6$, CdsL-GST and FliH-GST co-immunoprecipitation from *E. coli*. Each GST fusion protein was coexpressed in the absence (L and E lanes) or presence (CL and CE lanes) of CT398-His$_6$ and purified over Ni^{2+}-affinity resin. Samples of preincubated lysate (L and CL lanes) and eluate (E and CE lanes) were probed with α-CT398 (red) and α-GST (green). Immunoblot is representative of multiple independent experiments.

Proper recognition, unfolding, and secretion of substrates in both systems are regulated by interactions between the T3SS ATPase and ATPase-regulator proteins. The predicted presence of both flagellar- and NF-T3SS ATPase/ATPase-regulator pairs in *Chlamydia* (FliI/FliH and CdsN/CdsL,[37] respectively) is of particular note, given the proposed interaction between FlgZ and FliH. Additionally, while yet to be functionally characterized, one of the three sigma factors encoded by *Chlamydia* is RpoN.

Putative interactions were tested between CT398 and RpoN, both ATPases (FliI and CdsN) and both ATPase-regulators (FliH and CdsL). Each of the eight possible BACTH vector combinations[36,38] were tested, with these vectors facilitating the expression of hybrid proteins that are fused to the N- or C-terminus of the T25 or T18 *cya* fragment. Control transformations [Fig. 3(A)] included empty vectors (only T25 or T18 *cya* fragment) as a negative control along, while two sets of positive controls were a leucine zipper fragment combination and the chlamydial T3S chaperone/effector pair, Slc1/TARP$^{1.150}$.[39,40]

Several interactions between CT398 and the tested proteins were detected [Fig. 3(A,B) and Supporting Information Table S1], including RpoN and both T3SS ATPase-regulators (FliH and CdsL). The crystal structure of CT398 revealed a modular structure consisting of two domains [Fig. 1(A)]: an N-terminal coiled coil (CC, residues 1-171) and a C-terminal C_4-type Zn-ribbon (ZN, residues 172-254). To better understand the physical basis for the elucidated CT398 interactions, the CC and ZN domains were separately cloned into each BACTH vector and systematically tested for interaction with RpoN, FliH, and CdsL [Fig. 3(A)]. Only the CT398 CC domain was required to maintain interaction with RpoN. Domain-based interaction studies have not been performed with *H. pylori* FlgZ, but the FlgZ ZN domain was demonstrated to interact with nucleic acids (specifically *flaA* mRNA[35,41]). This indicates that the CT398 Zn-ribbon could potentially bind nucleic acids while simultaneously interacting with RpoN through the coiled-coil domain.

In contrast to domain-based interactions with RpoN, interactions with FliH and CdsL were detected between each individual domain of CT398 [Fig. 3(A)]. As structural information is lacking for both ATPase-regulator paralogs, computational modeling (I-TASSER[42]) was employed to better understand domain-based interactions between CT398 and FliH/CdsL. ATPase-regulator models were generated by I-TASSER based upon homology with the V-type ATP synthase subunit E.[43] Both models (Fig. 4) are comprised of an extended length α-helix followed by a

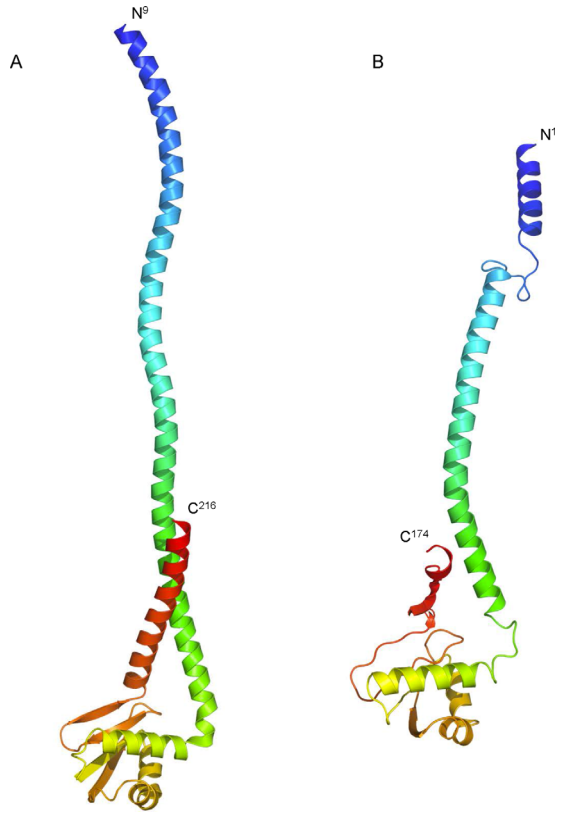

Figure 4. Computational models of *C. trachomatis* T3SS ATPase Regulators generated by I-TASSER. A: Computational model of CdsL (CT561) from *C. trachomatis* generated by I-TASSER[44] depicted in cartoon ribbon format using common rainbow colors (slowly changing from blue N-terminus to red C-terminus). Estimated accuracy of CdsL model: C-score: −1.25, TM-score: 0.56 ± 0.15. B: Computational model of FliH (CT718) from *C. trachomatis* generated by I-TASSER[44] depicted in cartoon ribbon format using common rainbow colors (slowly changing from blue N-terminus to red C-terminus). Estimated accuracy of FliH model: C-score: −2.00, TM-score: 0.48 ± 0.15.

globular C-terminal domain. The CdsL model (C-score: −1.25, TM-score: 0.56 ± 0.15) shares 15% sequence identity with the V-type ATP synthase subunit E from *Pyrococcus* (PDB ID: 4DT0[45]) and aligns with an RMSD of 2.40 Å. Despite only sharing 5% sequence identity with this same protein, the FliH model (C-score: −2.00, TM-score: 0.48 ± 0.15) aligns with an RMSD of 2.87Å. This suggests that CT398 and the ATPase-regulators are structurally similar and it seems plausible that interactions could occur between equivalent domains within each protein. These interactions would then facilitate the release of mRNAs bound to CT398 resulting in translation and secretion upon targeting to the T3SS export apparatus.[35]

To provide orthogonal support for the identified interactions, CT398 protein–protein interactions were probed using a GST pull-down assay. Attempts to express RpoN in *E. coli* failed to yield soluble protein regardless of the presence of CT398. Compatible plasmids encoding fusion proteins with N-terminal GST fused to full-length RpoN, FliH, or CdsL were cotransformed with His_6-tagged CT398. To demonstrate interaction specificity, two controls were performed. First, each GST fusion protein was purified on a Ni^{2+}-column in the absence of CT398-His_6 and was not retained by the column [Fig. 3(C)]. Second, GST alone failed to interact with CT398-His_6 when copurified over this column (data not shown), indicating CT398 does not have natural affinity for GST. Upon coexpression and purification of CT398-His_6 and each GST fusion protein partner on a Ni^{2+}-column, coelution was tested by Western Blot [Fig. 3(C)]. Both ATPase-regulators were only retained within the Ni^{2+}-column in the presence of CT398, indicating specific protein–protein interaction had occurred and providing support for the identified BACTH interactions.

