ARTICLE

Structural basis for the homotypic fusion of chlamydial inclusions by the SNARE-like protein IncA

Gino Cingolani[1,2], Michael McCauley [3,5], Anna Lobley[3], Alexander J. Bryer[4], Jordan Wesolowski[3], Deanna L. Greco[4], Ravi K. Lokareddy[1], Erik Ronzone[3,6], Juan R. Perilla [4] & Fabienne Paumet[3]

Many intracellular bacteria, including *Chlamydia*, establish a parasitic membrane-bound organelle inside the host cell that is essential for the bacteria's survival. *Chlamydia trachomatis* forms inclusions that are decorated with poorly characterized membrane proteins known as Incs. The prototypical Inc, called IncA, enhances *Chlamydia* pathogenicity by promoting the homotypic fusion of inclusions and shares structural and functional similarity to eukaryotic SNAREs. Here, we present the atomic structure of the cytoplasmic domain of IncA, which reveals a non-canonical four-helix bundle. Structure-based mutagenesis, molecular dynamics simulation, and functional cellular assays identify an intramolecular clamp that is essential for IncA-mediated homotypic membrane fusion during infection.

[1] Thomas Jefferson University, Department of Biochemistry and Molecular Biology, Philadelphia, PA 19107, USA. [2] Institute of Biomembranes and Bioenergetics, National Research Council, Via Amendola 165/A, 70126 Bari, Italy. [3] Thomas Jefferson University, Department of Microbiology and Immunology, Philadelphia, PA 19107, USA. [4] The University of Delaware, Department of Chemistry and Biochemistry, Newark, DE 19716, USA. [5] Present address: Janssen Research and Development, Spring House, PA 19477, USA. [6] Present address: VUE Health, Boston, MA 02110, USA. Correspondence and requests for materials should be addressed to G.C. (email: gino.cingolani@jefferson.edu) or to F.P. (email: fabienne.paumet@jefferson.edu)

The obligate intracellular pathogen *Chlamydia trachomatis* is the most frequent cause of bacterial sexually transmitted disease and infectious blindness worldwide, yet it is still considered a neglected disease pathogen by the World Health Organization[1]. *Chlamydia's* life cycle depends on the establishment of a fast-growing parasitic organelle inside the host cell called the "inclusion", the development of which is poorly understood. Inclusion membranes are decorated with ~60 transmembrane Inc proteins that are known to directly interact with host cell components and play a critical role in sustaining *Chlamydia's* life cycle[2–4]. Despite their importance, Incs have remained relatively uncharacterized and little is known about their function at the molecular level. The low sequence conservation and minimal similarity that Incs share with other proteins limits the usefulness of conventional bioinformatic tools to predict their structure and function. The best-characterized chlamydial Inc is IncA. It contains two extended 3,4-hydrophobic heptad repeat segments similar to the coiled-coil regions of the eukaryotic "SNAREs" (soluble N-ethylmaleimide-sensitive factor attachment receptors), which are proteins involved in cellular transport and membrane fusion[5–9]. IncA is involved in the homotypic fusion of chlamydial inclusions[10,11] and is the best-characterized bacterial SNARE-like protein to date.

During infection, each *Chlamydia* bacterium establishes its own inclusion inside a host cell. At high multiplicities of infection (MOI), cells contain multiple inclusions that ultimately fuse together to form one large inclusion per cell. This homotypic fusion event is important for the pathogenicity of *C. trachomatis* because natural non-fusogenic IncA mutants are replication-defective in humans and cause significantly milder disease compared with patients infected with normal fusogenic strains[12–14]. Measurements of relative chlamydial rRNA quantities in the multiplying organisms have also revealed that non-fusogenic strains grow more slowly than fusogenic strains of *Chlamydia*, which is supported by a reduced rate of protein synthesis and decreased multiplication efficiency[15]. A direct role for IncA in membrane fusion was demonstrated using microinjection of anti-IncA antibody during infection, which resulted in multiple inclusions that were unable to undergo homotypic fusion[10]. Cells infected with an IncA-deficient strain of *C. trachomatis* similarly displayed multiple inclusions at a high multiplicity of infection, further establishing that IncA is required for the homotypic fusion of inclusions[11,16].

IncA localizes to the inclusion membrane where it can potentially interact with host and inclusion-associated proteins. Yeast two-hybrid analysis demonstrated that IncA can bind itself[10], while immunoprecipitation also showed that IncA associates *in trans* when present on opposite membranes in a cell[5,8]. In *Chlamydia*-infected HeLa cells, the expression of ectopic IncA on the endoplasmic reticulum (ER) membrane impacts inclusion integrity and ER morphology, suggesting that IncA present on the ER interacts homotypically with IncA expressed on the inclusion and induces the fusion of both compartments[5]. While it is well-established that IncA is involved in the homotypic fusion of inclusions, the molecular mechanism of its fusogenic activity remains unknown.

Here, we describe the crystal structure of IncA, which we probed using biophysical, computational, and functional methods. We demonstrate that IncA folds into a stable, non-canonical four-helix bundle that is maintained as a monomer by intramolecular interactions. We also show that the monomeric conformation of IncA is critical for its activity during membrane fusion. Our work sheds light on a class of bacterial transmembrane proteins that control membrane fusion during infection, which is critical for *Chlamydia* pathogenicity.

Results

IncA folds into a non-canonical four-helix bundle. IncA is a protein of 273 amino acids composed of a short, cytoplasmic N-terminal moiety (residues 1–34), a bilobed transmembrane domain (residues 35–84), and a long cytoplasmic C-terminal domain that ends in a tail with low complexity (residues 247–273) (Fig. 1a). The bacterial SNARE-like domains are found in the cytoplasmic C-terminal domain[5,7]. To better understand the mechanism of IncA-mediated membrane fusion, we generated high-quality crystals of a chymotryptic fragment of IncA spanning residues 87–246 (IncA$_{87-246}$) (Supplementary Fig. 1b) and determined a crystal structure of IncA$_{87-246}$ to an $R_{work/free}$ of 14.1/16.8% at 1.12 Å resolution (Fig. 1b and Table 1). The electron density for IncA$_{87-246}$, including the six N-terminal histidines of the affinity tag, is exceptionally clear (Supplementary Fig. 2), consistent with the low B-factor of the crystal structure (~10.8 Å^2). We found that IncA$_{87-246}$ adopts an asymmetric and slightly blocky conformation, somewhat similar to a four-helix bundle (Fig. 1b). Both in crystal and in solution, IncA$_{87-246}$ exists as a monomer (Supplementary Fig. 3a, b and Supplementary Table 1), as previously observed for the full-length cytosolic domain of IncA (ΔTMD-IncA - Supplementary Fig. 1a), that sediments at equilibrium as a single species of 23.9 ± 0.8 kDa[8]. Altogether, these results suggest that the C-terminal protease-sensitive tail
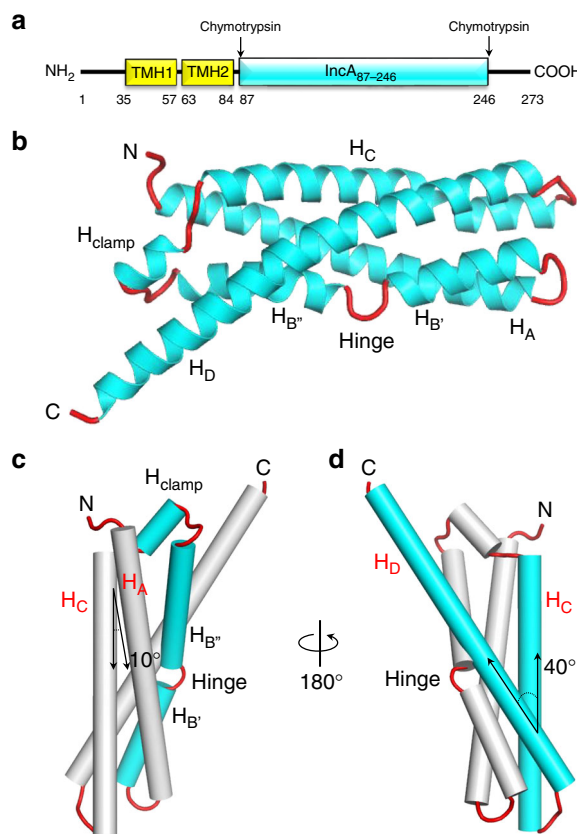

Fig. 1 The cytosolic domain of IncA folds into a non-canonical α-helical bundle. **a** Schematic diagram of *C. trachomatis* IncA domain organization. Transmembrane helices (TMH) were predicted using the HMMTOP server[50]. The stable core encompassing residues 87–246 (IncA$_{87-246}$) was identified by limited proteolysis of the cytosolic domain as described[8]. **b** Ribbon diagram of IncA$_{87-246}$ with α-helices and random coiled linkers colored in cyan and red, respectively. **c**, **d** Two rotated views of IncA$_{87-246}$, orthogonal to the representation in **b**, with α-helices shown as cylinders. **c** highlights the segmented structure of the helix H$_B$, while **d** shows the relative angle between helices H$_C$ and H$_D$

Table 1 Crystallographic data collection, phasing, and refinement statistics

	IncA$_{87-246}$	(IncA$_{87-246}$)^{NaI}	IncA$_{87-246(G144A)}$
Data collection			
X-ray source	SSRL 9–2	MicroMax-007 HF	MicroMax-007 HF
Detector	Pilatus 6M PAD	Pilatus3 R 200K	Pilatus3 R 200K
Space group	P2$_1$	P2$_1$	P1
Cell dimensions			
a, b, c (Å)	38.4, 48.8, 41.7	38.4, 48.7, 41.7	41.0, 43.5, 45.8
α, β, γ (°)	90.0, 103.7, 90.0	90.0, 103.8, 90.0	92.9, 95.9, 93.9
Wavelength (Å)	0.98	1.54	1.54
Resolution (Å)	15–1.12 (1.16–1.12)	15–1.80 (1.83–1.80)	15–1.95 (2.02–1.95)
No. reflections (tot/unique)	875,396/ 52,606	1,788,976/ 13,917	282,344/ 20,042
R_{sym}	5.5 (29.8)	7.6 (22.4)	4.8 (13.1)
R_{pim}	2.3 (14.8)	1.9 (2.2)	4.8 (13.0)
Mean$I$ / σI	47.9 (4.7)	104.5 (23.9)	20.4 (5.8)
CC1/2	0.963	0.997	0.950
Completeness (%)	91.3 (55.0)	99.5 (99.4)	88.1 (66.3)
Redundancy	7.3 (4.5)	16.0 (12.4)	1.5 (1.5)
Wilson B-factor (Å^2)	10.8	13.3	13.4
SAD phasing			
Number Iodine sites		16	
FOM		0.37	
Corr. of local RMS density		0.61	
Refinement			
PDB ID	6E7E		6E6A
Resolution (Å)	15–1.12		15–1.95
No. reflections	50,533		20,013
R_{work}/R_{free}[a]	14.1/16.8		16.9/21.2
No. of complexes in AU	1		2
No. of protein atoms	1375		2650
Ramachandran (favored/allowed/outliers)	100/0.0/0.0		99.7/0.3/0.0
R.M.S.D. from ideality			
Bond lengths (Å)	0.012		0.005
Bond angles (°)	1.649		0.860
MolProbity Score/ranking[b]	1.15/94th percentile		1.24/99th percentile
MolProbity ClashScore/ ranking[b]	3.65/5th percentile		4.71/98th percentile

