Published in final edited form as:

Science. 2015 September 4; 349(6252): 1120–1124. doi:10.1126/science.aaa9046.

Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages

Maria A. Schumacher^{1,*}, Nam K Tonthat¹, Jeehyun Lee¹, Fernando A. Rodriguez-Castañeda², Naga babu Chinnam¹, Anne K. Kalliomaa-Sanford², Irene W. Ng², Madhuri T. Barge², Porsha L.R. Shaw¹, and Daniela Barilla^{2,*}

¹ Department of Biochemistry, Duke University School of Medicine, 243 Nanaline H. Duke, Box 3711, Durham, NC 27710, U.S.A.

² Department of Biology, University of York, York YO10 5DD, United Kingdom

Abstract

Although recent studies have provided a wealth of information about archaeal biology, nothing is known about the molecular basis of DNA segregation in these organisms. Here we unveil the machinery and assembly mechanism of the archaeal *Sulfolobus* pNOB8 partition system. This system employs three proteins; ParA, an atypical ParB adaptor and a centromere-binding component, AspA. AspA utilizes a spreading mechanism to create a DNA superhelix onto which ParB assembles. This supercomplex links to the ParA motor, which contains a bacteria-like Walker motif. The ParB C-domain harbors structural similarity to CenpA, which dictates eukaryotic segregation. Thus, this archaeal system combines bacteria-like and eukarya-like components, suggesting the possible conservation of DNA segregation principles across the three domains of life.

DNA segregation or partition is an essential biological process ensuring faithful genomic transmission. The best understood segregation systems at the molecular level are those employed by bacterial plasmids. These simplified systems consist of a DNA centromere, ParA NTPase and centromere-binding protein (CBP), ParB (1-6). The most common partition apparati use Walker-box NTPases (7). The segregation systems that have been identified on bacterial chromosomes also harbor Walker NTPases, although their mechanisms are less clear (1-6). In contrast to bacteria, eukaryotic partition is highly complex. However, the linchpin in eukaryotic segregation is the histone protein, CenpA, which is deposited in place of histone H3 at DNA centromeres and dictates assembly of the segregation machinery (8-11). While some progress has been made in understanding DNA

*Correspondence to: maria.schumacher@duke.edu and daniela.barilla@york.ac.uk.

ACCESSION NUMBERS. Coordinates and structure factors have been deposited with the Protein Data Bank (4RS7, 4RS8, 4RSB, 4RSF, 4RU7 and 4RU8).

SUPPLEMENTARY MATERIALS

Materials and Methods

Figures S1-S11

Tables S1-S2

segregation in eukarya and bacteria, virtually nothing is known about the molecular process of segregation in archaea, the third domain of life (12-14).

To gain insight into the underpinnings of archaeal segregation we performed a molecular dissection of the proteins encoded on the plasmid pNOB8 partition cassette harbored in *Sulfolobus* NOB8H2 (15). This cassette contains three open reading frames; *orf44*, *orf45* and *orf46*. *orf46* and *orf45* encode 315 and 470 residue proteins, respectively, which show 33-37% and 42-58% sequence similarity to bacterial ParA and ParB proteins. *orf44* generates a 93 residue protein which shows no sequence similarity to any characterized partition protein. The organization of the *orf44-parB-parA* genes (Fig. S1A) is reminiscent of bacterial partition cassettes, however it is tricistronic unlike typical bicistronic bacterial systems (1-4). Hypothesizing, that the pNOB8 centromere may be located either 5' or 3' of the cassette as in bacteria, we analyzed pNOB8 protein binding to these regions. Unlike its bacterial counterparts, pNOB8 ParB only bound DNA nonspecifically (Fig. S1B). Orf44, however, bound the upstream DNA with high affinity ($K_{dapp} \sim 50$ nM) (Fig. 1A). Therefore, we named this CBP, Archaeal Segregation Protein A (AspA). DNase I footprinting showed that AspA interacted with a 23 bp palindrome in the upstream region and increasing AspA concentrations led to spreading around this site (Fig. 1B). Hence, like bacterial ParB proteins, AspA spreads nonspecifically to DNA adjacent to its centromere to form an extended "partition-complex". Although poorly characterized, higher-order partition-complexes are central to segregation as they mediate stabilizing, dynamic interactions with ParA assemblages (1-4). Indeed, the CBP-NTPase interaction is key to the partition process. Biochemical experiments, however, showed that AspA does not bind pNOB8 ParA. Rather, ParB bound both AspA ($K_d = 12$ μ M) and ParA ($K_d = 17$ μ M), indicating that it acts as an adaptor (Fig. S1C). Thus, the pNOB8 partition system is comprised of the CBP AspA, adaptor ParB and ParA NTPase. This finding prompted us to perform BLAST searches for the occurrence of *aspA-parB-parA* cassettes on archaeal genomes. The results revealed that this tripartite cluster of genes is widespread across different crenarchaeal genera and is harbored on both chromosomes and plasmids (Fig. S2).