To investigate the potential for flagellar/NF-T3SS cross-reactions within *C. trachomatis*, all possible ATPase/ATPase-regulator interactions were tested via BACTH screening (i.e. CdsN-CdsL, CdsN-FliH, FliI-FliH and FliI-CdsL). As expected, interactions between CdsN-CdsL and FliI-FliH were observed [Fig. 3(A,B)]. Intriguingly, we also detected an interaction between CdsL-FliI, but were unable to detect an interaction between FliH and CdsN [Fig. 3(A,B)]. Previous studies within *C. pneumoniae* identified equivalent interactions between orthologs of the proteins described herein.[18] These authors speculated that the small set of flagellar chlamydial proteins could form a hybrid structure with NF-T3SS components. The interactions reported here support the potential for this type of assembly to occur.

CT398 interacted with RpoN and core RNA polymerase in vivo

Given the difficulties expressing recombinant chlamydial RpoN in *E. coli*, as well as the absence of *Chlamydia*-specific RpoN antibodies, we took advantage of the recently described inducible expression system within *Chlamydia*[46] to perform *in vivo* protein pull-down experiments. This also provides a more physiologically relevant context to evaluate the proposed interactions. To provide high-specificity for protein association studies, the tandem affinity purification (TAP) system was employed. The TAP tag encodes epitopes to both FLAG and Strep-tag II antibodies and can be placed at the amino (N) or carboxy (C) terminus of a protein of interest within the pTL2 vector system. *C. trachomatis* was transformed with a modified shuttle vector capable of controlling the expression of tandem affinity-tagged RpoN [Fig. 5(A); pTL2-RpoN-TAPN]. Expression of RpoN was induced from 16–20 hpi with 2 ng ml^{-1} ATc. These developmental cycle time points correlate with a rise in endogenous levels of RpoN.[47] Infected cells were collected, washed, and treated with a

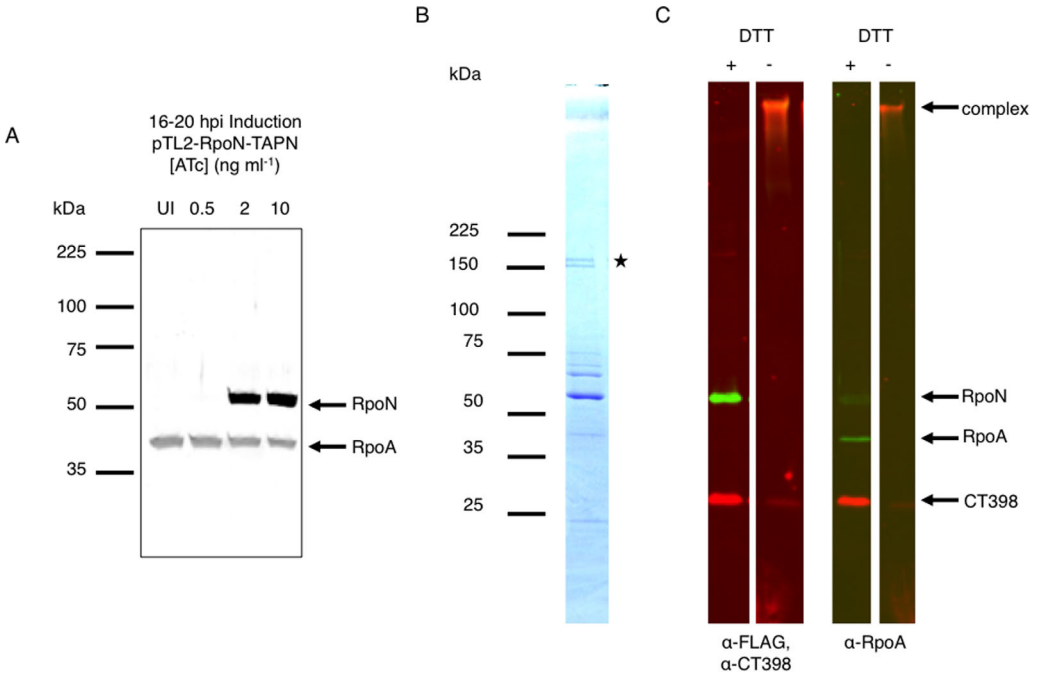

Figure 5. Co-Immunoprecipitation of RpoN-TAPN demonstrates interaction with CT398 and core RNA polymerase within *C. trachomatis*. A: Expression levels of pTL2-RpoN-TAPN transformed *C. trachomatis* were analyzed with increasing levels of ATc (0.5–10 ng ml^{-1}, induced from 16 to 20 hpi). Western blot was performed with the following antibodies, rat anti-RpoA, and mouse anti-FLAG (recognizing RpoN-TAPN). UI, uninduced. B: Co-immunoprecipitated RpoN-TAPN (described in Materials and Methods) was analyzed by SDS-PAGE. Image contrast has been manipulated to more easily visualize protein bands. Black star indicates β and β′ subunits of RNA polymerase. C: Immunoblot analysis of gel from panel B. Western blot was first performed with mouse anti-FLAG and rabbit anti-CT398 (left panel) and then subsequently probed with rat anti-RpoA (right panel) to establish specifity. Thiol-sensitive crosslinker DSP was utilized prior to lysis and indicates that immunoprecipitated proteins form an extremely high MW complex (DTT negative (-) lanes).

thiol-cleavable intracellular amine-to-amine crosslinker (DSP; 12.0 Å spacer) to maintain weak or transient protein–protein interactions during purification. Infected cells were lysed, clarified, and sequentially applied to affinity resins for each of the two epitopes in the TAP tag. Additionally, this entire process was concurrently performed on untransformed *C. trachomatis* (data not shown).

The final eluate was analyzed by SDS-PAGE [Fig. 5(B)] and immunoblot [Fig. 5(C)], both with and without reducing agent, for the presence of coeluted protein partners. Both CT398 and RpoA (α subunit of RNA polymerase) were detected only when tagged-RpoN was purified (i.e. not in wild-type infections), indicating the sequential purification process results in very low levels of background contamination. Furthermore, all three immunodetected proteins were found as a single, extremely high MW band in the unreduced crosslinker lane. Additionally, two bands unique to the RpoN co-IP [black stars in Fig. 5(B) at ~150 kDa] were extracted and analyzed by LC-MS/MS (data not shown), which indicated they were the β and β′ subunits of chlamydial RNA polymerase. Altogether, these observations indicate that CT398, RpoN and core RNA polymerase formed a single complex within *Chlamydia*.