Values in parentheses are for highest-resolution shells
[a]R_{free} was calculated using ~5% randomly selected reflections
[b]Percentile ranking relative to X-ray structures solved at similar resolution

(res. 247–273) does not promote IncA self-association in vitro. The tertiary structure of IncA$_{87-246}$ consists of four down-up-down-up antiparallel α-helices, named H$_A$-H$_D$. However, IncA$_{87-246}$ deviates from canonical four-helix bundles in at least three aspects. First, the helix H$_B$ is interrupted at position 144 by a glycine that generates two shorter helices, named H$_{B'}$ and H$_{B''}$ (Fig. 1b, c). We termed this break in helicity as the 'hinge'. Second, the loop connecting helices H$_B$ and H$_C$ contains a short insertion helix (res. 165–169), which we termed the 'clamp' helix (H$_{clamp}$) (Fig. 1b, c). The clamp makes numerous contacts with H$_A$, H$_{B''}$, H$_C$, and H$_D$ that account for a total of eight hydrogen bonds and 94 non-bonded interactions (Supplementary Fig. 4 and Supplementary Table 2). Third, while helices H$_A$ and H$_C$ run parallel to each other making an acute angle of ~ 10° and bonding throughout their entire length (Fig. 1c), the longest helix H$_D$ makes a 40° angle from helix H$_C$ pointing away from the H$_{clamp}$ (Fig. 1d) and gives an asymmetric appearance to the helical bundle. Consistent with the unique four-helix-composition of IncA$_{87-246}$, a search for structural relatives using DALI[17] did not identify four-helix bundles with high structural similarities to IncA$_{87-246}$ despite the abundance of this fold in nature. Instead, DALI found IncA$_{87-246}$ bears structural similarity (Z-score = 9.0) to the talin-HIP1/R/Sla2p actin-tethering C-terminal homology (THATCH) domain core of the Huntingtin Interacting Protein 12, (PDB ID 1R0D) (Supplementary Fig. 5), which superimposes to IncA$_{87-246}$ with a Cα root-mean-square deviation

(RMSD) ~ 3.6 Å. This helical bundle, involved in the association between actin and clathrin-coated structures at the plasma membrane and trans-Golgi network[18], also contains a clamp helix between H$_B$ and H$_C$, but, unlike IncA$_{87-246}$, its C-terminal helix H$_D$ is split into two α-helices, H$_D$ and H$_E$. Thus, the high-resolution crystal structure of the cytosolic domain of IncA reveals a non-canonical helical bundle.

Intramolecular contacts maintain IncA as a monomer. To investigate how the structural determinants in IncA that deviate from a classical four-helix bundle topology affect protein flexibility and the ability to mediate homotypic fusion, we analyzed the anisotropically-refined B-factor of IncA$_{87-246}$ (Fig. 2a). The hinge and the amino acids preceding H$_{clamp}$ were found to have significantly higher than average B-factors (~ 26 Å^2 vs. ~ 16 Å^2) (Fig. 2a), possibly underscoring intrinsic flexibility. To characterize these regions, we first generated a G144A hinge mutant variant and found that IncA$_{87-246(G144A)}$ (Supplementary Fig. 1c) remains monomeric in solution (Supplementary Fig. 3c, d and Supplementary Table 1) and it has comparable structural stability as IncA$_{87-246}$ (Supplementary Fig. 6). IncA$_{87-246(G144A)}$ crystallized in a triclinic space group with two IncA$_{87-246(G144A)}$ protomers in the unit cell. Though this crystal form did not diffract as well as the previous one, we were able to collect 88.1% complete data to 1.95 Å resolution (Table 1) and determine an accurate atomic model of
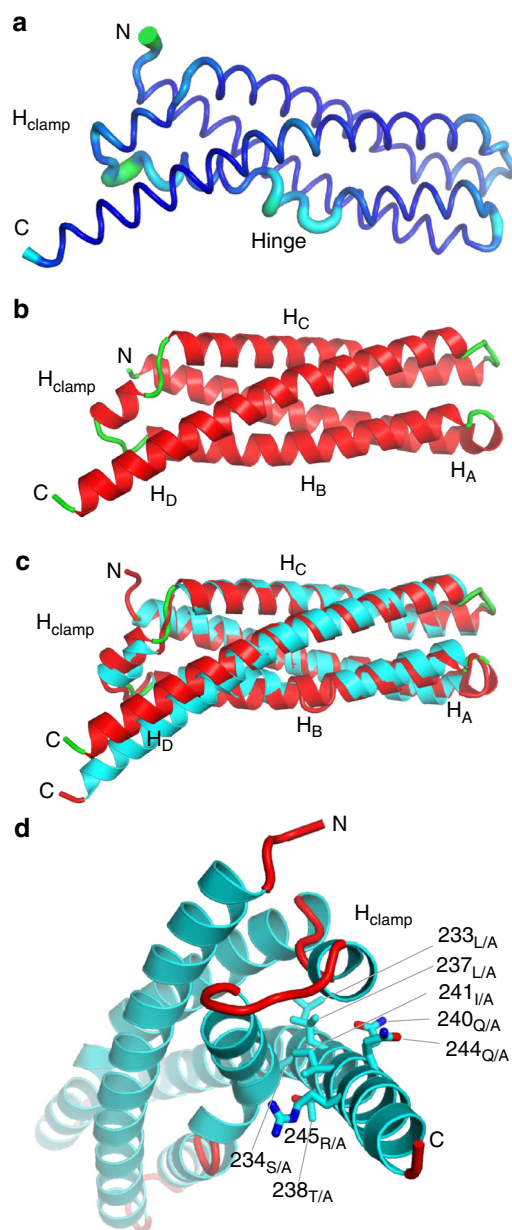
3

Fig. 2 Analysis of the conformational flexibility of IncA$_{87-246}$ reveals two features. **a** Graphic representation of IncA$_{87-246}$ anisotropically refined B-factor plotted onto the 3D-structure: the diameter of the tube is proportional to the B-factor. The hinge region and the moiety C-terminal of the H$_{clamp}$ (in cyan) deviate by more than two standard deviations from the mean B-factor. **b** Crystal structure of IncA$_{87-246(G144A)}$ (protomer A), refined at 1.95 Å resolution (Table 1) with α-helices and random coiled linkers colored in red and green, respectively. **c** Secondary structure superimposition of IncA$_{87-246(G144A)}$ (green and red) with IncA$_{87-246}$ (red and cyan). **d** Magnified view of IncA$_{87-246}$ viewed down the main helical axis. Residues mutated to Ala in IncA$_{87-246(polyA)}$ are shown as sticks

IncA$_{87-246(G144A)}$ using molecular replacement. The structure, solved to a R$_{work/free}$ of 16.9/21.2 at 1.95 Å resolution (Table 1), contains two four-helix bundles (Fig. 2b) where the interrupted α-helix H$_B$ of IncA$_{87-246}$ (Fig. 1b) is replaced by a straight continuous helix. Interestingly, the two IncA$_{87-246(G144A)}$ protomers in the triclinic unit cell are noticeably dissimilar (RMSD 1.65 Å), with protomer B characterized by a local unfolding of helix A between residues ^{100}VGSL^{103} (Supplementary Fig. 7). In both protomers,

this region is not implicated in crystal contacts, which suggests IncA$_{87-246(G144A)}$ may exist in solution in different microstates, also populated in the crystal lattice. Secondary structure super-imposition of IncA$_{87-246}$ with the reference IncA$_{87-246(G144A)}$ protomer A reveals the G144A mutation results in significant conformational changes throughout the molecule (RMSD 1.68 Å) (Fig. 2c), especially in the helix H$_D$, which is shifted upwards in IncA$_{87-246(G144A)}$. Thus, the G144A mutation in the segmented helix B of IncA plays a global structural role in the architecture of the bundle.

To probe the second region of IncA that has a higher-than-average B-factor (Fig. 2a), we generated a mutant, IncA$_{87-246(polyA)}$, where eight residues in H$_D$ that make contact with the helices surrounding the H$_{clamp}$ are mutated to alanine, namely L233A, S234A, L237A, T238A, Q240A, I241A, Q244A, and R245A (Fig. 2d and Supplementary Fig. 1d). Although these mutations result in a loss of 12 intramolecular bonds (Supplementary Fig. 8), IncA$_{87-246(polyA)}$ had comparable structural stability as IncA$_{87-246}$ (Supplementary Fig. 6). Unlike the previous two IncA constructs, IncA$_{87-246(polyA)}$ failed to crystallize, limiting our understanding of its structure. In solution, when analyzed by AUC-SV, IncA$_{87-246(polyA)}$ migrated as a slightly larger species than IncA$_{87-246}$ or IncA$_{87-246(G144A)}$, possibly consistent with a trimer or an elongated dimer (Supplementary Fig. 3e, f and Supplementary Table 1). Interestingly, IncA$_{87-237}$, a shorter deletion construct of IncA that lacks many of the same residues on Helix D mutated in IncA$_{87-246(polyA)}$ (Supplementary Fig. 1f), is also predominantly dimeric in solution[8]. Thus, when the intramolecular contacts generated by helix H$_D$ are broken, then a higher oligomeric species spontaneously forms, suggesting that the non-canonical conformation of the H$_{clamp}$ may function by locking the helical core of IncA in a monomeric conformation that prevents IncA from self-assembling on the inclusion membrane.