The AspA structure was obtained to gain insight into its function (Table S1; Fig. S3A). The structure consists of a winged-helix-turn-helix module followed by a C-terminal dimerization helix. Biochemical data supported that AspA is dimeric (Fig. S3B). The AspA fold is distinct from bacterial CBPs and shows structural similarity with the PadR family of transcription regulators, found in bacteria and archaea (16). Three AspA-DNA structures were determined using centromere containing 26mer and 32mer DNA sites (Fig. S3C; Table S1). The structures all contain multiple interacting AspA dimers bound to DNA, that when extended leads to the generation of the same periodically arranged left-handed protein-DNA superhelix (Fig. 1C-E; Fig. 2A). This superhelix bears general resemblance to the left-handed *V. cholera* ParA2-ATP-DNA filament (17). The protein and the DNA cannot be distinguished in the low-resolution ParA2-DNA electron microscopy structure, making a more direct comparison difficult. However, the AspA and ParA2 protein folds and superhelices are distinct; AspA and ParA2 are PadR and Walker-box proteins, respectively and the ParA2-ATP-DNA superhelix has a pitch of 120 Å compared to 70 Å for the AspA-DNA superhelix. The high-resolution of the AspA-DNA structure allows details of binding and superhelix formation to be discerned. Specifically, the insertion of two AspA $\alpha 3$

recognition helices of adjacently bound AspA dimers into each expanded major groove (~13 Å compared to 11 Å for B-DNA) leads to superhelix generation. The small Ala53 side chain appears key in the concurrent occupation of both helices in the same major groove. Leu52-Ala53 and Glu54-Lys59 contacts bolster this interaction (Fig. 1C). To be capable of DNA spreading, a protein must show adaptability in DNA binding. Almost all DNA base and phosphate contacts come from two adaptable AspA elements; the N-terminal arms and glutamines (Fig. S3D) from the recognition helices. The AspA 8-residue N-terminal arms fold only upon DNA binding and adopt a surprising range of conformations, allowing them to insert into either the major or minor grooves where they makes base interactions, phosphate contacts or both (Fig. 1D). Consistent with the superhelical structure, isothermal titration calorimetry (ITC) experiments revealed a stoichiometry of wtAspA dimer to 32mer duplex of 3:1. By contrast, a AspA(A53K) mutation, which replaced the small Ala53 side chain with a bulky residue bound DNA with a stoichiometry of 1:1 (Fig. 2B). Further, DNase I protection revealed that AspA(A53K) was unable to spread on DNA (Fig. 2C). Thus, multiple data support the AspA-DNA superhelix structure.

The pNOB8 ParB adaptor was found to be composed of two domains, an N-domain (ParB-N; residues 1-320) and C-domain (ParB-C; residues 370-470). Full-length ParB bound ParA in the presence or absence of ATP. However, neither ParB-C nor ParB-N bound ParA, indicating that the flexible, ~50-residue interdomain linker mediates ParA binding (Fig. S4A-C). AspA did not bind ParB-C (Fig. S5A) but bound ParB-N and ITC showed the binding stoichiometry was 1 AspA subunit:1 ParB-N molecule (Fig. 3A-B). The nonspecific DNA binding function of ParB was mapped to ParB-C (Fig. S5B). Thus, pNOB8 ParB carries out multiple adaptor functions; binding AspA, ParA and nonspecific DNA (Fig. S4C). We next performed structural studies on ParB-N. As pNOB8 ParB-N crystals diffracted poorly, the *Sulfolobus solfataricus* 98/2 chromosomal ParB (40% identity to pNOB8 ParB) N-domain was used for crystallographic analyses. The ParB-N structure contains two subdomains separated by a central helix (Fig. 3C). DALI searches revealed weak structural similarity between ParB-N and one protein; the N-terminal domain of the bacterial *Thermus thermophilus* ParB homolog, Spo0J (rmsd=3.3 Å) (Fig. 3C, right) (18). However, there are large structural differences in the N-terminal subdomains and central regions of these proteins. The latter region contains a loop in ParB-N and an HTH in Spo0J (Fig. 3C). The lack of an HTH in ParB-N is consistent with its inability to bind DNA and suggests that this domain must have evolved for alternative function(s), e.g. binding to AspA.

