5438–5449 *Nucleic Acids Research, 2016, Vol. 44, No. 11* *Published online 16 April 2016*
doi: 10.1093/nar/gkw248

Molecular insights into DNA binding and anchoring by the *Bacillus subtilis* sporulation kinetochore-like RacA protein

Maria A. Schumacher*, Jeehyun Lee and Wenjie Zeng

Department of Biochemistry, Duke University School of Medicine, 255 Nanaline H. Duke, Durham, NC 27710, USA

Received February 25, 2016; Revised March 21, 2016; Accepted April 2, 2016

ABSTRACT

During *Bacillus subtilis* sporulation, segregating sister chromosomes are anchored to cell poles and the chromosome is remodeled into an elongated structure called the axial filament. Data indicate that a developmentally regulated protein called RacA is involved in these functions. To gain insight into how RacA performs these diverse processes we performed a battery of structural and biochemical analyses. These studies show that RacA contains an N-terminal winged-helix-turn-helix module connected by a disordered region to a predicted coiled-coil domain. Structures capture RacA binding the DNA using distinct protein–protein interfaces and employing adjustable DNA docking modes. This unique DNA binding mechanism indicates how RacA can both specifically recognize its GC-rich centromere and also non-specifically bind the DNA. Adjacent RacA molecules within the protein–DNA structure interact leading to DNA compaction, suggesting a mechanism for axial filament formation. We also show that the RacA C-domain coiled coil directly contacts the coiled coil region of the polar protein DivIVA, which anchors RacA and hence the chromosome to the pole. Thus, our combined data reveal unique DNA binding properties by RacA and provide insight into the DNA remodeling and polar anchorage functions of the protein.

INTRODUCTION

DNA segregation ensures the continuation of progeny from one generation to the next and hence is one of the most fundamental of biological processes. While the mechanisms behind DNA segregation in bacteria are generally poorly understood, sporulating *Bacillus subtilis* have served as a tractable model system to study this process (1–5). During sporulation, *B. subtilis* divides asymmetrically by forming a septum near one pole of the developing cell or sporangium (6–8). The *B. subtilis* polar septum divides the sporangium into the forespore and mother cell. Cellular studies have revealed the main steps of this process, which are characterized by the organization of the chromosome into an elongated DNA mass, which has been called the axial filament, and the anchorage of chromosomes at opposite ends of the sporangium (9–15). While these general steps in *B. subtilis* sporulation have been well described from a cellular level, the molecular details are less understood. Insight into axial filament formation was obtained by the Losick laboratory, which identified a developmentally regulated, kinetochore-like protein that is involved in its formation. This protein was also shown to function in tethering chromosomes to the cell poles and therefore was named RacA for 'remodeling and anchoring of the chromosome' (16,17).

Data revealed that RacA binds specifically to GC-rich centromere sites with the palindromic consensus, TGACGCCGGCGTCA, located near the chromosome origin (17). This DNA site was thus named the RacA binding motif (*ram*). The binding stoichiometry of RacA to its *ram* sites is not known nor is its oligomerization state. Interestingly, however, *ram* sites appear to exist in clusters, which leads to enhanced binding by RacA (17). The interaction between RacA and the centromere-like elements was found to be required to anchor the chromosomes to the poles, presumably through interactions between RacA and proteins localized to the cell poles (16,17). RacA expression is only turned on during sporulation, upon which the intracellular levels of RacA rise to ∼3000 molecules/cell (∼3 μM) (16,17). RacA binds with high specificity to centromere DNA but also displays non-specific DNA binding activity, the latter likely enabled by the high intracellular levels of protein present during sporulation (17). Sporulating cells that harbor a *racA* mutation do not produce an axial filament and fail to trap the DNA in the forespore (16). To gain insight into how a single protein can perform the diverse functions of cell pole attachment, specific centromere recognition, non-specific DNA binding and DNA compaction at the molecular level we carried out structural and biochemical analyses. These combined data re-

*To whom correspondence should be addressed. Tel: +1 919 684 9468; Fax: +1 919 684 8885; Email: maria.schumacher@duke.edu

veal a unique range of molecular functionalities exhibited by RacA that allow it to perform its key roles in chromosome segregation during *B. subtilis* sporulation.

MATERIALS AND METHODS

Expression and purification of *B. subtilis* RacA, RacA domains and DivIVA

An artificial gene encoding *B. subtilis racA*, codon optimized for expression in *Escherichia coli* was purchased from Genscript Corporation, Piscataway, NJ, USA; Web:www.genscript.com. The gene was subcloned into the pET15b vector such that it expressed an N-terminal hexahistidine-tag (his-tag) for purification. The resultant vector was transformed into *E. coli* C41(DE3) cells. For protein expression, the *racA* expressing cells were grown to an OD_{600} of 0.6–0.8 and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37°C. Cells were lysed in buffer A (25 mM Tris pH 7.5, 300 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol) using a microfluidizer and cell debris removed by centrifugation at 17 000 rpm. The lysate was loaded onto a Ni-NTA column and the column washed with increasing concentrations of imidazole in buffer A. The protein was eluted with 100–500 mM imidazole and was >95% pure as assessed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. The purified protein was buffer exchanged into 50 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol and 1 mM DL-dithiothreitol (DTT) for crystallization. Constructs encoding the RacA DNA-binding domain (RacA DBD; residues 1–70) and the RacA C-terminal region (residues 71–184) were polymerase chain reaction amplified from the *racA* gene and subcloned into pET15b. An artificial gene encoding the DivIVA C-terminal domain (residues 66–164), codon optimized for *E. coli* expression was also obtained from Genscript and subcloned into pET15b. The resultant expressed protein domains were all 95–98% pure after Ni-NTA chromatography and used in biochemical and structural studies after buffer exchange into 50 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol and 1 DTT.

Crystallization and structure determination of RacA–DNA complexes

The his-tag was removed from the full length (FL) RacA protein via thrombin cleavage. Well diffracting crystals were obtained of RacA in complex with the 14mer palindrome, 5′-TGACGCCGGCGTCA-3′. For crystallization, RacA (at 10 mg/ml) was mixed with the DNA at a ratio of 2:1 ratio of RacA to DNA duplex and this solution was mixed 1:1 with a crystallization solution consisting of 10% PEG 3000, 0.1 M Tris HCl pH 8.0, 0.2 M lithium sulphate. Crystals took 4–10 months to grow and take the space group $P2_1$ with $a = 56.6$ Å, $b = 68.5$ Å, $c = 117.4$ Å and $\beta = 97.5°$. The structure was solved by multiple wavelength anomalous diffraction (MAD) using a DNA duplex in which the thymines were replaced with 5-bromouracil. RacA–DNA complexes with the brominated DNA produced crystals isomorphous with those obtained with the wild type (WT) DNA. Both crystals were cryo-preserved by dipping them in a solution supplemented with 20% glycerol for 4 s prior to placement in the cryo-stream.