The IncA monomer is highly stable. To determine if the crys-tallographic structure of IncA$_{87-246}$ represents a metastable intermediate, we subjected the 1.12 Å crystallographic model of IncA$_{87-246}$ to equilibrium molecular dynamics simulations, along with models of IncA$_{87-237}$ and IncA$_{87-246(G144A)}$ (Supplementary Fig. 1). These three systems were subjected to 4 μs of equilibrium sampling to explore conformational flexibility and stability of IncA, and to probe the effects of the hinge, H$_{clamp}$, and the C-terminus of helix H$_D$ on dynamics and structure. Our simu-lations showed that the structure of IncA$_{87-246}$ is highly stable and is unlikely to be a metastable intermediate. This conclusion is supported by the rigid, unchanging conformation of IncA$_{87-246}$ throughout the equilibrium sampling at physiological conditions (310 K, 150 mM NaCl) and over the entirety of the simulation (Supplementary movie 1). In contrast, IncA$_{87-237}$, which lacks nine additional C-terminal residues, underwent a structure-wide increase in root-mean-square fluctuation (RMSF) compared to both IncA$_{87-246}$ and IncA$_{87-246(G144A)}$ (Fig. 3a), and experienced an eventual conformational change after ~3.2 μs of sampling (Fig. 3b and Supplementary movie 2). This conformation change involved a repositioning of H$_{clamp}$ away from helix H$_D$, toward helix H$_A$, where it ultimately formed H-bonds with residues 94 to 100 of H$_A$, hereinafter referred to as H$_{A(94-100)}$. Hydrogen bonds were identified using a contact-analysis with a cutoff at 3.2 Å, between hydrogen, oxygen, and nitrogen atoms of the H$_{clamp}$ (res. 165–169) and H$_{A(94-100)}$. This hydrogen bond analysis was conducted for all three systems (Fig. 3c). Hydrogen bonds in IncA$_{87-237}$ appeared transiently around 2 μs following positional fluctuations of H$_{A(94-100)}$, and they persisted once H$_{A(94-100)}$ repositioned permanently at 3.2 μs. Beyond the notable structural
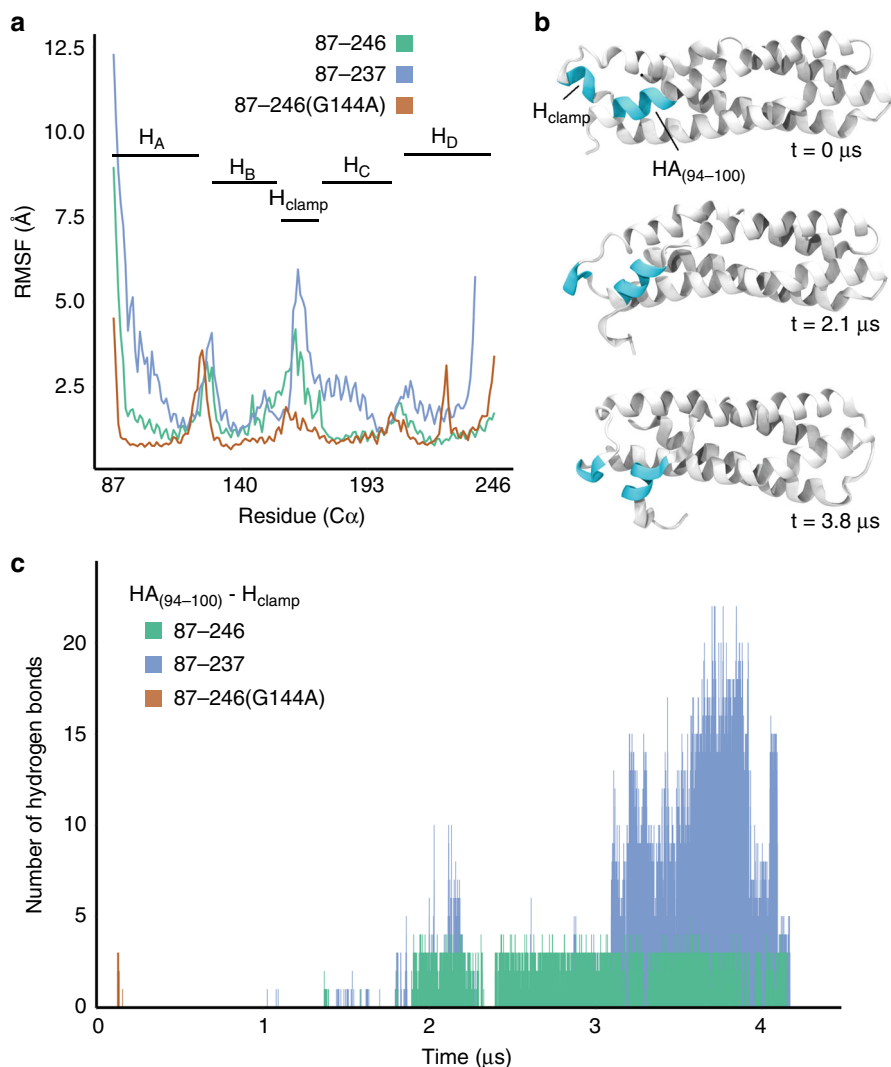
Fig. 3 Molecular dynamics simulations show that IncA$_{87-246}$ monomer is stable. **a** Root-mean-square fluctuation (RMSF) analysis of the three simulations performed: IncA$_{87-246}$ (green line), IncA$_{87-237}$ (blue line), and IncA$_{87-246(G144A)}$ (orange line). Prior to calculating RMSF, each of the three trajectories was aligned by minimizing the root-mean-square deviation (RMSD) of backbone atoms between the equilibrated structure and every subsequent frame. **b** Three snapshots ($t = 0\,\mu s$, $t = 2.1\,\mu s$, $t = 3.8\,\mu s$) of the observed conformational change during the IncA$_{87-237}$ simulation. The two regions of interest, H$_{clamp}$ and H$_{A(94-100)}$, are colored in cyan. **c** Bar plot of the number of hydrogen bonds occurring between the H$_{clamp}$ and H$_{A(94-100)}$ over the course of all three simulations. Hydrogen bonds were identified using a contact-analysis between hydrogen, oxygen, and nitrogen atoms within a cutoff of 3.2 Å

instability compared to IncA$_{87-246}$, this result implies that the C-terminus of helix H$_D$ (res. 237–246) plays a significant role in regulating the conformation of IncA$_{87-246}$. The RMSF analysis shows that the H$_{clamp}$ is highly flexible in all three structures (Fig. 3a). Moreover, the relatively higher magnitude of RMSF in the H$_{clamp}$ region of IncA$_{87-237}$ compared with that of IncA$_{87-246}$ and IncA$_{87-246(G144A)}$ again makes the structural effect of the C-terminus of helix H$_D$ apparent.

Structural changes of IncA$_{87-246(G144A)}$ during simulation are minimal compared to simulations of IncA$_{87-246}$ (Supplementary movies 3 and 4). The latter observation supports the hypothesis that helix H$_D$ is a key determinant of structural stability. Furthermore, noting both the lower RMSF of IncA$_{87-246(G144A)}$ (Fig. 3a) and the dissimilarity in the position of H$_D$ among IncA$_{87-246(G144A)}$ and IncA$_{87-246}$ (Fig. 2c), it is evident that the orientation of helix H$_D$ with respect to H$_{clamp}$ is coupled to the overall stability of each construct. Closer proximity between H$_{clamp}$ and H$_D$ in IncA$_{87-246(G144A)}$ prevents the H$_{clamp}$ from

forming hydrogen bonds with H$_{A(94-100)}$ resulting in a more stable conformation (Fig. 3c). The H$_{clamp}$ of IncA$_{87-246}$, with a significantly higher RMSF than that of IncA$_{87-246(G144A)}$, forms hydrogen bonds with H$_{A(94-100)}$ after 2 μs of sampling. While less-numerous and denoting no clear structural significance compared with those of IncA$_{87-237}$, the lack of hydrogen-bonding in IncA$_{87-246(G144A)}$ implies that the position of the C-terminus of H$_D$ relative to the H$_{clamp}$ directly modulates the behavior of H$_{clamp}$. This finding suggests that the overall flexibility of the H$_{clamp}$, potentially its ability to recognize regulatory proteins, is sensitive to the network of hydrogen bonds it can form with adjacent regions of H$_D$ and H$_A$.

Markov State models (MSMs) provide information regarding the thermodynamic stability of a protein and are typically used to characterize long timescale dynamic modes[19], such as folding coordinates. For the trajectories in the present study (4.18 μs for IncA$_{87-246}$ and 5.80 μs for IncA$_{87-237}$), the high temporal resolution of the input datasets (40,000 states with each

representing an interval of 100 ps), coupled with an appropriate lag time of 10 ns, resulted in four- and five-state models for $IncA_{87-237}$ and $IncA_{87-246}$, respectively. Models were scored and ranked for $IncA_{87-237}$ and $IncA_{87-246}$ separately, according to the generalized mixed Rayleigh quotient (GMRQ); the latter, a scalar value measuring the fit quality of a given MSM's learned dynamical eigenvectors against its input data (training score) or new data (testing score)[20]. Importantly, the resulting Markov states (shown in Supplementary Fig. 9a) were different for $IncA_{87-237}$ and $IncA_{87-246}$. In particular, the resulting MSM of $IncA_{87-246}$ describes just four states that are structurally similar beyond fluctuation at the termini and H_{clamp}, in contrast to the five models observed for $IncA_{87-237}$, which demonstrate the construct's flexible conformation observed during simulation. For both of the top-scoring models, the longest implied relaxation time was <200 ns (Supplementary Fig. 9a), asserting that each of the identified Markov state interconverts on relatively short timescales. The free energy landscape (Supplementary Fig. 9b) comparing the two MSMs shows the RMSD of each state mapped to its computed free energy (ΔG), illustrating that the sampled states of $IncA_{87-246}$ belong to a single free energy basin. This information, coupled with the high structural similarity of each full-length Markov State and associated relaxation times, draws us to conclude that the structure of $IncA_{87-246}$ represents a thermodynamic minimum (Supplementary Fig. 9c).