Discussion

Nearly 40% of the genes within *Chlamydia trachomatis* lack an annotated function (as determined by BLAST analysis), greatly hindering our understanding of the basic biology, development and pathogenesis of this human pathogen. Historically, obligate intracellular bacteria such as *C. trachomatis* have been highly resistant to genetic manipulation. To gain insight into how and why these hypothetical proteins function, we have utilized X-ray crystallography within this study to identify functionally characterized structural homologs to CT398. In addition to guiding subsequent studies, crystal structures provide the atomic level details critical to understanding the molecular mechanisms of protein function. Structural analysis of CT398 guided further biological investigation of this hypothetical protein that ultimately provided support for a functional role in the regulation of both transcription and T3S-mediated events. This study further underscores the potential that structural proteomics can provide in the study of an obligate intracellular pathogen.

Chlamydia are obligate intracellular pathogens that are essentially nonmotile (despite exhibiting intrainclusion Brownian-like motion) and lack flagella. Thus, the presence of several predicted flagellar

orthologs (notably FlhA, FliI, FliH, and FliF) is confounding, especially in regard to the high degree of reductive evolution *Chlamydia* has undergone. However, orthologs of these proteins within flagella-possessing organisms are involved in formation of the T3SS export apparatus at the inner membrane.[4] Within *Chlamydia* structural components of both the flagellar and NF-T3SS are temporally regulated, with transcript levels increasing throughout the developmental cycle. Quantitative gene expression analysis indicated that transcript levels of the flagellar genes were several fold lower than each NF-T3SS paralog throughout the developmental cycle.[8] Despite these differences, the high degree of structure/function conservation across both types of T3SSs, chlamydial flagellar proteins could feasibly play a role in T3S substrate recognition and export (through FliI) and/or recruitment of secreted targets (through the observed CT398-FliH interaction). It remains to be determined whether these events could occur apart from or in tandem with existing NF-T3SS assemblies.

CT398 interacts with RpoN, suggesting a potential role in regulating the function of this alternative sigma factor in *Chlamydia*. Two approaches support this interaction: bacterial two-hybrid and *in vivo* affinity pull-down. However, the possibility that this interaction may be indirect (i.e. requires another protein to bridge interaction) cannot be ruled out as there is no evidence of interaction with only the two proteins (e.g. *in vitro* recombinant protein association). The interaction between *H. pylori* FlgZ and RpoN was detected within a yeast two-hybrid system and could also require a mediating factor.[34] While the possibility of an indirect interaction cannot be dismissed, the observation that all FlgZ/CdsZ interactions with RpoN were detected in various microbial backgrounds (yeast, *E. coli*, and *Chlamydia*) more strongly supports a direct interaction. Domain-based BACTH analysis of the CT398-RpoN interaction indicates that only the coiled-coil domain is required for binding [Fig. 3(A)], potentially allowing the Zn-ribbon domain to concurrently contact DNA and provide an additional level of RpoN-transcriptional regulation. In the context of the RpoN-TAPN co-IP results this scenario seems increasingly likely as CT398 coeluted with RNA polymerase holoenzyme, which suggests this complex is formed prior to RNA polymerase elongation and sigma factor release. Additionally, CT398 may stabilize RpoN, facilitating an increased ability to compete with other sigma factors for core RNA polymerase. While the chlamydial RpoN-regulon remains to be described, within *H. pylori* class II flagellar genes are controlled by RpoN.[32,33] As *Chlamydia* do not produce flagella, it is unclear what parallels can be drawn, but the apparent structure/function similarities between CT398 and FlgZ allow us to speculate that chlamydial RpoN is potentially involved in regulating genes that are secreted (e.g. T3SS effector proteins). T3SS effector proteins typically are pathogen-specific, likely reflecting the tissue tropism of the pathogen, and as a result are often hypothetical proteins. Intriguingly, a previous bioinformatic prediction of RpoN promoter sites within *Chlamydia*[48] identified elements upstream of two hypothetical proteins (CT652.1 and CT683); however, experimental validation of these putative RpoN promoters is yet to be reported. Future studies utilizing newly described genetic tools within *Chlamydia* will likely provide critical insights into the functional outcomes of the CT398-RpoN interaction.

The expression pattern of CT398 supports that it is likely involved in regulating NF-T3S events during early time points in the developmental cycle as well. Interactions between CT398 and both ATPase-regulators (FliH and CdsL) were observed via BACTH and GST coprecipitation studies. FliH and CdsL function as negative regulators of T3SS ATPases (FliI and CdsN, respectively), preventing the hydrolysis of ATP until substrate export can be energetically coupled. Interactions with ATPase-regulators are consistent with those previously described within *H. pylori*. Douillard *et al*. proposed that the FlgZ-FliH interaction, coupled with the ability of the FlgZ Zn-ribbon to bind *flaA* mRNA, would enable rapid secretion upon translation at the NF-T3SS export system.[35] While *H. pylori* utilize a functionally distinct Type IV Secretion System for pathogenic purposes,[49] the extreme similarities between flagellar and NF-T3SS export systems suggests that CT398 might retain this functional characteristic (i.e. binding of mRNA transcripts and targeting to the T3SS export system). This aspect is highly speculative, as no evidence has been discovered that CdsZ interacts with mRNA for any genes in *Chlamydia*.

Another observation in this study warranting further discussion is the *C. trachomatis* transformation protocol. The protocol we have described is based upon the collective experimental observations of several recently published reports[50–53] and has been highly successful with a variety of buffers, chlamydial EB preparations and perhaps most importantly, gene targets (40+ unique ORFs have currently been transformed in our hands). It appears that the critical variables for success in our hands involve the following steps: the use of methyltransferase deficient *E. coli* cloning strains, ethanol precipitation/concentration of the modified shuttle vector, the addition of antibiotics between 12 and 16 hpi within the initial transformation and early lysis (~32 hpi) of the second passage (e.g. S2). These last two steps serve to greatly reduce the amount of wild-type (e.g. nontransformed) progeny from the initial infection and decrease the likelihood that transformed *Chlamydia* are lost to host cell cytotoxicity during S1 and S2 passages. The varying diversity in published *Chlamydia* transformation