MAD RacA–DNA data were collected at Advanced Light Source (ALS) beamline 8.3.1–2.7 Å at three wavelengths that corresponded to the bromine peak and inflection and a remote wavelength. The data were processed using MOSFLM and the heavy atom substructure was determined via SOLVE. The resulting figure of merit was 0.48 prior to density modification. Density modification (RESOLVE) generated a readily traceable electron density map. Although the FL protein was used to grow the crystals, the map revealed that the asymmetric unit (ASU) was comprised of three DNA duplexes and 10 RacA subunits. Only residues RacA 1–64 were visible. This was a result of proteolysis as when the crystals were run on a gel, only fragments and not the FL protein were observed. Refinement of the structure was performed using Crystallography & NMR System (CNS) (18). The final model includes 10 RacA molecules as well as all nucleotides of the three DNA duplexes and 168 water molecules. Final model refinement statistics are presented in Table 1.

A second RacA-14mer crystal was grown by mixing RacA(1–70) at a ratio of 2:1 with the 14mer DNA duplex. Crystals were grown using 30% PEG 4000, 0.1 M Hepes pH 7.5, 0.2 M NaCl as a crystallization reagent. The crystals were cryo-preserved using the crystallization reagent supplemented with 20% glycerol and data were collected at ALS beamline 8.3.1. The crystals displayed high mosaicity, which limited the usable data to 3.0 Å. A successful molecular replacement solution was obtained using a RacA wing dimer as a search model and clear density was revealed for the DNA in the F_o–F_c map following refinement of this starting model. The structure was refined to convergence using CNS and Phenix (18,19).

Crystallization and structure determination of apo RacA DBD

The RacA DBD, RacA(1–70), was used for crystallization after the his-tag was removed by thrombin cleavage. The protein was buffer exchanged into 20 mM Tris pH 7.5, 100 mM NaCl and concentrated to 15 mg/ml for crystallization. Crystals were obtained overnight by mixing the protein solution 1:1 with 20% PEG monomethylether 5000, 0.1 M Bis Tris pH 6.5. The crystals take the orthorhombic space group $P2_12_12_1$. Data were collected at ALS beamline 8.3.1–1.80 Å and processed with MOSFLM. The structure was solved by molecular replacement using the DBD from the RacA–DNA structure as a search model. The structure was refined using Phenix (Table 1).

Size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) was used to probe the molecular weights of RacA and the RacA DBD. All SEC experiments were performed using a HiLoad 26/600 Superdex 75 prep grade column. Experiments were performed in a buffer containing 300 mM NaCl, 5% glycerol, and 20 mM Tris HCl pH 7.5, 1 mM beta-mercaptoethanol (BME).

Table 1. Data collection and refinement statistics for RacA–DNA and apo RacA

	Apo RacA	RacA-14mer	RacADBD-14mer
Space group	$P2_12_12_1$	$P2_1$	$P3_1$
Cell dimensions			
a, *b*, *c* (Å)	33.8, 34.5, 42.7	56.6, 68.5,117.4	40.1, 40.1,136.5
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 97.5, 90.0	90.0, 90.0,120.0
Resolution (Å)	24.15–1.80	68.5-2.62	33.6-3.00
R_{sym} or R_{merge}	0.081 (0.562)[a]	0.057 (0.523)[a]	0.055 (0.153)[a]
I / σI	12.5 (2.6)	15.5 (2.7)	6.5 (2.0)
Completeness (%)	95.4 (77.7)	98.6 (95.4)	97.9 (97.6)
Redundancy	6.1 (4.4)	5.0 (4.4)	2.5 (2.4)
Refinement			
Resolution (Å)	24.15–1.80	68.5–2.65	33.6-3.00
R_{work}/R_{free} (%)	17.9/21.3	23.9/26.5	19.2/21.9
R.M.S. deviations			
Bond lengths (Å)	0.003	0.009	0.010
Bond angles (°)	0.550	1.23	1.53

[a]Values in parentheses are for highest-resolution shell.

Circular dichroism

Far UV circular dichroism (CD) spectra were collected on an AVIV 202 CD spectrophotometer. Before CD data collection, RacA was exchanged into a buffer consisting of 20 mM phosphate buffer and 150 mM NaF. Before analyses, the protein was diluted to ∼0.2 mg/mLl and the spectrum collected from 190–300 nm.

Fluorescence polarization (FP) binding assays

Fluorescence polarization (FP) assays were carried out to examine DNA binding by FL RacA, the RacA DBD, the RacA C-domain and RacA DBD mutants (RacA DBD(K15A), RacA DBD(R19A), RacA DBD(R30A) and RacA DBD(G34E)). The reactions were performed at 25°C in a buffer consisting of 150 mM NaCl and 25 mM Tris HCl pH 7.5. For these experiments, proteins were titrated into a reaction tube with the buffer containing 1 nM fluoresceinated DNA until binding saturation was reached. All FP data were plotted and fit using KaleidaGraph.

Isothermal titration calorimetry (ITC)

For isothermal titration calorimetry (ITC), protein samples were dialyzed into binding buffer consisting of 25 mM Tris pH 7.5 and 150 mM NaCl and thoroughly degassed. The protein concentrations were determined by Bradford assays. All ITC measurements were performed at 25°C using a VP-ITC calorimeter (Microcal). A total of 5-μl injections (28 total) of the 450 μM DivIVA(66–164) solution were injected into 470 μl of 10 μM RacA (either RacA(1–70) or RacA(71–184)) with an interval of 250 s at a constant stirring rate of 250 rpm. Similarly, 28 total injections of 5 μl of 680 μM RacA were injected into 470 μl of DivIVA(66–164). All data were collected and analyzed using the manufacturer-supplied software package, Origin 7.0. The resulting isotherms were fit to one-site, independent binding models.

Glutaraldehyde crosslinking and sedimentation assays

To assess the molecular weight of RacA we performed glutaraldehyde crosslinking. Purified RacA at 0.5 mg/ml in 50 mM Hepes pH 7.0, 50 mM NaCl was crosslinked with 0.2% glutaraldehyde (final concentration) and time points collected. The reaction was quenched by addition of SDS running buffer. For sedimentation assays, 10 μm of purified RacA(1–70) was combined with 54-mer DNA in 0.1 M Tris pH 7.5, 100 mM NaCl. Reaction mixtures (100 μl) were centrifuged using a TLA100.3 rotor at 40 000 rpm for 15 min. Equivalent amounts of supernatant and pellet fractions were separated on 15% SDS-PAGE gels and visualized by staining with Simply Blue Safe Stain.