Simulation of $IncA_{87-246(G144A)}$ revealed a spike in RMSF at residue 231 (Fig. 3a), a feature which is absent in both the $IncA_{87-246}$ and $IncA_{87-237}$ RMSF profiles. This result agrees with the structural plasticity of $IncA_{87-246(G144A)}$ protomers seen in the triclinic crystal form (Table 1), which deviate significantly at the C-terminus of helix H_D (Fig. 2c). Interestingly, the relative spike in RMSF of the H_{clamp} region is substantially lower in $IncA_{87-246(G144A)}$ than the other two structures. The repositioning of H_D that results from the G144A mutation (Fig. 2c) and its subsequent effect of lowering the flexibility of the H_{clamp} could possibly explain why $IncA_{87-246(G144A)}$ was unable to oligomerize in solution during experiments. Altogether, these results indicate that the structure of $IncA_{87-246}$ is highly stable and is not likely to spontaneously unfold to engage in homotypic fusion. Evidence for the importance of the C-terminal helix H_D in regulating $IncA_{87-246}$ stability is illustrated through the simulation of the shorter $IncA_{87-237}$, which underwent a conformational change, and subsequent comparative analysis of the trajectory against $IncA_{87-246}$ and $IncA_{87-246(G144A)}$. For all three structures, the H_{clamp} is a region of high flexibility, suggesting that it may play a role in the function of IncA. However, when G144 of the hinge is mutated to alanine, the H_D–H_{clamp} distance narrows and results in a significantly less flexible H_{clamp} during simulation compared with $IncA_{87-246}$ and $IncA_{87-237}$, which may explain experimental results regarding the lack of $IncA_{87-246(G144A)}$ oligomerization in solution (Supplementary Fig. 3 and Supplementary Table 1).

Intramolecular contacts are critical for IncA function. To probe the functional importance of IncA oligomerization, we assessed the fusogenic activities of the multiple IncA constructs ($IncA_{87-246}$, $IncA_{87-243}$, $IncA_{87-237}$, $IncA_{87-246(G144A)}$, and $IncA_{87-246(polyA)}$, Supplementary Fig. 1) during *Chlamydia* infection. We complemented an IncA knock-out (KO) *Chlamydia* strain[11] with the different IncA mutants. All IncA mutants expressed the intact N-terminal transmembrane domain (residues 1–86) for efficient secretion to the inclusion membrane, as well as a C-terminal FLAG-tag for rapid identification (see Supplementary Table 3). Expression levels for each $IncA_{mutant}$-FLAG protein were comparable as validated by

western blot analysis (Fig. 4a and Supplementary Fig. 10). HeLa cells were infected with these strains at an MOI of 5, and the homotypic fusion of inclusions was quantified 24 h post-infection (hpi). IncA KO *Chlamydia* complemented with $IncA_{WT}$ undergoes comparable levels of homotypic fusion as wild-type *Chlamydia* (Supplementary Fig. 11) and was used as a positive control, while the IncA KO strain was used as negative control. As shown in Fig. 4b, c, cells infected with *Chlamydia* expressing $IncA_{1-246}$, which lacks 27 C-terminal residues (Supplementary Fig. 1), had 23% fewer cells containing single, fused inclusions compared with $IncA_{WT}$. While this decrease is statistically significant, it demonstrates that the protease-resistant fragment $IncA_{1-246}$ remains largely fusion-competent. When the C-terminus was further shortened, the fusogenicity of IncA continued to drop with the deletion of only three additional residues ($IncA_{1-243}$) resulting in a drastic 43% inhibition of fusion. When six additional residues were deleted ($IncA_{1-237}$), we observed ~82% inhibition of fusion (Fig. 4c), demonstrating a nearly complete loss of fusogenicity. Since $IncA_{87-237}$ is mostly dimeric in solution[8], this finding suggests that the self-assembly of IncA could impact its fusogenic activity. Altogether, these data reveal that progressive deletion of the C-terminal residues of helix H_D in IncA results in an increased loss of fusogenicity, which correlates with a shift towards self-assembly and a loss of monomeric IncA in solution.

To further address the potential impact of IncA oligomerization on homotypic fusion, we complemented the IncA KO *Chlamydia* strain with $IncA_{1-246(polyA)}$, in which the residue-forming contacts on helix H_D have been mutagenized. We observed that $IncA_{1-246(polyA)}$ forms oligomers in solution (Supplementary Fig. 3e, f and Supplementary Table 1). We also observed that $IncA_{1-246(polyA)}$ was unable to drive homotypic fusion efficiently, and only ~ 33% of infected cells contained single inclusions (Fig. 4b, d). This loss of function was specific for these mutations. Cells infected with *Chlamydia* expressing another mutated IncA, $IncA_{1-246(G144A)}$, in which the hinge was genetically eliminated, were fully fusogenic and mostly displayed a single inclusion (Fig. 4b, d). Glycine 144 is not conserved amongst the fusogenic IncA proteins, further supporting our functional data, which indicate this particular amino-acid does not play a critical role in fusion.

Altogether, these data suggest that the intramolecular contacts generated by helix H_D likely maintain IncA in a monomeric fusion-competent state. We observed that the IncA mutants that form oligomers in solution are all non-fusogenic when expressed on the inclusion membrane. Although the oligomerization state of IncA on the membrane is unknown, these results suggest that IncA self-association may lead to its inactivation.

Wild-type IncA rescues fusion of non-fusogenic IncA. To further probe the mechanism controlling IncA function and assess the importance of *trans*-interactions between IncA present on opposing membranes, we conducted a series of co-infection experiments in which cells were infected with both non-fusogenic $IncA_{mutant}$ complemented-IncA KO *C. trachomatis* expressing GFP (MOI 5) and wild-type *C. trachomatis* expressing mCherry ($L2_{mCherry}$, MOI 2) (Fig. 5). Wild-type *C. trachomatis* transformed with mCherry ($L2_{mCherry}$) undergoes the same level of fusion as wild-type *Chlamydia* (Supplementary Fig. 11). We used the IncA KO *Chlamydia* strain complemented with $IncA_{1-246}$, and the IncA KO (non-induced) strain as positive and negative controls, respectively. As expected, $IncA_{1-246}$ was able to promote fusion with $IncA_{WT}$, resulting in a mixed inclusion containing both GFP and mCherry *Chlamydia* (Fig. 5a). Strikingly, the
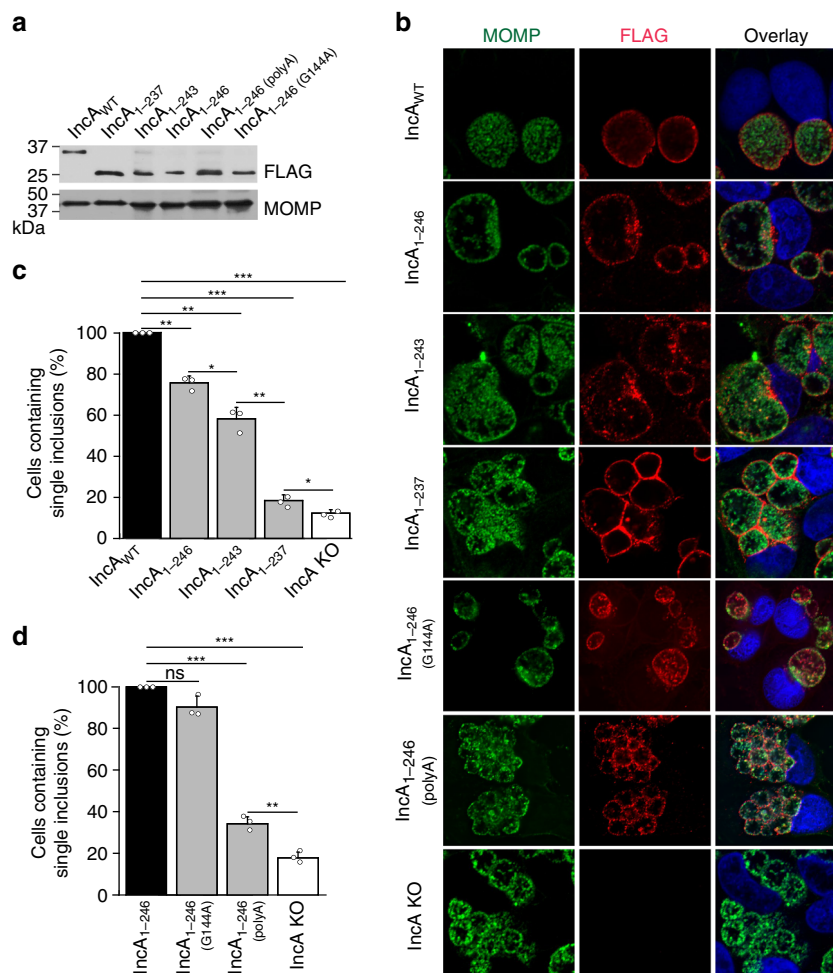
Fig. 4 The intramolecular interactions in IncA$_{87-246}$ are critical for its fusogenic activity. **a** IncA$_{mutant}$-FLAG complemented-IncA KO strains express comparable amounts of IncA$_{mutant}$-FLAG protein (Supplementary Fig. 10). HeLa cells were infected with the indicated strains for 24 h and then lysed in sample buffer. Samples were analyzed by western blot using anti-FLAG and anti-MOMP primary antibodies. *C. trachomatis* MOMP served as a loading control for infection. **b** Immunofluorescence microscopy analysis of HeLa cells infected with the indicated IncA$_{mutant}$-FLAG complemented IncA KO strain at 24 hpi. Bacteria were labeled with anti-MOMP (green) antibody and the expression of IncA$_{mutant}$-FLAG on the inclusion was revealed with anti-FLAG (red) antibody. DNA was stained with Hoechst (blue). The ring-like FLAG staining (red) shows that the IncA$_{mutant}$-FLAG constructs are secreted onto the inclusion surface, indicating that any loss of fusogenic function is not due to their mislocalization. Scale bar = 10 μm. Images are representative of three independent experiments. **c** and **d** Quantification of homotypic fusion with the indicated IncA$_{mutant}$-FLAG complemented IncA KO strain. Data were normalized to IncA$_{WT}$-FLAG (**c**) or IncA$_{1-246}$-FLAG (**d**) *Chlamydia*. IncA KO *Chlamydia* served as a negative control. Graphs display the means of three independent experiments ± the standard deviation. Asterisks indicate statistical significance where, * denotes *p*-values < 0.05, **denotes *p*-values < 0.01, and *** denotes *p*-values < 0.001 (two-tailed student *t*-test). Source data are provided as a Source Data file

mutants for which homotypic fusion was significantly impaired (IncA$_{1-237}$, IncA$_{1-243}$, and IncA$_{1-246(polyA)}$), were all able to promote "heterotypic" fusion when IncA$_{WT}$ was present on the opposing membrane, thus resulting in complete rescue with almost all of the cells (97%, 96%, 94%, respectively) displaying single red wild-type inclusions that fused with green mutant inclusions (Fig. 5), compared with far fewer of the cells (~ 18%, 55%, and 35%, respectively) having homotypically fused inclusions (Fig. 4). These results suggest that IncA$_{WT}$ was able to form *trans*-interactions, likely by disassembling *cis*-oligomers (see Discussion). Inclusions generated by the IncA KO *Chlamydia* strain were unable to undergo fusion even with WT inclusions (Fig. 5a, b, non-induced). This observation confirms that IncA needs to be present on both inclusion membranes to promote fusion. Altogether, these data suggest a model whereby IncA$_{WT}$ is able to disassemble *cis* IncA$_{mutant}$ oligomers to promote IncA$_{WT}$: IncA$_{mutant}$ fusion.