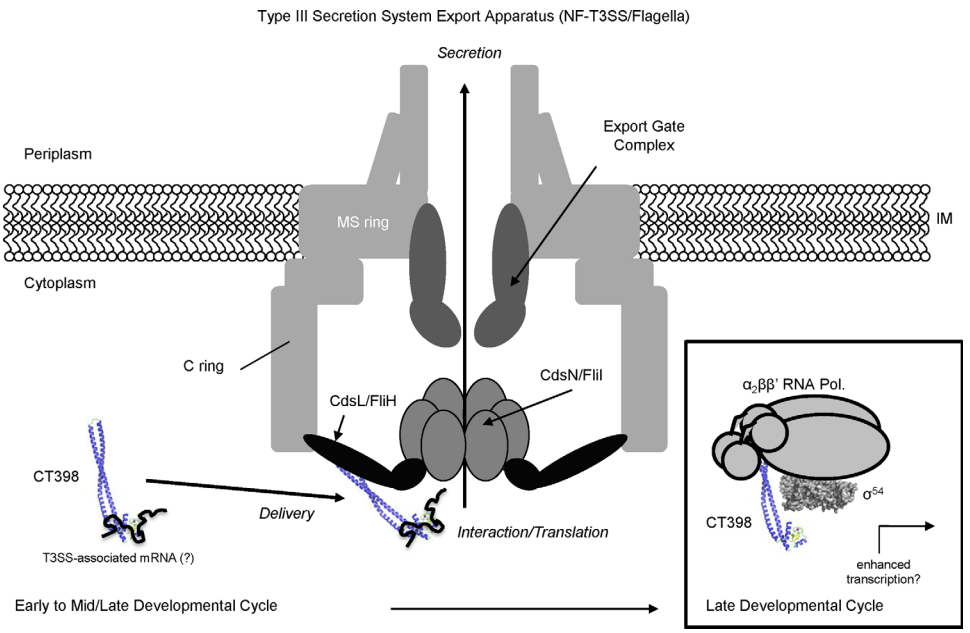

Figure 6. Proposed mechanism of CT398 biological role within *C. trachomatis*. Interactions with T3SS ATPase regulators and structural similarities to *H. pylori* FlgZ suggest a role for CT398 in the regulation of the NF-T3SS export apparatus, which could involve targeting specific mRNAs for translation and secretion and occur early in the developmental cycle. Upon expression of RpoN later in the developmental cycle, CT398 may switch roles and modulate (in an as yet unknown manner) RpoN-dependent binding events/transcription. Relevant NF-T3SS and flagellar orthologs are labeled. IM, inner membrane. Chlamydial T3S sorting platform cartoon model is adapted from *Shigella*.[54]

protocols suggests a detailed methodological study outlining the critical variables for success is needed.

CT398 is a multifaceted protein capable of interfacing with several protein partners. The studies described herein have allowed us to develop a model for how CT398 functions within *Chlamydia* (Fig. 6). Even though CT398 is constitutively expressed throughout the chlamydial developmental cycle [**Fig. S3(A)** and Ref. 40], the identified protein partners are not. Both RpoN and the NF-T3SS export apparatus are developmentally regulated, with expression levels increasing as the developmental cycle progresses. EBs are, however, equipped with active, assembled NF-T3SS injectisomes,[55] indicating that CT398 likely interacts with CdsL/FliH predominantly at early time points in the developmental cycle. As the developmental cycle progresses and RpoN protein levels begin to rise,[47] CT398 could then switch roles and function as an RpoN-chaperone at late time points in the developmental cycle. Alternatively, CT398 appears to be an extremely abundant protein in both EBs and RBs, and may be in a large enough excess to interact with all partners throughout the developmental cycle. The ability for the same protein to function as a protein chaperone and gene expression regulator, specifically with respect to the NF-T3SS of other bacteria, is well documented. Within *Shigella* for example, after the class II chaperone IpgC delivers the first and second hydrophobic translocators (IpaB and IpaC, respectively) to the T3S injectisome for secretion,[56] it then binds as a cofactor to the AraC-family transcriptional activator MxiE to upregulate the expression of late effector proteins.[57,58]

Overall, it appears that CT398 is involved in many aspects of chlamydial biology and pathogenesis. However, numerous intriguing questions regarding this role still remain, and include what genes are regulated by RpoN, how does CT398 affect transcription, and is CT398 able to target mRNA transcripts to the NF-T3SS export apparatus, and if so, what are these targets. Recently developed molecular and genetic tools in *Chlamydia* can likely address these questions and further expand our understanding of NF-T3SS.

Materials and Methods

Cloning, overexpression, and purification of recombinant Chlamydia *proteins*

A gene fragment encoding the entire open reading frame (residues 1–254) of CTL0655 was amplified from *C. trachomatis* (serovar L2 434/Bu) genomic DNA via PCR and subcloned into *Ssp*I-digested pTBSG through ligation independent cloning.[59] As CTL0655 and CT398 from C. *trachomatis* serovar D/UW-3 are 99% identical (V178I and A237S substitutions) references throughout the text to this gene product will utilize CT398 in order to keep with standard nomenclature in the *Chlamydia* field. Upon DNA sequence confirmation, the vector was transformed into Rosetta 2(DE3) pLysS *E. coli* competent cells. This strain was grown to an OD_{600} of 0.8 at 37°C

within terrific broth supplemented with Ampicillin (100 μg mL^{-1}) and Chloramphenicol (30 μg mL^{-1}), and protein expression was induced overnight at 16°C by the addition of isopropyl 1-thio-β-D-galactopyranoside (IPTG) to a 1 m*M* final concentration. Bacterial cells were harvested by centrifugation, resuspended in lysis buffer [20 m*M* Tris-HCl (pH 8.0), 500 m*M* NaCl, and 10 m*M* imidazole], and then lysed by sonication. The soluble tagged protein was collected in the supernatant following centrifugation of the cell homogenate and purified on a Ni^{2+}-NTA-Sepharose column according to published protocols.[60] Recombinant tobacco etch virus (TEV) protease was used to digest the fusion affinity tag from the target protein. After desalting into 20 m*M* NaP_i (pH 6.0), final purification was achieved by ResourceS cation-exchange chromatography followed by size exclusion chromatography (GE Healthcare). The purified protein was concentrated to 10 mg mL^{-1}, washed with 10 μ*M* $ZnCl_2$ and buffer exchanged by ultrafiltration into 20 m*M* NaP_i (pH 6.0), 200 m*M* NaCl, and stored at 4°C for further use.

Overexpression vectors for gluthione S-transferase (GST) pulldown assays were constructed as follows. Gene fragments encoding the entire open reading frames of CT398 (residues 1–254), CT561 (residues 1–223), CT718 (residues 1–174) and RpoN (residues 1–436) were amplified from *C. trachomatis* (serovar L2 434/Bu) genomic DNA via PCR, digested with *BamHI* and *NotI* and subcloned into the expression plasmid pT7HmT.[60] Upon DNA sequence confirmation, CT398/pT7HmT was digested with *NcoI* and *NotI*, so as to retain the His_6 affinity tag, gel purified, and cloned into pACYC-DUET (Novagen). Sequence confirmed vectors encoding CT561, CT718 and RpoN were digested with *BamHI* and *NotI* and cloned into pGEX-4T1 (generously provided by Dr. Magnus Hook, Texas A&M Health Science Center).