RESULTS

RacA structures reveal multiple modes of binding leading to DNA coating

To deduce the molecular mechanism behind the unusual DNA binding properties of RacA, we crystallized the *B. subtilis* RacA protein in the presence of a 14mer palindromic centromere DNA site, 5′-TGACGCCGGCGTCA-3′ and solved the structure by MAD ('Materials and Methods' section). The structure, obtained to 2.62 Å, contains multiple RacA molecules bound to three DNA duplexes in the crystallographic ASU (Figure 1A and B; Table 1). The DNA stacks end to end to create a pseudocontinuous helix with the protein wrapping the DNA about its electropositive surface (Figure 1B). Although the FL 184 residue RacA protein was utilized in crystallization trials in a 2:1 (RacA subunit: DNA duplex) ratio, only residues 1–64 are visible for each RacA subunit. Subsequent analyses showed that over the extended time required for crystal growth, RacA had broken down into N- and C-terminal regions. The N-terminal fragment encompasses the DBD. This finding was consistent with sequence alignments of RacA homologs, which showed that RacA proteins contain a highly conserved N-terminal domain (N-domain) connected to a C-terminal region (C-domain) predicted to contain a coiled coil by a linker of ∼40–60 residues that is divergent not only

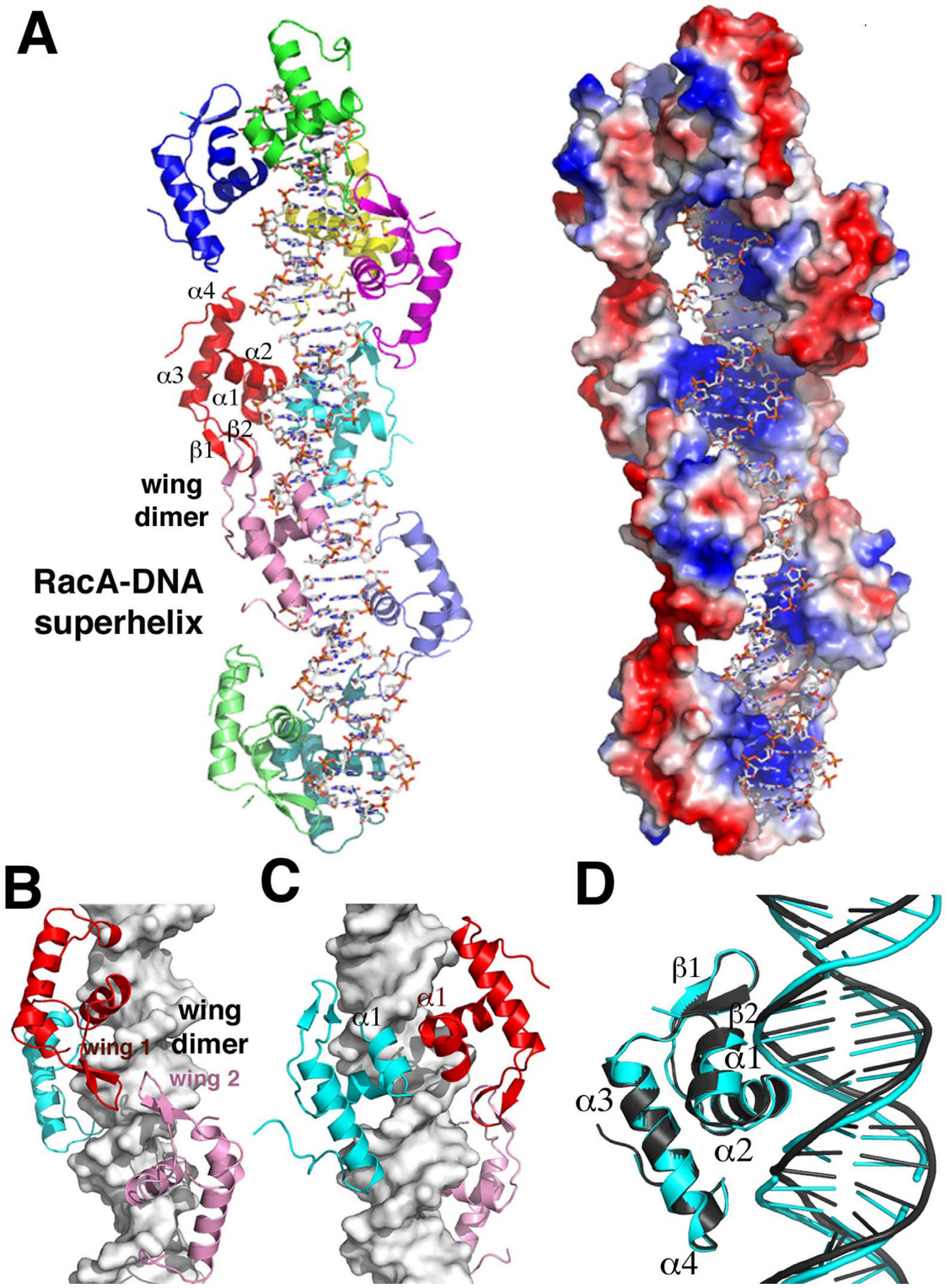

Figure 1. Structure of *Bacillus subtilis* RacA–DNA complex. (**A**) Overall structure of the crystallographic asymmetric unit of the RacA–DNA complex. Each RacA subunit is shown as a ribbon and colored differently. The DNA stacks pseudocontinuously. Labeled is one 'wing dimer' in which RacA subunits dimerize on the DNA through contacts between their two fold related wings. (**B**) Close up of the red-pink wing dimer from (A) with the wings labeled. (**C**) Close up of the α1–α1 interaction that assembles additional RacA molecules on the DNA. (**D**) Superimposition of one subunit from a wing dimer (dark gray) and a subunit forming α1–α1 contacts (cyan). Although these molecules make different RacA–RacA and DNA interactions, they dock on the DNA in a similar overall manner with their recognition helices inserted in the major groove and their wings in the minor grooves.

in sequence but also length and is predicted to be disordered (Supplementary Figure S1).

The RacA molecules in the structure harbor the topology: α1 (residues 2–10), α2 (14–23), β1 (28–32), β2 (34–38), α3 (40–54), α4 (58–61). Helices 1–2 form a helix-turn-helix (HTH) motif and β-strands 1–2 and the short turn between them comprise a DNA binding 'wing' (Figure 1A–C). In the structure, 10 RacA subunits bind the three 14mer DNA sites arranged pseudocontinuously in the crystallographic ASU. Each 14mer palindrome site is contacted by two RacA molecules that form a weak dimer via contacts between their wings (herein termed 'wing dimers') (Figure 1A and B). Unexpectedly, however, four additional RacA subunits dock onto the DNA and insert between the wing dimer molecules. These additional RacAs are anchored onto the DNA by α1–α1 interactions with the DNA bound RacA wing dimer subunits (Figure 1A and C). As a result each 14mer site is bound by up to 4 RacA molecules. Little surface area is buried in the wing–wing or α1–α1 interfaces (∼200–400 Å) suggesting that they form only upon DNA binding. This was supported by SEC analyses, which showed that the RacA DBD, RacA(1–70), is monomeric (Supplementary Figure S2). Despite the different juxtapositions of the RacA molecules, superimpositions show they dock onto the DNA similarly, whereby helix 2 of each subunit inserts into the major groove and the wings bind near the minor groove and contact the DNA phosphate backbone (Figure 1D).