Discussion

Fusogenic viral proteins[21] and eukaryotic SNAREs[22–24] have been extensively studied, and much is known about the key structural determinants and, in some cases, mechanisms of membrane fusion. A common denominator of fusogenic proteins is their intrinsic structural plasticity and ability to undergo dramatic conformational changes upon fusion, typically adopting a meta-stable conformation in the pre-fusogenic state and a thermo-dynamically stable structure post-fusion[25,26]. In contrast, the key players and mechanisms that mediate fusion of chlamydial inclusions remain largely unknown. Chlamydial inclusions are extraordinarily challenging to study due to the fragility of the lipid membrane, the poor conservation of Inc proteins, and the lack of an in vitro fusion assay. In addition, the homotypic fusion of chlamydial inclusions presents a formidable topological challenge because the proteins responsible for this event are identical on both membranes[27]. In this paper, we describe the molecular
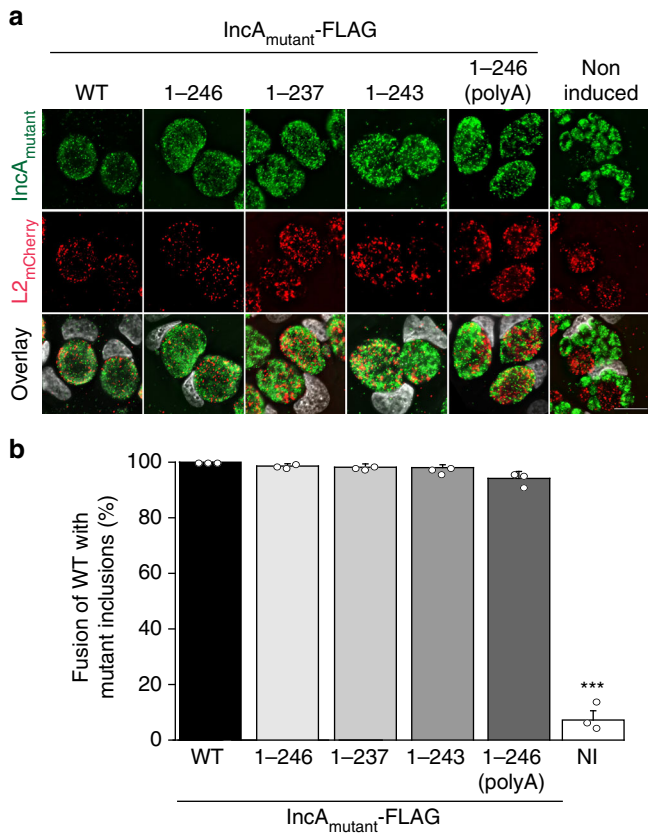
a

b

Fig. 5 Wild-type IncA rescues the activity of non-fusogenic IncA when expressed in *trans*. **a** HeLa cells were co-infected with the indicated IncA$_{mutant}$-FLAG complemented IncA KO strains (IncA mutant, green) at a MOI of 5 and wild-type *C. trachomatis* expressing mCherry (L2$_{mCherry}$, red) at a MOI of 2. IncA$_{mutant}$-FLAG expression was induced with anhydrotetracycline at 5 hpi. Not induced (DMSO) served as a negative IncA KO control. Cells were fixed at 24 hpi and DNA was stained with Hoechst (gray). L2$_{mCherry}$ (red) and IncA$_{mutant}$ (green) inclusions that have undergone fusion contain both red and green *Chlamydia*. Scale bar = 10 μm. **b** Quantification of heterotypic fusion between L2$_{mCherry}$ and IncA$_{mutant}$ inclusions. Data were normalized to IncA$_{WT}$-FLAG. Graphs display the means of three independent experiments ± the standard deviation. Asterisks (***) denote a *p*-value < 0.001 (two-tailed student *t*-test). Source data are provided as a Source Data file

architecture of the prototypical bacterial SNARE-like protein IncA, which mediates the homotypic fusion of chlamydial inclusions[10,11].

Our work sheds light on three aspects that are important for deciphering the mechanisms of chlamydial inclusion membrane fusion. *First*, the cytosolic domain of IncA folds into a non-canonical four-helix bundle characterized by a segmented helix H$_B$, a clamp, and a long C-terminal helix H$_D$. This structure differs profoundly from the four-helix bundle structure of eukaryotic SNAREs previously used to describe IncA topology (Supplementary Fig. 12) but shares similarity with the THATCH domain core of the Huntingtin Interacting Protein 12 that mediates associations between actin and clathrin-coated structures.

Combining biochemical, biophysical, and functional methodologies, we provide evidence that the C-terminal helix H$_D$ of IncA makes key intramolecular contacts with H$_{clamp}$ and helix H$_B$, thus locking the IncA bundle into a stable monomeric pre-fusion conformation. These contacts, in particular, S234, Q240,

I241, and Q245, are conserved in all known fusogenic IncA proteins (i.e., *Ctr*IncA expressed by *C. trachomatis*, *Cmu*IncA expressed by *C. muridarum*, and *Cs*IncA expressed by *C. suis*), supporting the idea that this region is important for IncA function. *Second*, using molecular dynamics simulations, we establish that the crystallographic structure of the IncA cytosolic core is thermodynamically stable and unlikely to undergo major tertiary structural conformational changes spontaneously. Local conformational changes in IncA occur during membrane fusion, but they likely remain confined to regions of the bundle characterized by higher RMSD, such as the H$_{clamp}$ and the C-terminal helix H$_D$. *Third*, by assessing the fusion of the inclusions in HeLa cells infected with genetically manipulated *Chlamydia* strains, we establish a direct correlation between the monomeric state of IncA in solution and its fusogenic activity. Deletion of just nine amino acids of the C-terminal helix H$_D$ results in the formation of stable IncA dimers in vitro and the drastic loss of fusogenic activity during infection. This inhibition of fusion was nonetheless completely rescued when cells infected by *Chlamydia*-expressing IncA$_{mutants}$ were co-infected with *Chlamydia*-expressing IncA$_{WT}$, suggesting that the full-length bundle exerts a dominant function in *trans* to restore the fusogenic activity of IncA$_{mutants}$.

Using these data, we postulate a model in which IncA$_{WT}$ promotes the fusion of inclusion membranes by engaging in homotypic interactions in *trans* (i.e., *trans* homodimer in Fig. 6) mediated by the C-terminal helix H$_D$. Such an interaction is unlikely to be a simple swap of the C-terminal helix because molecular dynamic simulations argue against a global unfolding of the IncA four-helix bundle. Instead, this interaction appears limited to a molecular contact between two IncA bundles. The ability of IncA$_{WT}$ to rescue IncA$_{mutants}$ in *trans* suggests that the association between IncAs is not restricted to the C-terminal helix and may use determinants in the H$_{clamp}$.

The non-canonical helical structure of IncA is key to solving the topological conundrum of homotypic fusion. In this context, the intramolecular contacts within IncA C-terminal helix H$_D$ and H$_{clamp}$ are essential for its fusogenic activity as it allows for homotypic interactions to occur only in *trans*. By removing the intramolecular interactions through point mutations (IncA$_{1-246(polyA)}$) or truncation (IncA$_{1-237}$), we show that IncA$_{mutants}$ self-assemble in solution, likely leading to the formation of *cis* homodimers on the inclusion membrane (Supplementary Fig. 13). Consequently, the *cis* homodimers cannot interact with IncA in *trans* and thus are unable to promote homotypic fusion.

This model is supported by a previous analysis of the oligomeric state of IncA during infection, which found that IncA can naturally switch between monomeric and oligomeric states[5,8,10], but we cannot rule out the existence of an additional host or *Chlamydia* factor that assists in the fusogenic bundling of IncA. One possibility is that a soluble or membrane-bound regulatory protein could catalyze the formation of the IncA *trans*-complex, leading to membrane fusion. It is also possible that the unstructured moiety at the C-terminus of IncA helix H$_D$ or perhaps an intramolecular interaction between the helical bundle and the N-terminal membrane-embedded hairpin stabilizes the auto-inhibitory state that keeps IncA monomeric and competent for fusion.