Each of the three pGEX-4T1 vectors (CT561, CT718, and RpoN) were transformed into BL21(DE3) *E. coli* cells, both individually and in tandem with CT398/pACYC-DUET. Expression and lysis of each cotransformant was accomplished in a manner similar to that described above. Clarified supernatants were applied to Ni^{2+}-NTA-Sepharose column and extensively washed [20 m*M* Tris-HCl (pH 8.0), 500 m*M* NaCl, and 50 m*M* Imidazole]. Loading and elution fractions were analyzed by immunoblot using α-CT398 and α-GST antibodies (Thermoscientific).

Construction of xpTL2-TAP shuttle vector

The shuttle vector pTL2-TAP is a modified version of pASK-GFP-L2.[46] The multiple cloning site of pASK-GFP-L2, including GFP, was replaced with a single *Age*I cloning site to create pTL2. The *Age*I site was positioned such that any insert would provide the start codon immediately downstream of the *Age*I junction. To make pTL2-TAP, a tandem affinity pulldown (TAP) tag-encoding sequence (codon-optimized for *C. trachomatis*) was inserted at the *Age*I through ligation independent cloning.[59] The TAP tag amino acid sequence is as follows: M**DYKDDDDK**G SAASWSHPQFEKGGGSGGGSGGGSWSHPQFEKG SAASAS, with the FLAG epitope in bold and both Strep-tag II epitopes underlined. Also, the *Age*I site was recreated at the 5′ end of the TAP tag, but not at the 3′ end where it was replaced with a *Nhe*I site. With pTL2-TAP, a gene of interest inserted at the *Age*I site will have a C-terminal TAP tag when expressed, while a gene inserted at the *Nhe*I site will have an N-terminal TAP tag. For this study, the RpoN ortholog from *C. trachomatis* L2 434/Bu (CTL0873) was inserted into the *Nhe*I site, creating pTL2-RpoN-TAPN.

Crystallization

Recombinant *C. trachomatis* CT398 was crystallized by vapor diffusion in Compact Jr. (Emerald Biosystems) sitting drop plates at 20°C. Specifically, 0.5 μL of protein solution [10 mg mL^{-1} in 20 m*M* NaP_i (pH 6.0), 200 m*M* NaCl) was mixed with 0.5 μL of reservoir solution containing 200 m*M* ammonium acetate, 100 m*M* sodium acetate (pH 4.0), and 15% PEG 4K, from the ProPlex HT screen condition B7 (Molecular Dimensions), and equilibrated against 75 μL of the latter. Single block-shaped crystals appeared after 2 days and continued to grow for ~5 days. Crystals were flash-cooled in a cryoprotectant solution consisting of mother liquor with 20% (*v*/*v*) ethylene glycol.

Diffraction data collection, structure determination, refinement, and analysis

X-ray diffraction data were collected at both 1.000 Å and 1.283 Å at 100K using a Dectris Pilatus 6M pixel array detector at beamline 17ID at the APS IMCA-CAT (Table I). Following data collection, partial reflections were integrated with XDS.[61] Laue class analysis and data scaling were performed with Aimless,[20] which suggested the space group was *C*2.

Experimental phase information was obtained for the CT398 structure by Zn-SAD using the SHELX suite.[62] SHELXD was used to find two unique Zn atoms. SHELXE was used to choose the correct hand of the heavy-atom positions and to prepare solvent-flattened maps from the experimental SAD phases. Phases were combined with the complete 2.12 Å native diffraction dataset using CAD.[63] Buccaneer[64] was used to trace 456/514 of the expected amino acids from the combined experimental maps.

Structure refinement was carried out using Phenix.[65,66] One round of individual coordinates and isotropic atomic displacement factor refinement was conducted, and the refined model was used to calculate both $2F_o - F_c$ and $F_o - F_c$ difference maps. These maps were used to iteratively improve the model by manual building with Coot[67,68] followed by subsequent refinement cycles. TLS refinement[69] was

incorporated in the final stages to model anisotropic atomic displacement parameters. Ordered solvent molecules were added according to the default criteria of Phenix and inspected manually using Coot prior to model completion. Additional information and refinement statistics are presented in Table I. Regions of poor map quality exist within the C-terminus that prevented the modeling of residues 234–254.

BACTH vector construction

Gene fragments encoding the entire open reading frames of CT398, CT609 (RpoN or σ^{54}), CT561 (CdsL), CT669 (CdsN), CT717 (FliI) and CT718 (FliH) were amplified from *C. trachomatis* serovar L2 434/Bu genomic DNA via PCR and subcloned (using restriction sites *BamHI* and *KpnI*) into vectors either upstream (pKNT25/pUT18) or downstream (pKT25/pUT18C) of the two catalytic domains (T25/T18) of *B. pertussis* adenylate cyclase.[36] Any natural *BamHI* or *KpnI* restriction sites within open reading frames of the targeted genes were silently mutated using the megaprimer PCR method[70] so as to not disrupt the resulting amino acid sequence. All constructs were verified to be free of mutations and in frame with the T25/T18 domain by DNA sequencing.

BACTH complementation assays

CT398 BACTH vectors were cotransformed into BTH101 *E. coli* cells with various combinations of the RpoN, T3SS and flagellar BACTH vectors described above, plated on LB-Amp-Kan agar plates and incubated at 30°C for 24–36 hours. Interaction efficiencies between hybrid proteins were determined by growth and color change on minimal media (M63) plates. For this measurement, multiple colonies were used to inoculate LB-Amp-Kan liquid media supplemented with 0.5 m*M* IPTG at 30°C for 12–18 hours. For each interaction combination, 5 μL of turbid culture was plated on M63 media supplemented with Amp (50 μg mL^{-1}), Kan (25 μg mL^{-1}), IPTG (0.5 m*M*) and X-Gal (40 μg mL^{-1}) and incubated for 2-3 days at 30°C prior to analysis. The vectors pKT25 and pUT18, each harboring a leucine zipper gene fragment, were used as a positive control, while the same vectors in the absence of any fusion partner were used as the negative control. An additional *Chlamydia*-specific control was also used which included the vectors CT043/pKT25 and $TARP^{1.150}$/pUT18.[39,40]

Transformation of Chlamydia trachomatis

The following protocol was adapted from[50–53] and modified according to observations detailed below. L929 mouse fibroblast (ATCC CCL-1) and *Chlamydia trachomatis* L2 434/Bu cells were used in these studies. Transformed plasmids were purified from (dam^-/dcm^-) *E. coli* (Bioline) using a Qiagen mini prep kit, ethanol precipitated using standard procedures, and resuspended in sterile ddH_2O (~3 μg μL^{-1} final concentration). Infected cells were cultured in RPMI 1640 media supplemented with 300 μg mL^{-1} L-glutamine, 5% (*v*/*v*) heat-inactivated FBS (Tetracycline free) and 10 μg mL^{-1} Gentamycin. Working concentrations of Ampicillin were 1 μg $mL^{-1}$and 5 μg mL^{-1} for selections S1-S3 and S4, respectively, and 1 μg mL^{-1} for Cyclohexamide throughout. With the exception of RPMI media, ice cold buffers were used throughout.