The DNA docking mechanism employed by RacA appears unique. However, DALI searches show that the RacA DBD harbors significant structural homology to the *B. subtilis* transcription regulators GlnR and TnrA (r.m.s. deviation (rmsd) = 1.2 and 1.4 Å, respectively), which are a subfamily of the MerR superfamily of DNA binding proteins (Figure 2A) (20). MerR and TnrA/GlnR proteins have similar DBDs but differ in their overall DNA binding and oligomerization modes; MerR proteins dimerize by a coiled coil directly connected to their DBDs, while the TnrA/GlnR proteins only dimerize upon DNA binding via interactions between N-terminal regions that are not present in MerR proteins or RacA (Figure 2B) (21,22). Instead of a coiled coil directly attached to their DBDs as in MerR proteins, TnrA/GlnR members contain a long disordered region that is attached to a short C-terminal tail, which folds into a helix upon binding to glutamine synthetase (20). RacA is similar to TnrA/GlnR proteins in containing a long disordered region C-terminal to its DBD. However, RacA differs from TnrA/GlnR proteins in its oligomeric interactions, which are formed between its wings and α1 helices present in its DBD (Figure 2B). These interactions also show no similarity to any oligomer in the protein database.

Mechanism for specific and non-specific DNA binding by RacA

An important yet enigmatic aspect of RacA's function is its ability to bind DNA both specifically and non-specifically. Sequence specific binding by RacA was revealed by ChIP-on-chip analyses, which showed that it interacts with 25 distinct regions of the *B. subtilis* chromosome that extend from −412 to + 200 kb (17). These *ram* sites contain 5′ and 3′ TG/CA steps, which were shown to be important for specific binding (Figure 3A) (17). While the two modes of DNA docking by RacA subunits revealed in the structure both involve insertion of the HTH into the major groove and the wing into the minor groove, these subunits make different sets of DNA contacts. Strikingly, however, the same residues are employed in making the different sets of contacts. The RacA subunits that participate in forming wing dimers make seven phosphate contacts and 3–4 base interactions while the other RacA molecules each contribute nine phosphate and two base interactions (Figure 3A). Thus, both docked subunits make an unusually large number of phosphate interactions for such a small DNA-binding module. This DNA binding feature likely permits promiscuity in DNA binding. In both cases, these phosphate contacts are provided by residues located in the N-terminus of α1, α2, the wing and the N-terminus of the short C-terminal helix, α4. The RacA molecules that are anchored onto the DNA by wing–wing interactions make specific contacts with the centromere DNA (Figure 3A) and hence appear to mediate centromere binding specificity. The Arg19 side chains in these RacA subunits read the TG/CA steps via bidentate hydrogen bonds to guanine2 and stacking contacts with the 5′ thymine (thymine1) in the DNA major groove (Figure 3B). These dual interactions perfectly fit the criteria for a specific protein–DNA contact called a 5′-pyrimidine-guanine-3′ (5′-YpG-3′) interaction (23). This contact mode is thought to arise because of the inherent flexibility of pyridimine-guanine steps, which allows them to become unstacked more readily than other bps (23,24).

In a 5′-YpG-3′ interaction, an arginine side chain makes specific hydrogen bonds to the major groove face of the 5′-guanine nucleobase, while simultaneously contacting the unstacked preceding pyrimidine (23). This recognition mode was initially discovered in the complex between the *Saccharomyces cerevisiae* sporulation specific transcription factor, Ndt80 and the middle sporulation element (MSE) DNA site (24). A subsequent survey by the Glover lab showed that several different classes of DNA-binding motifs used this form of DNA recognition (23). The thymine1 nucleobases in RacA–DNA structures are all extensively displaced from stacking with guanine2 to a point where they stack nearly optimally (∼3.5 Å) with the Arg19 side chain (Figure 3B). The Arg19Cζ atom also makes van der Waals contacts with the thymine methyl group (3.5 Å), thus providing specificity for a TpG bp over a CpG bp step. The side chains of Lys15 from the wing dimers also contacts guanine bases in the major groove. However, the density for these residues is somewhat spurious indicating that these may not be tight interactions. The Lys15 residues from the RacA α1–α1 interacting subunits make base interactions to two guanines, guanine5 and guanine9', which might impart limited specificity in binding (Figure 3C). Interestingly, the two residues in RacA that contact bases, Lys15 and Arg19, are modular in their approach to DNA binding as the Lys15 side chain can contact different bases depending on the DNA sequence and Arg19 can make bipartite contacts that recognize the *ram* site while also, through side chain rotations, interact with phosphates in non-*ram* DNA.