In summary, the structure of the cytoplasmic domain of IncA described in this study sheds light on how this unique bacterial protein engages in functional complexes to control membrane fusion. Our studies also provide an invaluable template for homology modeling of Inc-orthologs from other fusogenic and non-fusogenic *Chlamydia* strains[10,28], and will help guide us towards a more mechanistic understanding of how mutations and sequence insertions in IncA affect its fusogenic activity.
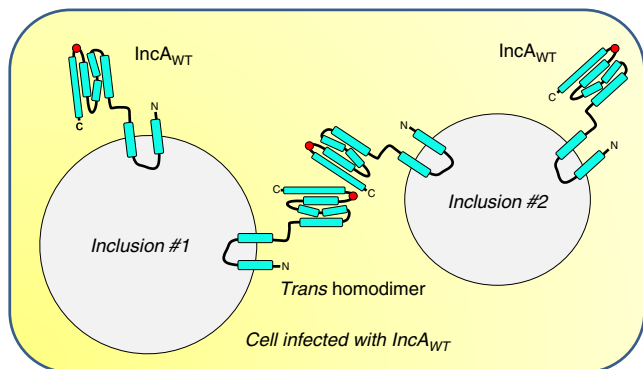
Fig. 6 Model of IncA-mediated homotypic fusion of *C. trachomatis* inclusions. Schematic diagram of a cell infected by *C. trachomatis* that expresses IncA$_{WT}$. Each inclusion displays monomeric IncA$_{WT}$. The fusion of inclusion membranes requires the formation of *trans*-homodimers of IncA$_{WT}$, which interact via the C-terminal helix H$_D$ and the H$_{clamp}$ (highlighted with a red circle)

Methods

Gene cloning and mutagenesis. All primers and plasmids used in this study are detailed in Supplementary Tables 4 and 5, respectively. IncA$_{87-246}$-His$_{6x}$ (FD578) was constructed by PCR amplification of the region Thr87 to Lys246 using primers FO515/FO516 and full-length IncA as a template. The PCR product was digested with NcoI and XhoI and ligated into pET28a. IncA$_{87-246(G144A)}$-His$_{6x}$ (FD915) was made using the QuikChange mutagenesis kit (Agilent) and the primers FO1067/FO1068 following the manufacturer instructions. IncA$_{87-246(polyA)}$-His$_{6x}$ (FD985) was constructed by PCR amplification of the region spanning Thr87 to Lys246 using primers FO1116/FO1185 and a synthetic gene containing the alanine mutations (GenScript) as the template. The PCR product was digested with NcoI and XhoI and ligated into pET28a. IncA$_{1-246}$-FLAG (FD930) was constructed by PCR amplification of the region spanning Met1 to Lys246 using primers FO1023/FO1024 and full-length IncA as the template. The PCR product was digested with NotI and SalI and ligated into pBOMB3-Tet. The pBOMB3-Tet plasmid (FD929) was constructed by removing the tetracycline promoter from pBOMB4-Tet with XhoI and NotI followed by ligation into pBOMB3 (FD923)[29]. IncA$_{1-243}$-FLAG (FD944) was constructed by PCR amplification of the region spanning Met1 to Leu243 using primers FO1023/FO1143 and full-length IncA as template. The PCR product was digested with NotI and SalI and ligated into pBOMB3-Tet. IncA$_{1-246(G144A)}$-FLAG (FD945) was constructed by PCR amplification of the region spanning Met1 to Lys246 using primers FO1023/FO1024, and full-length IncA served as the template. The PCR product was digested with NotI and SalI and ligated into pBOMB3-Tet. Next, the G144A mutation was introduced using the QuikChange mutagenesis kit (Agilent) and the primers FO1067/FO1068 following the manufacturer instructions. IncA$_{1-246(polyA)}$-FLAG (FD948) was constructed by PCR amplification of the region spanning Met1 to Lys246 using primers FO818/FO1184 and a synthetic gene containing the alanine mutations (GenScript) as template. The PCR product was digested with NotI and SalI and ligated into pBOMB3-Tet.

Recombinant protein expression and purification. IncA$_{87-246}$-His$_{6x}$ (FD578), IncA$_{87-243}$-His$_{6x}$ (FD933), IncA$_{87-246(G144A)}$-His$_{6x}$ (FD915), and IncA$_{87-246(polyA)}$-His$_{6x}$ (FD985) were transformed into BL21(DE3) *E. coli* (Invitrogen, cat #44-0049) and grown in Luria-Bertani media to an optical density (at 600 nm) of 0.8. Protein expression was induced with 0.2 mM isopropyl β-D-thiogalactopyranoside at 16 °C for 20 h. After centrifugation, bacteria pellets containing IncA$_{87-246}$-His$_{6x}$ were resuspended in 25 mM Hepes pH 7.4, 75 mM NaCl, 10% (w/v) glycerol, 0.5 mM Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl), 1 mM phenylmethylsulfonul fluoride (PMSF), 10 μM Leupeptin, and 1.5 μM Pepstatin A; IncA$_{87-243}$-His$_{6x}$ was resuspended in 8.45 mM Potassium phosphate mono-basic/40.5 mM Sodium phosphate di-basic pH 7.4, 200 mM NaCl, 10% (w/v) glycerol, 0.5 mM TCEP-HCl, 1 mM PMSF, 10 μM Leupeptin, and 1.5 μM Pepstatin A; IncA$_{87-246(G144A)}$-His$_{6x}$ was resuspended in 25 mM Hepes pH 7.4, 200 mM NaCl, 10% (w/v) glycerol, 0.5 mM TCEP-HCl, 1 mM PMSF, 10 μM Leupeptin, and 1.5 μM Pepstatin A; and IncA$_{87-246(polyA)}$-His$_{6x}$ was resuspended in 8.45 mM Potassium phosphate mono-basic/40.5 mM Sodium phosphate di-basic pH 7.4, 75 mM NaCl, 10% (w/v) glycerol, 0.5 mM TCEP-HCl, 1 mM PMSF, 10 μM Leupeptin, and 1.5 μM Pepstatin A. Resuspended bacteria were lysed by passage through an EmulsiFlex C3 high pressure homogenizer (Avestin). Lysates were then clarified by centrifugation at 35,000 rpm for 30 min at 4 °C, before being incubated with NiNTA beads for 1 h at 4 °C. The NiNTA beads were then loaded on to gravity chromatography columns and washed with resuspension buffer containing increasing amounts of imidazole (50 mM, 75 mM, and 100 mM). Proteins were

eluted with resuspension buffer containing 400 mM imidazole and concentrated in centrifugal concentrators (3000 Dalton cutoff). Purified proteins were resolved on a HiLoad 16/60 Superdex prep-grade 200 column (GE Healthcare) equilibrated with their indicated buffers without glycerol containing 0.5 mM TCEP. The fractions containing protein were collected and concentrated. The purity of each protein was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie Blue staining. Protein concentrations were determined by measuring the absorbance at 280 nm.

Complementation of IncA KO *C. trachomatis* with IncA mutants. IncA KO *C. trachomatis* L2 and the IncA KO complemented with IncA$_{WT}$-FLAG and IncA$_{(1-237)}$-FLAG were previously generated[11]. To complement IncA KO *Chlamydia* with IncA-FLAG mutants, 5 x 10^7 IncA KO elementary bodies and 10 μg of unmethylated DNA were mixed with 100 μl of 2x transformation buffer (20 mM Tris, pH 7.5, 100 mM CaCl$_2$) and brought up to 200 μl with sterile water. Transformations were incubated at room temperature for 35 min and then mixed with 13 ml of cell culture medium (Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% calf serum, 2 mM L-glutamine, non-essential amino acids, 1 mM sodium pyruvate). Two milliliters of transformation was added to each well of a 6-well plate containing confluent Vero cells. Antibiotic selection was initiated at 18 hpi with 100 ng ml^{-1} chloramphenicol. At 46 hpi, the cells were lysed in sterile water and the lysate was used to infect four T$_{175}$ flasks containing confluent Vero cells. The chloramphenicol concentration was increased to 200 ng ml^{-1} and 1 μg ml^{-1} cycloheximide was added. At 46 hpi, the cells were lysed in sterile water and the lysate was used to infect two T$_{175}$ flasks containing confluent Vero cells. The chloramphenicol concentration was increased to 400 ng ml^{-1}, and cycloheximide was maintained at 1 μg ml^{-1}. At 72 hpi, the cell lysate was used to infect one T$_{175}$ flask containing confluent Vero cells. The transformations were subsequently harvested every 46 hpi and the lysate was used to infect one T$_{25}$ flask containing confluent Vero cells until the flask was 30%–60% infected. The selection medium contained 400 ng ml^{-1} chloramphenicol and 1 μg ml^{-1} cycloheximide for all remaining passages. Individual clones were isolated by limited dilution[30] on Vero monolayers for 7–10 days in cell culture medium supplemented with 400 ng ml^{-1} chloramphenicol and 1 μg ml^{-1} cycloheximide. To confirm IncA-FLAG expression, cells were infected with clonal isolates for 24 h at 37 °C, fixed, labeled with anti-FLAG antibody and analyzed by immunofluorescence microscopy. IncA-FLAG expression was induced with 10 ng ml^{-1} anhydrotetracycline at 6 hpi. Selected clones were then purified by density gradient centrifugation[31]. Briefly, infected cells were harvested in K36 buffer (50 mM K$_2$HPO$_4$, 48.9 mM KH$_2$PO$_4$, 100 mM KCl, 14.9 mM NaCl, pH 7.0) and lysed by sonication. Lysates were clarified by centrifugation at 500 x *g* for 15 min at 4 °C and the supernatants were layered on top of a 30% renografin (Mallinckrodt) cushion, before being centrifuged at 40,000 x *g* for 30 min at 4 °C. Each pellet was then resuspended in K36 buffer and layered on top of a discontinuous renografin gradient consisting of 54% (bottom layer), 44% (middle layer), and 40% (top layer) renografin solutions. All renografin solutions were prepared in K36 buffer. The gradients were centrifuged at 40,000 x *g* for 1 h at 4 °C. The band at the 54%/44% interface was harvested, diluted in SPG (17.4 mM Na$_2$HPO$_4$, 2.6 mM NaH$_2$PO$_4$, 4.9 mM L-glutamic acid, pH 7.4) and pelleted by centrifugation at 40,000 x *g* for 30 min at 4 °C. The pellet was then resuspended in SPG and stored at −80 °C. A summary of all the *Chlamydia* strains used in this study is presented in Supplementary Table 3.