Timeline for transformation

Day 1: 5×10^5 L929 cells were seeded in each well of a six-well plate.

Day 2: *Chlamydia*-DNA transformation reaction (200 μL total volume) was prepared as follows: 10 μL of *Chlamydia trachomatis* L2 434/Bu EB seed stock in 1× SPG (~1×10^7 cfu), 6 μL concentrated (~3 μg μL^{-1}) modified shuttle vector (described above), 100 μL 2× $CaCl_2$ buffer (20 m*M* Tris-HCl pH 7.5, 100 m*M* $CaCl_2$) and ddH_2O to final volume. Reaction was gently pipetted up and down 10 times, incubated for 30 minutes at RT, and diluted into 12 mL of ice cold 1× SPG buffer. Media was aspirated from L929 cells (from **Day 1**), and each well received 2 mL of transformation reaction. The infection was centrifuged for 30 minutes (20°C, 550*g*). Transformation reaction was aspirated and 2 mL of RPMI media (with Cyclohexamide) was added per well. Infection was incubated at 37°C for ~42 hours.

Day 3: Ampicillin (1 μg ml^{-1}) was added to infection (from **Day 2**) between 12 and 16 hours post infection to significantly reduce the amount of untransformed EBs (infection termed *S1*). 5×10^5 L929 cells were seeded in each well of a new six-well plate.

Day 4: Nearly all inclusions (~100% infectivity) from *S1* appeared aberrant. Infected cells were lysed by scraping, collected in a sterile 15 mL conical with ~1 mL of autoclaved glass beads and vortexed for 45 seconds. Cellular debris was clarified by centrifugation for 5 minutes (20°C, 900*g*). Media was aspirated from L929 cells (from **Day 3**), and each well received ~1.5 mL of infected cell lysate (entire *S1* supernatant was used) and 0.5 mL of prewarmed RPMI media. The infection was centrifuged for 30 minutes (20°C, 550*g*). Cell lysate was aspirated and 2 mL of RPMI media (1 μg mL^{-1} of both Cyclohexamide and Ampicillin) was added per well (infection termed *S2*). Infection was incubated at 37°C for ~32 hours. 5×10^5 L929 cells were seeded in each well of a new six-well plate.

Day 5: Nearly all inclusions (~100% infectivity) from *S2* appeared aberrant. Infection was passaged and L929 cells (**Day 4**) were infected in an equivalent manner as described above. After centrifugation, cell lysate was aspirated and 2 mL of RPMI media (1 μg mL^{-1} of both Cyclohexamide and Ampicillin) was added per well (infection termed *S3*). Infection was incubated at 37°C for ~48–72 hours.

Day 7: Approximately 0.5% of cells from *S3* were infected and appeared to contain wild-type inclusions. 5×10^5 L929 cells were seeded in each well of a new six-well plate.

Day 8: Passage of *S3* and infection of L929 cells (**Day 7**) was carried out as described above (termed *S4*) with the exception that 5 μg mL^{-1} of Ampicillin was added to RPMI media prior to incubation at 37°C for ~48 hours.

Day 10: Nearly all inclusions (~100% infectivity) from S4 appeared to be wild-type. At this point transformants were considered to be obtained and expanded for future studies.

Coimmunoprecipitation of protein complexes from C. trachomatis *using Strep-FLAG tandem affinity purification (SF-TAP)*

1 L spinner flask (~1×10^6 L929 cells mL^{-1} in RPMI, 1 and 5 μg mL^{-1} of Cyclohexamide and Ampicillin, respectively) was infected with 150 μL of pTL2-TAPN-RpoN transformed *Chlamydia trachomatis* L2 434/Bu EB seed stock (~1×10^6 cfu/μL, diluted in 1× SPG buffer) and incubated at 37°C. Anhydrotetracycline (ATc, 2 ng ml^{-1}) was added at ~16 hours post infection in order to induce protein expression. All centrifugation steps were comprised of the following parameters unless otherwise noted: 10 minutes, 20°C and 900*g*. Cells were harvested at ~20 hours post infection by centrifugation and washed three times, first with 100 mL of HBSS with Ca^{2+}/Mg^{2+} followed by two rounds of 100 mL of PBS. To preserve weak or transient protein–protein interactions, cells were incubated with 1 m*M* DSP (Thermo) in PBS on a rocker for 30 minutes at RT. Intracellular crosslinking was quenched by the addition of 100 mM Tris-HCl (pH 7.5) for 15 minutes at RT. Cells were collected by centrifugation and resuspended/lysed in 6 mL of 1× RIPA buffer (lacking DTT and EDTA) with 1 m*M* PMSF on ice for 15 minutes. Cells were then further disrupted by sonication on ice (15 seconds on, 60 seconds off; five times) at power setting 6, prior to a high speed (30 minutes, 20°C and 30,966*g*) centrifugation step to remove cellular debris and intact EBs. The supernatant was retained and the SF-TAP immunoprecipitation process was initiated.

Lysed supernatant was combined with 400 μl (50% w/v slurry) Streptavidin sepharose resin (IBA Life Sciences) in a 15 ml conical and incubated overnight at 4°C on an overhead tumbler. Resin was collected by centrifugation and transferred to a 1.5 ml Eppendorf tube, washed three times with 1X TBS (10 mM Tris-HCl (pH 7.4), 150 mM NaCl) + 0.1% NP-40 and eluted with 1 ml of elution buffer (TBS + 2.5 mM desthiobiotin). Eluate was combined with 50 μl of Anti-FLAG M2 Magnetic Beads (Sigma) and incubated at RT for 60 minutes on an overhead tumbler. Beads were washed with 500 μl of TBS three times, 500 μl of low salt buffer (5 mM Tris-HCl (pH 7.4), 10 mM NaCl) two times and eluted with 400 μl of 250 μM of FLAG peptide (Sigma) in TBS. The eluate was then dried to a pellet in a SpeedVac Concentrator, resuspended in 40 μl of ddH_2O and analyzed by SDS-PAGE and immunoblot.

Western blotting

Infected L929 cells were harvested at indicated time points via the following method. Media was aspirated, cells were washed three times with PBS and 50 μL SDS-PAGE Laemmli buffer was added per 10 cm^2 well surface area. Wells were scraped with a pipet tip until the cell layer detached (~10 seconds). An equal volume of PBS was added in order to pipet the detached cells, which were then boiled at 100°C for 5 min prior to analysis (~10% of total lysate was analyzed per lane). The following antibodies were used in this study: rabbit anti-CT398 (affinity purified with recombinant CT398, custom production Proteintech), mouse anti-FLAG (Sigma), goat anti-GST (Thermoscientific), and rat anti-RpoA.