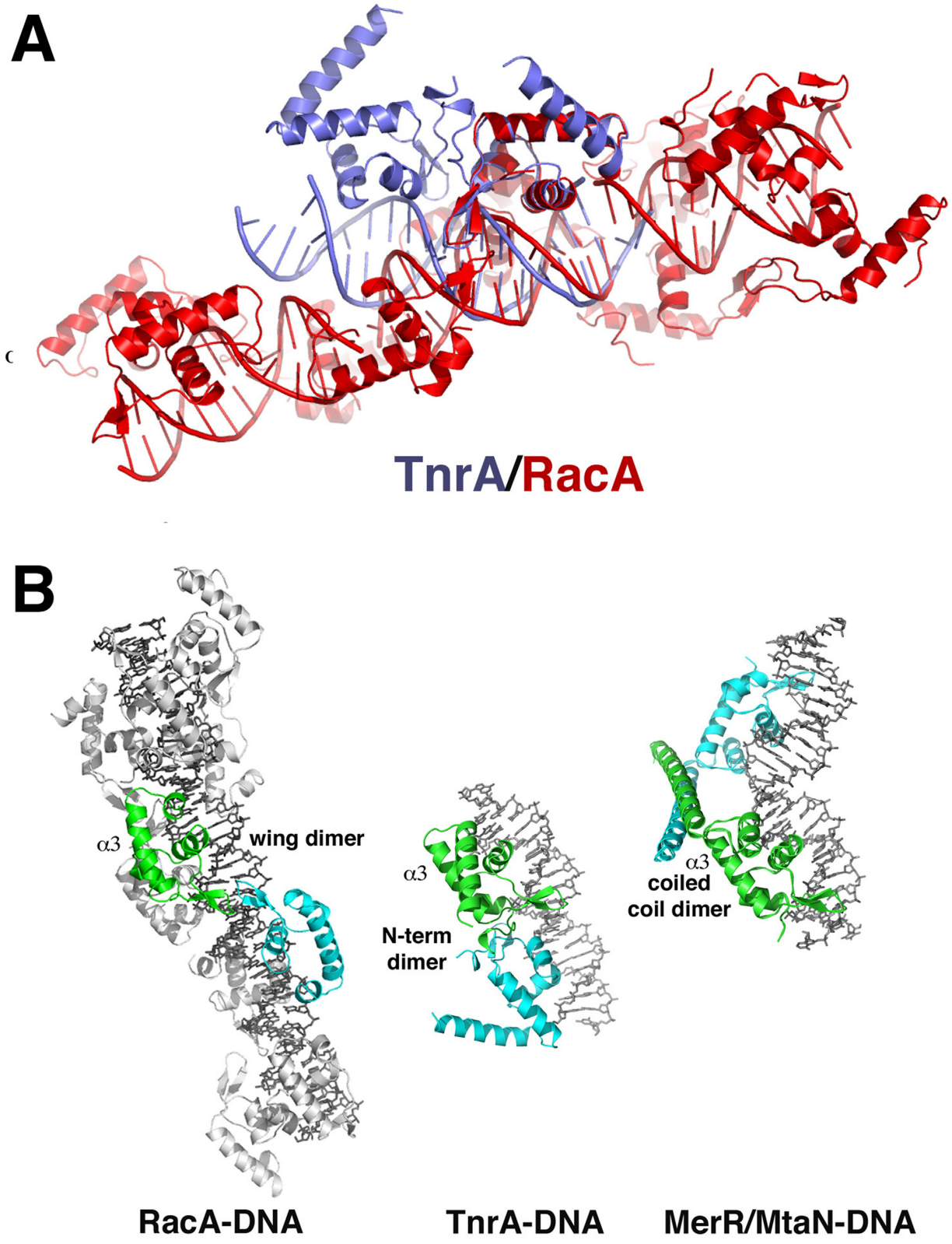

Figure 2. Comparison of RacA with TnrA/GlnR and canonical MerR proteins. (**A**) Overlay of one DNA-binding domain of RacA and TnrA underscoring that although the individual subunits superimpose well, the oligomerization and overall DNA binding modes are distinct. (**B**) Side by side comparison (using the green subunits as a reference) of RacA–DNA, TnrA–DNA and the canonical MerR protein MtaN–DNA complexes. The cyan subunit corresponds to the dimer partner of each green subunit. The figure highlights, again, the significant differences in how these proteins bind DNA and oligomerize.

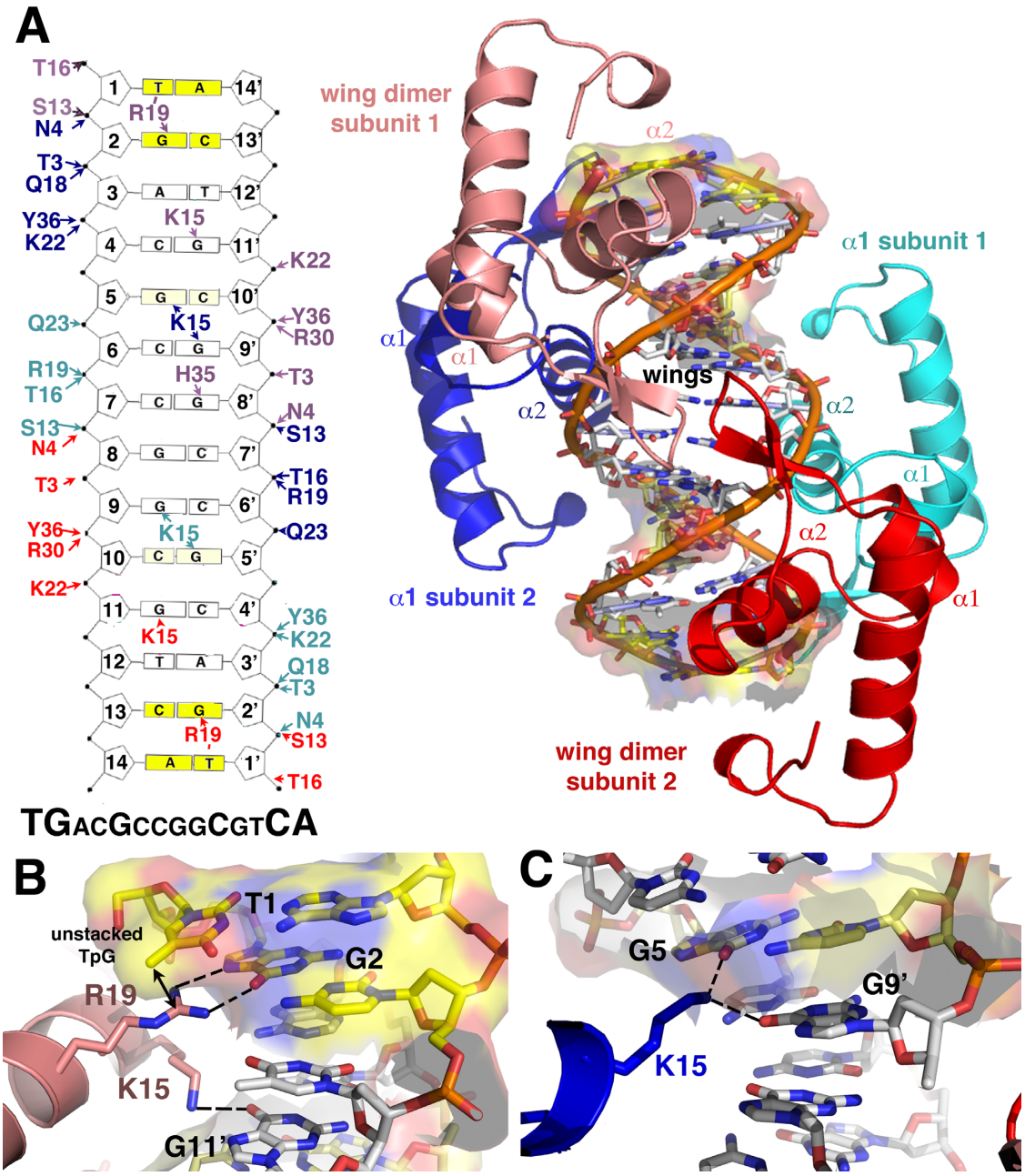

Figure 3. RacA DNA contacts. (**A**) Left, schematic showing the contacts formed by one RacA wing dimer (subunits colored red and pink) and two RacA subunits docked via α1–α1 contacts with the wing dimers. The contacts from the latter molecules are colored cyan and blue. The conserved nucleotides in the centromere that are essential for base specific binding are colored yellow and the G-C bps that provide some level of specificity are colored light yellow. This conservation is also indicated below by a MDscan logo (17), where the key bps for binding are shown in larger letters than less important/conserved bps. To the right is a ribbon diagram showing the organization of the wing dimers (red, pink) and α1–α1 dimers on the DNA. (**B**) Close up view of the base specific interactions that specify the TpG steps at each end of the centromere consensus and is mediated by the Arg19 side chains. (**C**) Close up of the sole base contacts provided by Lys15 from the α1–α1 interacting subunits.