Western blot. Infected cells were lysed 24 hpi by scraping into NuPage LDS sample buffer (Invitrogen) containing 2 μM pepstatin A, 10 μM leupeptin, 1 mM PMSF, 10 mM sodium fluoride, and 5.4 mM sodium orthovanadate. After boiling for 5 min at 95 °C, samples were separated by SDS-PAGE and transferred to Polyvinylidene difluoride (PVDF) membrane at 100 V for 1 h. Membranes were incubated in blocking buffer [TBS-T (25 mM Tris, pH 7.5, 150 mM NaCl, 0.1% Tween-20) containing 3% bovine serum albumin] for 1 h at room temperature. Blots were then incubated with mouse anti-FLAG (Sigma, F1804) or goat anti-MOMP (Virostat, 1621) antibodies diluted in blocking buffer overnight at 4 °C. Following several washes with TBS-T, the blots were incubated with anti-mouse and anti-goat HRP-conjugated secondary antibodies (Pierce) diluted in TBS-T containing 0.5% low-fat milk powder for 1 h at room temperature. Blots were then washed several times with TBS-T and revealed using SuperSignal West Dura substrate (Pierce).

Homotypic fusion assay. In all experiments, 6 x 10^4 HeLa cells (ATCC) were seeded on glass coverslips 24 h prior to being infected with wild-type *C. trachomatis* or various *C. trachomatis* mutants at a MOI of 5 in DMEM containing 10% calf serum, 2 mM L-glutamine, 1 mM sodium pyruvate, non-essential amino acids, and 10 μg ml^{-1} gentamicin. The infection was synchronized by centrifugation for 1 h at 1000 x *g* (20 °C). The infected cells were then transitioned to 37 °C and 5% CO$_2$ for the duration of the experiment. IncA-FLAG expression was induced at 6 h post infection (hpi) with 10 ng ml^{-1} (IncA$_{WT}$-FLAG, IncA$_{1-237}$-FLAG, IncA$_{1-243}$-FLAG, IncA$_{1-246}$-FLAG) or 20 ng ml^{-1} anhydrotetracycline (IncA$_{1-246(G144A)}$-FLAG and IncA$_{1-246(polyA)}$-FLAG). At 24 hpi, the cells were fixed with ice-cold 100% methanol for 10 min and blocked in 25 mM Hepes pH 7.4, 150 mM NaCl, 10% horse serum, 0.9 mM calcium chloride, 0.5 mM magnesium chloride, 0.1%

9

Triton X-100. Next, coverslips were stained with mouse anti-FLAG (Sigma, F180) and goat anti-MOMP (ViroStat, 1621) primary antibodies for 1 h at room temperature (RT). After several washes, the coverslips were incubated with secondary antibodies donkey anti-mouse Alexa Fluor 555 (Invitrogen) and donkey anti-goat Alexa Fluor 488 (Invitrogen) and 5 μg ml^{-1} Hoechst (Invitrogen) for 1 h at room temperature (RT). The coverslips were washed and mounted using ProLong Diamond AntiFade Mountant (Invitrogen). Single and multiple inclusions were quantified in a total of 250 infected cells for each *Chlamydia* strain. Images were acquired with a Nikon TiE inverted epi-fluorescence microscope equipped with a 60x oil-immersion objective and Nikon Elements AR software. The percentage of cells with single inclusions was calculated and normalized to either the IncA KO complemented with IncA$_{WT}$-FLAG or IncA$_{1-246}$-FLAG as described in the figure legend. Statistical significance was measured using a two-tailed student *t*-test.

Heterotypic fusion (co-infection) assay. HeLa cells were co-infected with mCherry-expressing *Chlamydia* (L2$_{mCherry}$, MOI 2) and IncA$_{mutant}$-FLAG complemented-IncA KO *Chlamydia* (MOI 5), which constitutively expresses GFP. The infection was synchronized by centrifugation for 1 h at 1000 x g (20 °C). The infected cells were then transitioned to 37 °C and 5% CO$_2$ for the duration of the experiment. IncA$_{mutant}$-FLAG expression was induced at 5 hpi as described above. DMSO (non-induced), which is effectively the IncA KO, served as negative control. The cells were fixed 24 hpi with 4% paraformaldehyde for 20 min. DNA was labeled with 5 μg ml^{-1} Hoechst (Invitrogen) for 15 min at RT. Following several washes, the coverslips were mounted with ProLong Diamond AntiFade Mountant (Invitrogen). The percentage of WT inclusions (red) that fused with IncA$_{mutant}$-FLAG inclusions (green) was enumerated. A total of 150 red WT inclusions for each co-infection condition were counted. Images were acquired with a Nikon TiE inverted epi-fluorescence microscope equipped with a 60x oil-immersion objective and Nikon Elements AR software. Data were normalized to IncA$_{WT}$-FLAG complemented IncA KO *Chlamydia*. Statistical significance was measured using a two-tailed student *t*-test.

Crystallographic methods. The proteolysis-resistant fragment IncA$_{87-246}$ and the mutant protein IncA$_{87-246(G144A)}$ were crystallized using the vapor diffusion hanging drop method by mixing 2 μl of purified protein (typically concentrated at 7.5 mg ml^{-1}) with an equal volume of crystallization solution containing 0.2 M Sodium acetate trihydrate pH 8.0 and 20% (w/v) Polyethylene Glycol 3350. IncA$_{87-246}$ crystals were harvested in nylon cryo-loops, cryo-protected with 27% ethylene glycol and flash-frozen in liquid nitrogen. For single-wavelength anomalous diffraction (SAD) phasing, IncA$_{87-246}$ crystals were soaked for one minute in 5 μl of cryo-protectant solution supplemented with 0.5 M sodium iodide (NaI), and back-washed to remove the excess of NaI prior to plunging in liquid nitrogen. High-resolution diffraction data for native IncA$_{87-246}$ were collected at beamline 9–2, at Stanford Synchrotron Radiation Lightsource (SSRL), on a Dectris Pilatus 6M detector. Crystals of IncA$_{87-246(G144A)}$ and IncA$_{87-246}$ soaked in NaI were diffracted on a Rigaku MicroMax-007 HF diffractometer equipped with a Pilatus3 R 200K direct detector. All steps of data indexing, integration, and reduction were carried out using HKL2000[32], HKL3000[33], and *CCP4* programs[34]. The structure was solved by SAD using an ultra-redundant dataset (i.e., overall redundancy >16) collected in-house. Sixteen iodine sites were located by *phenix.autosol*[35] and used to calculate an initial set of SAD phases with Figure of Merit (FOM) equal to 0.37 between 15–1.8 Å resolution. An initial electron density map calculated with SAD phases was improved by solvent flattening and histogram matching, auto-built alternating cycles of automated model building using *phenix.autobuild*[36] and manual rebuilding using COOT[37]. The completed model was then subjected to positional and anisotropic B-factor refinement to the highest resolution available (~ 1.12 Å) using *phenix.refine*[38] and subjected to final re-refinement using PDB-redo[39], which yielded the best R_{work}/R_{free} and stereochemistry. The triclinic structure of IncA$_{1-246(G144A)}$ was solved by molecular replacement using the wild-type IncA$_{1-246}$ structure as the search model, as implemented in PHASER[40]. The best solution was subjected to positional and B-factor refinement in *phenix.refine*[38] using all reflections between 15–1.95 Å resolution. The final models were refined to a R_{work}/R_{free} of 14.1/16.8% (IncA$_{87-246}$) and 16.9/21.2% (IncA$_{87-246(G144A)}$) using all diffraction data between 15–1.12 Å and 15–1.95 Å, respectively. Final model validation was done using MolProbity[41,42]. Crystallographic data collection and refinement statistics are shown in Table 1. Ribbon diagrams and surface representations were prepared using the program PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.). Intramolecular contacts were measured using the Pisa server[43] and PDBsum[44], and secondary structure superimpositions were carried out in Coot[37].

Analytical ultracentrifugation sedimentation velocity assay. AUC-SV analysis was carried out in a Beckman XL-A Analytical Ultracentrifuge. IncA samples were dissolved at 75 μM and 150 μM (corresponding to 1.5 and 3 mg ml^{-1}) in 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT and were spun at 45,000 rpm at 4 °C. Absorbance values between 285 nm were fit to a continuous sedimentation coefficient (c(s)) distribution model in SEDFIT[45]. Data were visualized and presented using GUSSI (University of Texas Southwestern Medical Center).

Molecular dynamics simulations. Simulations of IncA$_{87-246}$, IncA$_{87-237}$, and IncA$_{87-246(G144A)}$ were performed using NAMD2.12[46] and the CHARMM36m protein forcefield[47]. To prepare these structures into models for equilibrium sampling, they were solvated with the TIP3P water model using the Solvate plugin in VMD[48] and each system was charge-neutralized and ionized with NaCl to a concentration of 150 mM using VMD's AutoIonize plugin. Each model was then subjected to 10,000 steps of gradient minimization. After minimization, harmonic restraints were applied to the protein backbone with a force constant of 5 kcal mol^{-1} during gradual thermalization of the system: the initial temperature was set to 50 K and then incrementally increased by 5 K every 10 ps. Once the systems reached 310 K, the last phase of model preparation was a gradual release of the restrained protein backbone atoms. Ten iterations were employed, removing 0.1 kcal mol^{-1} of the harmonic restraint's force constant after intervals of 50 ps. Equilibrium simulations of the prepared models were run on the NCSA Blue Waters supercomputer and TACC Stampede2. Full electrostatics were calculated at every other timestep, while non-bonded interactions were calculated at every timestep. The Particle Mesh Ewald (PME) method was used to calculate long-range electrostatic interactions. Short-range non-bonded interactions were calculated within a cut-off of 12 Å. Bonds to hydrogen were constrained using the SHAKE and SETTLE algorithms for protein and solvent, respectively. Temperature regulation was accomplished using the Langevin thermostat method, with a damping coefficient of 1.0 ps^{-1}. System pressure was maintained at 1.0 bar using the Nose-Hoover Langevin piston barostat. Each equilibrium simulation was run for 4 μs with a timestep of 2 fs. Trajectory analysis was conducted with VMD. Before processing, the three trajectories were aligned by minimizing the RMSD between backbone atoms of the initial equilibrated structure and all subsequent frames.