Multiple sequence alignments and figure modeling

Multiple sequence alignments were carried out using ClustalW[71] and aligned with secondary structure elements using ESPRIPT.[72] Three-dimensional structures were superimposed using the Local-Global Alignment method (LGA).[73] Representations of all structures were generated using PyMol.[74] Calculations of electrostatic potentials at the CT398 molecular surface were carried out using DELPHI.[75]

Accession numbers

GenBank accession numbers, *Chlamydia trachomatis* serovar L2 434/Bu (**166154609**); *Treponema pallidum* (**15639485**); *Borrelia garinii* (**51573536**); *Chlamydia pneumoniae* (**15618436**); *Helicobacter pylori* (**15645574**); *Campylobacter jejuni* (**112360051**); *Mycobacterium tuberculosis* (**1261931**); *Clostridium acetobutylicum* (**15894586**); *Aquifex aeolicus* (**15606457**); and *Deinococcus radiodurans* (**6457963**).

HP0958 PDB ID (**3NA7**) was obtained from the PDB.[76] Coordinate and structure factors for CT398 have been deposited in the Protein Data Bank with accession number 4ILO.

Acknowledgment

The authors thank Amy N. Sinclair for technical assistance with the BACTH assays and for critical reading and discussion of the manuscript.

References

1. Darville T (2013) Recognition and treatment of chlamydial infections from birth to adolescence. Adv Exp Med Biol 764:109–122.

2. Thylefors B, Négrel AD, Pararajasegaram R, Dadzie KY (1995) Global data on blindness. Bull World Health Organ 73:115–121.
3. Hu VH, Holland MJ, Burton MJ (2013) Trachoma: protective and pathogenic ocular immune responses to *Chlamydia trachomatis*. PLoS Negl Trop Dis 7:e2020
4. Bardy SL, Ng SY, Jarrell KF (2003) Prokaryotic motility structures. Microbiology 149:295–304.
5. Jarrell KF, McBride MJ (2008) The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol 6:466–476.
6. Cornelis GR (2010) The type III secretion injectisome, a complex nanomachine for intracellular 'toxin' delivery. Biol Chem 391:745–751.
7. Hueck CJ (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62:379–433.
8. Hefty PS, Stephens RS (2007) Chlamydial type III secretion system is encoded on ten operons preceded by sigma 70-like promoter elements. J Bacteriol 189:198–206.
9. Blocker A, Komoriya K, Aizawa S (2003) Type III secretion systems and bacterial flagella: insights into their function from structural similarities. Proc Natl Acad Sci USA 100:3027–3030.
10. Abby SS, Rocha EP (2012) The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. PLoS Genet 8:e1002983
11. Wolf K, Betts HJ, Chellas-Gery B, Hower S, Linton CN, Fields KA (2006) Treatment of *Chlamydia trachomatis* with a small molecule inhibitor of the Yersinia type III secretion system disrupts progression of the chlamydial developmental cycle. Mol Microbiol 61:1543–1555.
12. Mueller KE, Plano GV, Fields KA (2014) New frontiers in type III secretion biology: the Chlamydia perspective. Infect Immun 82:2–9.
13. Scidmore MA, Hackstadt T (2001) Mammalian 14-3-3beta associates with the Chlamydia trachomatis inclusion membrane via its interaction with IncG. Mol Microbiol 39:1638–1650.
14. Hower S, Wolf K, Fields KA (2009) Evidence that CT694 is a novel *Chlamydia trachomatis* T3S substrate capable of functioning during invasion or early cycle development. Mol Microbiol 72:1423–1437.
15. Clifton DR, Fields KA, Grieshaber SS, Dooley CA, Fischer ER, Mead DJ, Carabeo RA, Hackstadt T (2004) A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. Proc Natl Acad Sci USA 101: 10166–10171.
16. Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, Behrens S, Niinikoski A, Mewes HW, Horn M, Rattei T (2009) Sequence-based prediction of type III secreted proteins. PLoS Pathog 5:e1000376
17. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, Mitchell W, Olinger L, Tatusov RL, Zhao Q, Koonin EV, Davis RW (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282:754–759.
18. Stone CB, Bulir DC, Gilchrist JD, Toor RK, Mahony JB (2010) Interactions between flagellar and type III secretion proteins in Chlamydia pneumoniae. BMC Microbiol 10:18
19. Rigden DJ (2011) Ab initio modeling led annotation suggests nucleic acid binding function for many DUFs. Omics 15:431–438.
20. Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Cryst D67:282–292.
21. Evans P (2006) Scaling and assessment of data quality. Acta Cryst D62:72–82.
22. Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol 4:269–275.
23. Weiss MS (2001) Global indicators of X-ray data quality. J Appl Cryst 34:130–135.
24. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. Science 336:1030–1033.
25. Evans P (2012) Biochemistry. Resolving some old problems in protein crystallography. Science 336:986–987.
26. Harding MM (2001) Geometry of metal-ligand interactions in proteins. Acta Cryst D57:401–411.
27. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res 38:W545–W549.
28. Caly DL, O'Toole PW, Moore SA (2010) The 2.2-Å structure of the HP0958 protein from Helicobacter pylori reveals a kinked anti-parallel coiled-coil hairpin domain and a highly conserved ZN-ribbon domain. J Mol Biol 403:405–419.
29. Barta ML, Dickenson NE, Patil M, Keightley A, Wyckoff GJ, Picking WD, Picking WL, Geisbrecht BV (2012) The structures of coiled-coil domains from type III secretion system translocators reveal homology to pore-forming toxins. J Mol Biol 417:395–405.
30. Mandelker D, Gabelli SB, Schmidt-Kittler O, Zhu J, Cheong I, Huang CH, Kinzler KW, Vogelstein B, Amzel LM (2009) A frequent kinase domain mutation that changes the interaction between PI3Kalpha and the membrane. Proc Natl Acad Sci USA 106: 16996–17001.
31. Mason JM, Arndt KM (2004) Coiled coil domains: stability, specificity, and biological implications. Chembiochem 5:170–176.
32. Pereira L, Hoover TR (2005) Stable accumulation of sigma54 in Helicobacter pylori requires the novel protein HP0958. J Bacteriol 187:4463–4469.
33. Ryan KA, Karim N, Worku M, Moore SA, Penn CW, O'Toole PW (2005) HP0958 is an essential motility gene in Helicobacter pylori. FEMA Microbiol Lett 248: 47–55.
34. Rain JC, Selig L, De Reuse H, Battaglia V, Reverdy C, Simon S, Lenzen G, Petel F, Wojcik J, Schachter V, Chemama Y, Labigne A, Legrain P (2001) The protein-protein interaction map of Helicobacter pylori. Nature 409:211–215.
35. Douillard FP, Ryan KA, Caly DL, Hinds J, Witney AA, Husain SE, O'Toole PW (2008) Posttranscriptional regulation of flagellin synthesis in Helicobacter pylori by the RpoN chaperone HP0958. J Bacteriol 190:7975–7984.
36. Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci USA 95:5752–5756.
37. Stone CB, Bulir DC, Emdin CA, Pirie RM, Porfilio EA, Slootstra JW, Mahony JB (2011) Chlamydia pneumoniae CdsL regulates CdsN ATPase activity, and disruption with a peptide mimetic prevents bacterial invasion. Front Microbiol 2:21
38. Karimova G, Dautin N, Ladant D (2005) Interaction network among Escherichia coli membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. J Bacteriol 187:2233–2243.
39. Brinkworth AJ, Malcolm DS, Pedrosa AT, Roguska K, Shahbazian S, Graham JE, Hayward RD, Carabeo RA (2011) Chlamydia trachomatis Slc1 is a type III secretion chaperone that enhances the translocation of its