Thus, the combination of adaptable phosphate and base interacting residues provides promiscuity in DNA binding by RacA. Indeed, use of one residue, Arg19, for generating the full complement of key base specific interactions builds in a readily adjustable transition from specific to non-specific contacts.

Sequence alignments of RacA homologs show that the DNA contacting residues are highly conserved and that Lys15 and Arg19 are completely conserved (Supplementary Figure S1). The presence of the specific 5′-YpG-3′ contacts from the Arg19 side chains of the RacA subunits in the wing dimer suggested that this represents the specific DNA binding mode of RacA. Support for this supposition was obtained from a second RacA–DNA structure that was solved from crystals grown using a 2:1 RacA(1–70):14mer DNA mixture (Table 1). The structure revealed the same

binding mode as the other RacA–DNA crystal form. However, in this structure only the wing dimer is bound to each 14mer (Figure 4A). This finding suggests that the first RacA molecules bind to the specific centromere to form the wing dimer and additional RacA molecules bind using α1–α1 contacts with the already bound RacA subunits. The DNA substrates bound by RacA show no significant global bend, however the average minor groove widths of the DNA are 8.1 Å and 7.2 Å (25) for the multiply bound and wing dimer only bound structures respectively, compared to 5.7 Å for B-DNA, indicating that DNA distortion is involved in RacA binding.

Structure of apo RacA DNA-binding domain

To deduce if any conformational changes occur in the RacA N-terminal domain upon DNA binding we determined the RacA DBD apo structure to 1.80 Å resolution (Table 1). No oligomer was revealed by crystal packing or by PISA analysis, consistent with our SEC analyses. Superimpositions of the apo and DNA bound form of the RacA DBD revealed that the structures are essentially identical with the exception of wing residues 29–35 (Figure 4B). In the apo conformation, the side chain of Asn31 rotates inward and makes a cluster of hydrogen bonds to the backbone atoms of residues 32–36 (Supplementary Figure S3). This coordination stabilizes the wing in a compact form and allows Arg30 and Tyr36 to stack together. Upon DNA binding, Arg30 and Tyr36 move to interact with the DNA phosphate backbone. This shift would facilitate formation of the wing dimer as it positions Arg30 to hydrogen bond with the backbone carbonyl oxygen of Glu32 from another wing. The 2-fold related arginine-carbonyl oxygen interaction glues the two ends of the wings together. The small Gly34 residue is critical for wing dimer formation as the two fold related Gly34 residues pack together in the wing dimer interface (Figure 4B).

Biochemical characterization of RacA DNA binding and C-domain structure

The structural data indicate that the RacA N-terminal domain is responsible for DNA binding. However, to directly test this hypothesis and to further probe the RacA DNA binding mechanism, we carried out FP experiments ('Materials and Methods' section). Binding experiments were carried out with FL RacA, RacA(1–70) and RacA(71–184) and revealed that FL RacA bound the 14mer centromere site with a K_d of 58.4 +/− 0.3 nM, the N-domain, RacA(1-70), bound the DNA with a K_d of 9.2 +/− 4 μM and the RacA C-domain showed no binding, even at high protein concentrations (Figure 4C). Consistent with previous studies, FP showed that RacA binds DNA nonspecifically; FL RacA bound a non-*ram* site (K_d = 0.55 +/− 0.03 μM) (Supplementary Figure S4). The RacA C-domain showed no binding to this site (Supplementary Figure S4).

The combined data indicate that the RacA N-domain mediates DNA binding but that the C-domain enhances this binding in an indirect manner. Oligomerization by a distant, physically separated domain has been shown to be one way to increase DNA binding affinity through mass action (24–26). For example, DNA binding by bacterial segregation ParB proteins containing HTH DNA-binding motifs are significantly enhanced by the oligomerization of domains connected to the HTH domains by flexible linkers (26–28). Consistent with this idea, residues ∼100-160 in the RacA C-domain are predicted to form a coiled coil, which is a common oligomerization module. CD can be used to detect the presence of coiled coils as such structures give a characteristic spectra with a 220 nm/209 nm of 1 or >1 compared to non-coiled-coil helical proteins, which produce spectra with ratios of 0.6–0.8 (29,30). CD studies on the RacA C-domain resulted in a CD spectrum with a 220/209 nm ratio of 1.02, supporting that it harbors a coiled coil (Supplementary Figure S5). To assess whether the RacA C-domain oligomerizes we employed glutaraldehyde crosslinking and SEC. Crosslinking revealed the presence of dimers and some tetramers in both the FL protein and the C-domain samples (Supplementary Figure S6A and B) and SEC analysis was consistent with a tetramer for the FL protein (Supplementary Figure S6C). Hence, the combined studies indicate that the RacA C-domain, which is attached by a flexible linker to the monomeric DBD, contains a coiled coil and forms dimers and possibly tetramers at increased protein concentrations.

The RacA–DNA structures reveal residues important in base specific contacts to the centromere as well as non-specific interactions. The structure also predicts that although the Gly34 side chain does not make direct contacts to the DNA, its small size would be important for DNA binding as any other residue would disrupt formation of the wing dimer and hence prevent RacA proper docking on the DNA. To test these structural hypotheses, we generated four mutations in the RacA DBD, K15A, R19A, R30A and G34E, and measured their DNA binding by FP. RacA(K15A) displayed weak DNA binding, with an estimated K_d of 50 μM. The R30A mutant displayed a K_d similar to the K15A while K_ds could not be accurately determined for the R19A and G34E mutants (Figure 4D). These findings are consistent with the structural data.

RacA–RacA contacts: possible model for axial filament formation

A role in the formation of the elongated axial filament suggests that RacA must harbor non-specific DNA ability and also mediate RacA–RacA interactions between DNA molecules (17). Interestingly, examination of the packing of both RacA–DNA crystal forms revealed intimate RacA–RacA interactions between clusters of RacA molecules on adjacent RacA–DNA complexes, suggesting a possible mechanism for axial filament formation (Figure 5A). Indeed, even small clusters of bound RacA molecules could make cross contacts and contribute to DNA organization while allowing space for the flexible linker (estimated at 40–120 Å) to extend well beyond the DNA molecule. As a test of this model, we performed sedimentation analyses to assess whether the RacA DBD alone interacting with a centromere could form filamentous structures that pelleted. In these experiments the RacA DBD was mixed at ratios of 4:1 and 2:1 with a 54mer DNA site containing three consecutive centromere repeats, incubated at room temperature for