Markov state model analysis. Markov State Model (MSM) construction and analysis were conducted using the MSMBuilder[49] application. To prepare the IncA$_{87-246}$ and IncA$_{87-237}$ trajectories, each frame was aligned to the first in its series by RMSD between backbone atom coordinates. Water and ions were removed, and copies were saved every ten frames, with each frame in both of the over 40,000-frame datasets representing a 100 ps interval between samples. Prior to constructing MSMs, the trajectories were clustered by RMSD between atomic coordinates according to the *k*-medoids algorithm; loop regions were ignored. The initial choice of *k*-value was determined by a parameter sweep against the resultant within-cluster sum of squares, or inertia, as a metric of fit quality. The selection criterion for optimal *k* was based on the "Elbow method," where returns on fit quality diminish steeply at *k* + 1. This pointed to an optimal value of *k* = 6. Owing to both the limitations of this approach and the sensitivity of the resulting MSMs to this parameter, the range of tested *k* values was *k* = {4,8}, followed by sparse sampling up to *k* = 4000 for both systems. After, MSMs were built from the clustered trajectories, testing differing model lag times using maximum likelihood estimation and (N_{states}−1) dynamical timescales for different input *k*-clusters. The resulting models of IncA$_{87-246}$ and IncA$_{87-237}$ were scored based on their generalized matrix Rayleigh quotient[20] (GMRQ) training and testing scores. The optimal lag time (τ) for both IncA$_{87-246}$ and IncA$_{87-237}$ models was determined to be τ = 10 ns and was used for all reported results. For the MSMs shown in Supplementary Fig. 9, panels (a) and (b): *k* = 6 clustered input trajectories for IncA$_{87-246}$ and IncA$_{87-246}$; panel c: *k* = 500 clustered input trajectory for IncA$_{87-246}$ only. Free energies were calculated according to $\Delta G = -k_b T \log p$, where k_b is the Boltzmann constant (kcal mol^{-1}) and *p* is the estimated equilibrium population of a given Markov state.

Thermal stability assay. The thermal stability of IncA$_{87-246}$, IncA$_{87-246(G144A)}$, and IncA$_{87-246(polyA)}$ was measured by recording changes in the ellipticity intensity at 220 nm as a function of temperature. Circular dichroism (CD) spectra were recorded using a Jasco J-810 spectropolarimeter equipped with a Neslab RTE7 refrigerated recirculator available at the Sidney Kimmel Cancer Center X-ray Crystallography and Macromolecular Characterization Shared Resource. IncA samples were dissolved in 20 mm sodium phosphate (pH 7.4), 100 mm NaCl and 1 mM DTT at a final concentration of 1.0 μM in a 1 cm rectangular quartz cuvette (Starna Cells, Inc.). CD spectra were recorded between 195 nm and 260 nm. The variations in ellipticity at 220 nm as a function of temperature in 1 °C increments were measured over the range 20−75 °C. Slow cooling to 20 °C followed by a CD scan for secondary structure demonstrated that the unfolding of all IncA samples are largely irreversible. The apparent Melting Temperature (*app*Tm) values for IncA$_{87-246}$, IncA$_{87-246(G144A)}$, and IncA$_{87-246(polyA)}$ are 38, 38.5, and 38.5 °C, respectively.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The atomic coordinates and structure factors for wild-type IncA$_{87-246}$ and IncA$_{87-246(G144A)}$ have been deposited in the Protein Data Bank with accession codes 6E7E and 6E6A, respectively. The authors declare that all other relevant data supporting the findings of this study are included in this published article and its Supplementary Information files, or from the corresponding authors upon request.

References

1. Gerbase, A., Rowley, J. & Mertens, T. Global epidemiology of sexually transmitted diseases. *Lancet* **351**(Suppl 3), 2–4 (1998).
2. Li, Z. et al. Characterization of fifty putative inclusion membrane proteins encoded in the Chlamydia trachomatis genome. *Inf. Imm.* **76**, 2746–2757 (2008).
3. Lutter, E., Martens, C. & Hackstadt, T. Evolution and conservation of predicted inclusion membrane proteins in Chlamydiae. *Comp. Funct. Genom.* **2012**, 362104 (2012).
4. Rockey, D. D., Scidmore, M. A., Bannantine, J. P. & Brown, W. J. Proteins in the chlamydial inclusion membrane. *Microbes Infect.* **4**, 333–340 (2002).
5. Delevoye, C., Nilges, M., Dautry-Varsat, A. & Subtil, A. Conservation of the biochemical properties of IncA from Chlamydia trachomatis and C. caviae: oligomerization of IncA mediates interaction between facing membranes. *J. Biol. Chem.* **279**, 46896–46906 (2004).
6. Paumet, F. et al. Intracellular bacteria encode inhibitory SNARE-like proteins. *PLoS ONE* **4**, e7375 (2009).
7. Ronzone, E. & Paumet, F. Two coiled coil domains of Chlamydia trachomatis IncA affect membrane fusion events during infection. *PLoS ONE* **8**, e69769 (2013).
8. Ronzone, E. et al. An α-helical core encodes the dual functions of the chlamydial protein IncA. *J. Biol. Chem.* **289**, 33469–33480 (2014).
9. Wesolowski, J. & Paumet, F. SNARE motif: A common motif used by pathogens to manipulate membrane fusion. *Virulence* **1**, 319–324 (2010).
10. Hackstadt, T., Scidmore-Carlson, M. A., Shaw, E. I. & Fischer, E. R. The Chlamydia trachomatis IncA protein is required for homotypic vesicle fusion. *Cell. Microbiol.* **1**, 119–130 (1999).
11. Weber, M. et al. A functional core of IncA is required for chlamydia trachomatis inclusion fusion. *J. Bacteriol.* **198**, 1347–1355 (2016).
12. Pannekoek, Y. et al. Interrelationship between polymorphisms of incA, fusogenic properties of Chlamydia trachomatis strains, and clinical manifestations in patients in the Netherlands. *J. Clin. Microbiol.* **43**, 2441–2443 (2005).
13. Suchland, R. J., Rockey, D. D., Bannantine, J. P. & Stamm, W. E. Isolates of Chlamydia trachomatis that occupy nonfusogenic inclusions lack IncA, a protein localized to the inclusion membrane. *Infect. Immun.* **68**, 360–367 (2000).
14. Geisler, W., Suchland, R., Rockey, D. & Stamm, W. Epidemiology and clinical manifestations of unique Chlamydia trachomatis isolates that occupy nonfusogenic inclusions. *J. Inf. Dis.* **184**, 879–884 (2001).
15. Xia, M. et al. Chlamydia trachomatis variant with nonfusing inclusions: growth dynamic and host-cell transcriptional response. *J. Inf. Dis.* **192**, 1229–1236 (2005).
16. Johnson, C. & Fisher, D. Site-specific, insertional inactivation of IncA in Chlamydia trachomatis using a group II intron. *PLoS ONE* **8**, e83989 (2013).
17. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucl. Acids Res.* **38**, W545–W549 (2010).
18. Brett, T., Legendre-Guillemin, V., McPherson, P. & Fremont, D. Structural definition of the F-actin-binding THATCH domain from HIP1R. *Nat. Struct. Mol. Biol.* **13**, 121–130 (2006).
19. Pande, V., Beauchamp, K. & Bowman, G. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **52**, 99–105 (2010).
20. McGibbon, R. & Pande, V. Variational cross-validation of slow dynamical modes in molecular kinetics. *J. Chem. Phys.* **142**, 124105 (2015).
21. Harrison, S. Viral membrane fusion. *Nat. Struct. Mol. Biol.* **15**, 690–698 (2008).
22. Chen, Y. A. & Scheller, R. H. SNARE-mediated membrane fusion. *Nat. Rev. Mol. Cell Biol.* **2**, 98–106 (2001).
23. Hu, C. et al. Fusion of cells by flipped SNAREs. *Science* **300**, 1745–1749 (2003).
24. Weber, T. et al. SNAREpins: Minimal machinery for membrane fusion. *Cell* **92**, 759–772 (1998).
25. Skehel, J. J. & Wiley, D. C. Coiled coils in both intracellular vesicle and viral membrane fusion. *Cell* **95**, 871–874 (1998).
26. Sollner, T. Intracellular and viral membrane fusion: a uniting mechanism. *Curr. Opin. Cell Biol.* **16**, 429–435 (2004).
27. Pérez-Vargas, J. et al. Structural basis of eukaryotic cell-cell fusion. *Cell* **157**, 407–419 (2014).
28. Rockey, D., Fischer, E. & Hackstadt, T. Temporal analysis of the developing Chlamydia psittaci inclusion by use of fluorescence and electron microscopy. *Inf. Imm.* **64**, 4269–4278 (1996).
29. Bauler, L. & Hackstadt, T. Expression and targeting of secreted proteins from Chlamydia trachomatis. *J. Bacteriol.* **196**, 1325–1334 (2014).
30. Mueller, K. & Fields, K. Application of β-lactamase reporter fusions as an indicator of effector protein secretion during infections with the obligate intracellular pathogen Chlamydia trachomatis. *PLoS ONE* **10**, e0135295 (2015).
31. Scidmore, M. Cultivation and laboratory maintenance of *Chlamydia trachomatis*. *Curr. Protoc. Microbiol.* **11**, Unit 11A.a (2005).
32. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzym.* **276**, 307–326 (1997).
33. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 859–866 (2006).
34. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760–763 (1994).
35. Terwilliger, T. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D. Biol. Crystallogr.* **65**, 582–601 (2009).
36. Terwilliger, T. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D. Biol. Crystallogr.* **64**, 61–69 (2008).
37. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
38. Adams, P. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 1948–1954 (2002).
39. Joosten, R., Long, F., Murshudov, G. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220 (2014).
40. McCoy, A. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
41. Chen, V. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
42. Laskowski, R., MacArthur, M., Moss, D. & Thornton, J. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
43. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2018).
44. Laskowski, R. PDBsum new things. *Nucl. Acids Res.* **37**, D355–D359 (2009).
45. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
46. Phillips, J. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
47. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
48. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
49. Beauchamp, K. et al. MSMBuilder2: modeling conformational dynamics at the picosecond to millisecond scale. *J. Chem. Theory Comput.* **7**, 3412–3419 (2011).
50. Tusnády, G. & Simon, I. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.* **283**, 489–506 (1998).

Acknowledgements

Author contributions

Conceived and designed experiments: G.C., J.W., J.R.P., F.P.; Performed the experiments: G.C., M.M., A.L., A.J.B., J.W., D.L.G., R.K.L., E.R.; Analyzed the data: G.C., M.M., A.L., A.J.B., J.W., D.L.G., R.K.L., E.R., J.R.P., F.P.; Contributed reagents/material/analysis tools: G.C., J.R.P., F.P.; Wrote the paper: G.C., A.J.B., J.W., J.R.P., F.P.