Small angle x-ray scattering (SAXS) (19-22) was employed to study the AspA-ParB-N interaction and revealed a symmetric ellipsoid-like structure for the complex that could be fit by placing an AspA dimer in the center flanked by ParB-N molecules (Radius of gyration (R_g) of model=34.2 Å compared to 33.2 ± 0.4 Å from experimental data) (Supplementary Materials and Methods). In this model, which is consistent with the ITC stoichiometry of 1:1, the AspA C-terminal helices are grasped by ParB-N helices $\alpha 11$ - $\alpha 12$ - $\alpha 13$ (Fig. 3D), which dimerize in the model; the same $\alpha 11$ - $\alpha 12$ - $\alpha 13$ dimer was observed in ParB-N structures. Importantly, the AspA-ParB-N model also leaves the N-terminal face of the AspA dimer free to contact DNA, consistent with our data (Fig. 3D). The AspA dimer-ParB-N SAXS model can be docked onto each AspA dimer in the AspA-DNA superhelix leading to

the generation of a multiprotein superstructure, in which ParB-N molecules conform to AspA-DNA superhelical grooves (Fig. 3E). Multiple modeling attempts, without taking into account the SAXS data, revealed that this organization provided not only the best fit to the AspA-DNA superstructure but also the only one without clash. These data suggest that the specific AspA-DNA superhelical structure may function as a template for ParB-N binding (Fig. 3E; left). To test the AspA-DNA-ParB-N model, we generated a stable ParB-N truncation mutant (Fig. S6), lacking helices $\alpha 11$ - $\alpha 12$ - $\alpha 13$, which the model indicates as important in AspA binding. ITC and FP experiments (Fig. 3A) revealed no binding by the mutant.

The AspA-DNA-ParB-N superstructure exposes the ParB C-terminus at the surface, which would project the long linker and attached ParB-C into the solvent, allowing these regions to contact ParA and nonspecific DNA, respectively. The ParB-C structure was solved and consists of five helices. ParB-C forms a tight dimer (burying 2530 Å²), which was supported by biochemical analyses (Fig. 4A; Fig. S7A). Strikingly, ParB-C shows structural similarity to histones, including CenpA (Fig. 4A, Fig. S7B). This is notable because CenpA is the histone variant that marks centromeres in eukaryotes (8). Although the ParB-C dimer is somewhat distinct from histone dimers, it contains an electropositive surface pattern similar to histones and as noted, it binds DNA (Fig. S5B). A ParB-C-DNA model constructed by superimposition with nucleosome structures shows the DNA is proximal to the ParB-C basic patches (Fig. 4B). A conserved feature of CBPs proteins that work in concert with Walker-box NTPases is their ability to both spread onto and bridge DNA (23-24). The presence of many dynamic CBP-ParA attachments could maintain association of CBP-centromere complexes with ParA but simultaneously allow movement. Unlike bacterial CBPs, the pNOB8 CBP, AspA forms a linear partition-complex and shows no bridging capability. The nonspecific DNA binding function of ParB-C could fulfill that role.

Partition mechanisms are dictated by their motor (NTPase) proteins, which bind CBP-centromere partition-complexes. The pNOB8 ParA structure shows strong similarity to bacterial Walker-box ParA proteins (25-27) (Fig. 4C; Fig. S8). Two conformations of pNOB8 ParA were obtained; apo and AMPPNP-bound (Fig. S9A-B). Comparison of the AMPPNP and apo ParA dimers reveals that the so-called signature lysines make the expected intersubunit phosphate contacts in the AMPPNP-bound state, which stabilizes dimerization and enhances ATP hydrolysis. However, these residues are too distant in the apo conformation to make the same interaction (Fig. 4D). Thus, AMPPNP (or ATP) binding to pNOB8 ParA appears to lock in a dimer state optimal for ATPase activity. Both ATP binding and hydrolysis are required for partition by Walker-box ParAs and current data suggests that these proteins move dynamically, en masse on the nucleoid DNA, which serves as a substratum (28-29). Cycles of ParA movement are driven by CBP binding, which stimulates ParA's ATPase activity. This represents a viable segregation mechanism for archaea, which like bacteria, lack a nuclear membrane. Consistent with this, pNOB8 ParA bound DNA nonspecifically in an ATP-dependent manner, similar to bacterial ParAs (Fig. S9C; Fig. S10).