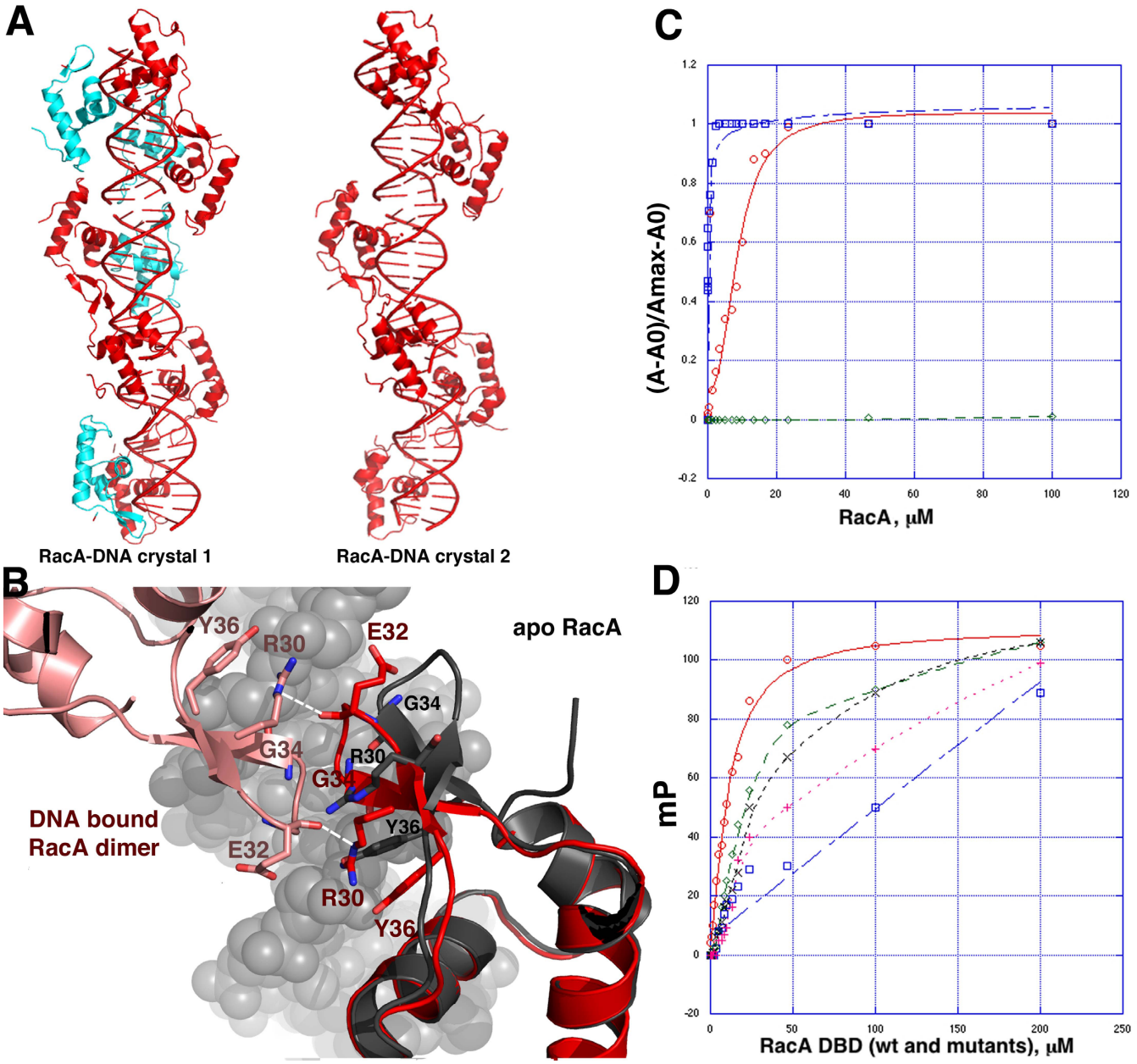

Figure 4. RacA DNA binding mechanism: RacA successively coats the DNA and DNA binding leads to conformational changes in the wings required for docking. (**A**) Comparison of the two RacA–DNA structures. The wing dimer bound molecules and DNA are colored red. The RacA molecules (cyan) that interact via α1–α1 contacts dock subsequent to the wing dimers. (**B**) Superimposition of one RacA apo structure (dark gray) onto the DNA bound RacA molecules. Shown is the close interaction between Gly34 residues required for DNA binding. This region adopts a different conformation in the apo state that appears to stabilize it in the absence of DNA. (**C**) Fluorescence polarization (FP) binding analysis of FL RacA (blue curve fit), RacA DBD (red) and the RacA C-domain (green). Measurements were performed in triplicate. (**D**) FP analysis of RacA DBD mutants. The red, green, blue, gray and pink curves correspond to binding by WT RacA DBD, RacA(K15A), RacA(R19A), RacA(30A) and RacA(G34E).

20 min and centrifuged at 40 000 rpm for 15 min. The results revealed that a significant fraction of the sample had sedimented (Supplementary Figure S7).

The RacA C-domain binds directly to the polar anchorage protein, DivIVA

Previous two-hybrid studies indicated an interaction between RacA and the DivIVA protein (31). As *B. subtilis* DivIVA localizes to the poles, this suggested that this interaction might function as the membrane tether for RacA (32,33). DivIVA consists of an N-domain, residues 1–65, which interacts with the membrane and a cytosolic located C-domain, residues 66–164 (33). Although the two-hybrid analysis suggested a linkage between RacA and DivIVA, to date no studies have determined whether this interaction is direct and the identity of the potential interacting domains. Interestingly, the DivIVA C-domain, which was implicated as the region involved in RacA interaction, con-