invasion effector substrate TARP. Mol Microbiol 82: 131–144.
40. Saka HA, Thompson JW, Chen YS, Kumar Y, Dubois LG, Moseley MA, Valdivia RH (2011) Quantitative proteomics reveals metabolic and pathogenic properties of Chlamydia trachomatis developmental forms. Mol Microbiol 82:1185–1203.
41. Pereira LE, Tsang J, Mrázek J, Hoover TR (2011) The zinc-ribbon domain of Helicobacter pylori HP0958: requirement for RpoN accumulation and possible roles of homologs in other bacteria. Microb Inform Exp 1:1–10.
42. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725–738.
43. Pallen MJ, Bailey CM, Beatson SA (2006) Evolutionary links between FliH/YscL-like proteins from bacterial type III secretion systems and second-stalk components of the FoF1 and vacuolar ATPases. Protein Sci 15: 935–941.
44. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:40
45. Balakrishna AM, Hunke C, Gruber G (2012) The structure of subunit E of the Pyrococcus horikoshii OT3 A-ATP synthase gives insight into the elasticity of the peripheral stalk. J Mol Biol 420:155–163.
46. Wickstrum J, Sammons LR, Restivo KN, Hefty PS (2013) Conditional gene expression in Chlamydia trachomatis using the tet system. PLoS One 8:e76743
47. Mathews SA, Volp KM, Timms P (1999) Development of a quantitative gene expression assay for Chlamydia trachomatis identified temporal expression of sigma factors. FEBS Lett 458:354–358.
48. Mathews SA, Timms P (2000) Identification and mapping of sigma-54 promoters in Chlamydia trachomatis. J Bacteriol 182:6239–6242.
49. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R (1999) Helicobacter pylori virulence and genetic geography. Science 284:1328–1333.
50. Wang Y, Kahane S, Cutcliffe LT, Skilton RJ, Lambden PR, Clarke IN (2011) Development of a transformation system for Chlamydia trachomatis: restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector. PLoS Pathog 7:e1002258
51. Agaisse H, Derré I (2013) A C. trachomatis cloning vector and the generation of C. trachomatis strains expressing fluorescent proteins under the control of a C. trachomatis promoter. PLoS One 8:e57090
52. Song L, Carlson JH, Whitmire WM, Kari L, Virtaneva K, Sturdevant DE, Watkins H, Zhou B, Sturdevant GL, Porcella SF, McClarty G, Caldwell HD (2013) Chlamydia trachomatis plasmid-encoded Pgp4 is a transcriptional regulator of virulence-associated genes. Infect Immun 81:636–644.
53. Gong S, Yang Z, Lei L, Shen L, Zhong G (2013) Characterization of Chlamydia trachomatis plasmid-encoded open reading frames. J Bacteriol 195:3819–3826.
54. Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, Picking WD, Liu J (2015) Visualization of the type III secretion sorting platform of Shigella flexneri. Proc Natl Acad Sci USA 112:1047–1052.
55. Fields KA, Mead DJ, Dooley CA, Hackstadt T (2003) Chlamydia trachomatis type III secretion: evidence for a functional apparatus during early-cycle development. Mol Microbiol 48:671–683.
56. Menard R, Sansonetti P, Parsot C, Vasselon T (1994) Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of S. flexneri. Cell 79: 515–525.
57. Mavris M, Sansonetti PJ, Parsot C (2002) Identification of the cis-acting site involved in activation of promoters regulated by activity of the type III secretion apparatus in Shigella flexneri. J Bacteriol 184: 6751–6759.
58. Mavris M, Page AL, Tournebize R, Demers B, Sansonetti P, Parsot C (2002) Regulation of transcription by the activity of the Shigella flexneri type III secretion apparatus. Mol Microbiol 43: 1543–1553.
59. Qin H, Hu J, Hua Y, Challa SV, Cross TA, Gao FP (2008) Construction of a series of vectors for high throughput cloning and expression screening of membrane proteins from Mycobacterium tuberculosis. BMC Biotechnol 8:51
60. Geisbrecht B, Bouyain S, Pop M (2006) An optimized system for expression and purification of secreted bacterial proteins. Protein Expr Purif 46:23–32.
61. Kabsch W (2010) XDS. Acta Cryst D Biol Crystallogr 66:125–132.
62. Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Cryst D Biol Crystallogr 66: 479–485.
63. Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: programs for protein crystallography. Acta Crystallography D Biol Crystallogr 50: 760–763.
64. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Cryst 62:1002–1011.
65. Adams P, Grosse-Kunstleve R, Hung L, Ioerger T, McCoy A, Moriarty N, Read R, Sacchettini J, Sauter N, Terwilliger T (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Cryst D Biol Crystallogr 58:1948–1954.
66. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst D Biol Crystallogr 66: 213–221.
67. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Cryst D Biol Crystallogr 60:2126–2132.
68. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Cryst D Biol Crystallogr 66:486–501.
69. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Cryst D Biol Crystallogr 62: 439–450.
70. Brøns-Poulsen J, Nøhr J, Larsen L (2002) Megaprimer method for polymerase chain reaction-mediated generation of specific mutations in DNA. Methods Mol Biol 182:71–76.
71. Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.
72. Gouet P, Courcelle E, Stuart D, Métoz F (1999) ESPript: analysis of multiple sequence alignments in postscript. Bioinformatics 15:305–308.

73. Zemla A (2003) LGA: A method for finding 3D similarities in protein structures. Nucleic Acids Res 31:3370–3374.
74. DeLano WL (2002) The PyMOL Molecular Graphics System. **2009**, http://www.pymol.org
75. Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem 23:128–137.
76. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank. A computer-based archival file for macromolecular structures. Eur J Biochem 80:319–324.