In summary, this work provides molecular insights into DNA segregation in archaea. This includes a delineation of the partition protein structures and higher-order complexes (Fig.

S10). The pNOB8 ParB component was shown to harbor structural similarity to histone/CenpA, which functions as the eukaryotic CBP. On the other hand, pNOB8 ParA harbors a Walker motif similar to bacterial ParAs. Recent data suggest that *Sulfolobus solfataricus* itself employs a chromosomal partition system with a predicted Walker-box NTPase (30). Thus, these findings suggest that Walker-box cassettes may be the most ubiquitous DNA segregation system in nature and also suggests the possible conservation of general segregation principles across the three domains of life.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by NIH grant GM074815 (M.A.S) and BBSRC (BB/F012004/1) and Leverhulme Trust (RPG-245) grants (D.B.). SAXS and crystallographic data were collected at ALS. ALS is a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the DOE, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies program, supported by the DOE Office of Biological and Environmental Research. Additional support comes from MINOS (R01GM105404). We thank Qunxin She for the kind gift of the *Sulfolobus* NOB8H2 strain and Nerea Matamala for initial contributions to the project.

REFERENCES

1. Gordon GS, Wright A. Annu. Rev. Microbiol. 2000; 54:681–708. [PubMed: 11018141]
2. Hayes F, Barillà D. Nat. Rev. Microbiol. 2006; 4:133–143. [PubMed: 16415929]
3. Thanbichler M, Shapiro L. J. Struct. Biol. 2006; 156:292–303. [PubMed: 16860572]
4. Schumacher MA. Curr. Opin. Struct. Biol. 2012; 22:72–79. [PubMed: 22153351]
5. Gerdes K, Møller-Jensen J, Jensen RB. Mol. Microbiol. 2000; 37:455–466. [PubMed: 10931339]
6. Pinto UM, Pappas KM, Winans SC. Nat. Rev. Microbiol. 2012; 10:755–765. [PubMed: 23070556]
7. Koonin EV. J. Mol. Biol. 1993; 229:1165–1174. [PubMed: 8445645]
8. Tachiwana H, et al. Nature. 2011; 476:232–235. [PubMed: 21743476]
9. Valdivia MM, Hamdouch K, Ortiz M, Astola A. Curr. Genomics. 2009; 10:326–335. [PubMed: 20119530]
10. Henikoff S, Furuyama T. Chromosoma. 2012; 121:341–352. [PubMed: 22552438]
11. Bernad R, Sanchez P, Losada A. Exp. Cell Res. 2009; 315:3233–3241. [PubMed: 19660450]
12. Gribaldo S, Brochier-Armanet C. Phil. Trans. R. Soc. 2006; 361:1007–1022.
13. Williams TA, Foster PG, Cox CJ, Embley TM. Nature. 2013; 504:231–236. [PubMed: 24336283]
14. McInerney JO, O-Connell MJ, Pisani D. Nat. Rev. Microbiol. 2014; 12:449–455. [PubMed: 24814066]
15. She Q, et al. Extremophiles. 1998; 2:417–425. [PubMed: 9827331]
16. Fibriansah G, et al. PLOS One. 2012; 7:e48015. [PubMed: 23189126]
17. Hui MP, et al. Proc. Nat. Acad. Sci. U.S.A. 107:4590–4595.
18. Leonard TA, Butler JG, Löwe J. Mol. Microbiol. 2004; 53:419–432. [PubMed: 15228524]
19. Svergun DI. Biophys. J. 1999; 76:2879–2886. [PubMed: 10354416]
20. Svergun DI. J. Appl. Crystallogr. 2000; 33:530–534.
21. Svergun DI, Petoukhov MV, Koch MH. Biophys. J. 2001; 80:2946–2953. [PubMed: 11371467]
22. Petoukhov MV, Svergun DI. Biophys. J. 2005; 89:1237–1250. [PubMed: 15923225]
23. Schumacher MA, Funnell BE. Nature. 2005; 438:516–518. [PubMed: 16306995]
24. Graham TG, et al. Genes & Dev. 2014; 28:1228–1238. [PubMed: 24829297]
25. Dunham TD, Xu W, Funnell BE, Schumacher MA. EMBO J. 2009; 28:1792–1802. [PubMed: 19461582]

26. Leonard TW, Butler PJ, Löwe J. EMBO J. 2005; 24:270–282. [PubMed: 15635448]
27. Schumacher MA, et al. J. Biol. Chem. 2012; 287:26146–26154. [PubMed: 22674577]
28. Ringgaard S, van Zon J, Howard M, Gerdes K. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:19369–19374. [PubMed: 19906997]
29. Vecchiarelli AG, Neuman KC, Mizuuchi K. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:4880–4885. [PubMed: 24567408]
30. Kalliomaa-Sanford AK, et al. Proc. Natl. Acad. Sci. U.S.A. 2012; 109:3754–3759. [PubMed: 22355141]

Fig. 1. Identification of pNOB8 centromere and AspA-centromere structure

(A) AspA (0-1000 nM) was incubated with three DNA fragments upstream of its gene or the pNOB8 *traG* gene and subjected to EMSAs. White triangles denote location of the palindromic centromere. Arrows: unbound DNA; brackets: high MW AspA-DNA complexes. The percentage of bound DNA was plotted against the concentration of AspA. (B) DNase I footprint identifying the DNA site bound by AspA (left, 0-750 nM) and its ability to spread at higher concentrations (right, 0-3000 nM). First lane, AG ladder; last lane, BSA control. The inverted centromere repeat is in green, upstream region in black and *aspA* gene in blue. Below footprints, the sequence shaded in grey indicates the extent of AspA footprint on top and bottom strands at low-medium protein concentrations. (C) AspA-32mer structure. Three interacting AspA dimers coat the DNA, generating a continuous superstructure. Below, close up showing contacts between recognition helices occupying the same major groove. (D) AspA N-terminal arm-DNA contacts. (E) Schematic diagram showing AspA-centromere interactions.

Fig. 2. AspA-DNA superhelix formation

(A) AspA-DNA superstructure. The centrally bound dimer is magenta and dimers spreading in the 5' and 3' directions are green and blue. (B) Top, ITC of wtAspA binding to the 32mer DNA, revealing a binding stoichiometry of 3 AspA dimers to 1 DNA duplex. Bottom, ITC analysis of AspA(A53K)-32mer interaction showing a stoichiometry of 1 AspA dimer to 1 DNA duplex. (C) DNase I footprints performed with wtAspA and AspA(A53K) binding to the pNOB8 *par* upstream region (as in Fig. 1B). Unlike wtAspA, AspA(A53K) does not spread.

Fig. 3. Structure of ParB-N and assembly onto the AspA-DNA superhelix

(A) Increasing concentrations of wtParB-N (blue) or ParB-N $\alpha 11$ - $\alpha 12$ - $\alpha 13$ truncation mutant (red) were added to a preformed AspA-DNA complex to analyze binding. (B) ITC analyses revealing that ParB-N does not bind DNA (left) but binds AspA-DNA (right). (C) Left, pNOB8 ParB-N structure. Right, overlay of ParB-N onto *T. thermophilus* Spo0J. Notable structural differences are denoted by asterisks. (D) AspA-ParB-N model obtained by docking the AspA and ParB-N crystal structures into the SAXS envelope. (E) AspA-DNA-ParB-N model obtained by fitting the ParB-N-AspA model onto the AspA-DNA

superhelix. The AspA-DNA superhelix is blue and ParB-N molecules are shown as coarse cartoon surfaces, with individual molecules colored red and magenta.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4. pNOB8 ParB-C contains a eukaryotic histone/CenpA-like fold and ParA harbors a bacterial Walker-box fold

(A) Superimposition of the ParB-C and *K. lactis* Cse4/CenpA structures (rmsd=2.5 Å for Ca atoms). (B) ParB-C modeled onto DNA showing its electrostatic surface. (C) pNOB8 ParAAMPPNP subunit structure. (D) Superimposition of apo and AMPPNP-bound ParA subunits showing relation of other subunit.