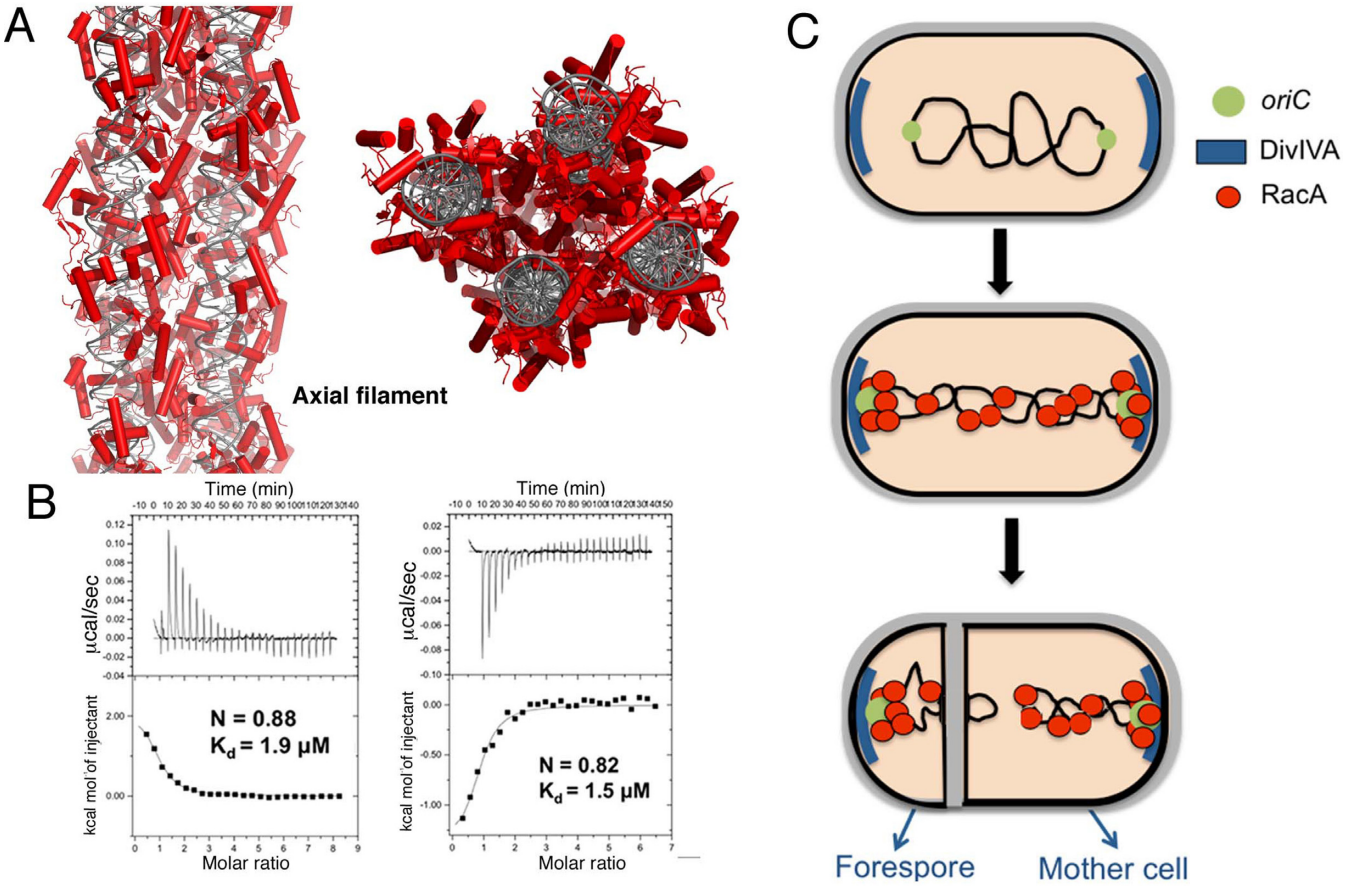

Figure 5. Structural mechanisms for axial filament formation and polar attachment by RacA. (**A**) Model of axial filament formation via RacA–RacA packing observed in both RacA–DNA structures. (**B**) ITC binding isotherms for, left, RacA C-domain binding to the C-domain of DivIVA and right, DivIVA C-domain binding to RacA C-domain. Both isotherms show a stoichiometry of one RacA CD-domain to one DivIVA C-domain with K_ds of 1.9 and 1.5 μM, respectively. (**C**) Molecular model for RacA anchoring, centromere binding and axial filament formation by RacA. Top shows a *Bacillus subtilis* cell before sporulation. Upon entry into sporulation, RacA (indicated by red circles) is expressed and binds specifically to the centromere near OriC (green) and also forms the axial filament by binding non-specifically to the chromosome. The non-specific contacts bring RacA clusters bound to the DNA in proximity and they then interact with each other as proposed by Figure 5A, leading to formation of the elongated, compact chromosome (the axial filament). The DNA is then pumped into the forespore.

tains a coiled coil that can tetramerize (33). This suggested that this DivIVA region might interact with the RacA C-domain, which also forms a coiled coil that can form higher order oligomers. To test this hypothesis we employed ITC. Specifically, we assessed the ability of the DivIVA C-domain (DivIVA(66–164)) to interact with the RacA DNA-binding and the RacA C-domain (71–184). These studies showed that the RacA DBD and DivIVA(66–164) do not interact (Supplementary Figure S8A). The RacA C-domain, however, bound robustly to the DivIVA(66–164) and *vice versa* (Figure 5B; Supplementary Figure S8B). These data establish that the RacA coiled coil interacts directly with DivIVA. Thus, this domain of RacA acts somewhat analogously to the C-terminal region of TnrA/GlnR, which bind glutamine synthetase as a helix and which is also connected to its DBD by a long, flexible linker (20).

DISCUSSION

Sporulation allows *B. subtilis* to survive conditions of severe environmental stress (1–5). Key to this process are cellular reprogramming events, including the remodeling of the chromosomes into the axial filament (4). RacA was isolated as a protein involved in this process as well as anchoring the origin regions at the cell poles. Here we report a detailed structural and biochemical characterization of this protein. These studies show that RacA is organized into two separable domains, an N-terminal DBD and C-terminal domain, which are connected by a proteolytically sensitive linker. This domain organization is unlike canonical MerR proteins but similar to TnrA/GlnR proteins. However, unlike MerR and TnrA/GlnR proteins, RacA functions as both a specific and non-specific DNA binding protein, the former role is in binding to the centromere and the latter in DNA organization. DNA binding architectural proteins that mediate the global organization of large DNA domains have not been as well characterized as those that mediate short range folding, such as integration host factor (IHF) and histone U93 (HU). Arguably, among the best studied global organizing DNA binding proteins are H-NS and MatP (32,33). Data indicates that H-NS bridges DNA substrates using both its C-terminal DNA-binding and N-terminal oligomerization domains (34). However, the molecular de-

tails remain unclear. Data showed that MatP organizes the Ter (terminus containing) macrodomain of the *E. coli* chromosome by binding to specific DNA sites. High resolution structures of MatP–DNA complexes revealed that is has a similar domain arrangement as RacA; it consists of an N-terminal DBD connected to a coiled coil C-domain (35). However, the DBD of MatP is unrelated to that of RacA and in MatP there is no disordered linker between its DNA-binding and coiled coil domains. The coiled coils of the MatP dimers associate to form tetramers, which mediate long range bridging within the Ter DNA domain, leading to loose compaction of the DNA. RacA appears to contribute to formation of a more densely packed DNA structure than MatP and H-NS (Figure 5C). In our structures RacA DBDs interact closely with other RacA DBDs bound on adjacent DNA, suggesting a mechanism for this process of DNA compaction (Figure 5A and C).

RacA is unusual for a DNA binding protein in that its roles require that it mediate both specific and non-specific DNA binding, the former for centromere binding and the latter to aid in formation of the axial filament. Our structures reveal that RacA contains a winged-HTH DNA binding motif and the structures captured RacA binding to different sequences, the specific *ram* centromere site and non-specific sequences. Strikingly, the same residues are utilized in both specific and non-specific binding. A key feature of this binding mode is the large number of phosphate contacts that permit docking onto the DNA. In addition, once bound to the DNA the RacA DBD can interact in multiple ways with other RacA DBDs to aid in DNA binding. One residue, Arg19 is observed to make the key specific contacts to the TG bps at the ends of the *ram* site. When bound to non-*ram* sites, the Arg19 rotates and participates in phosphate contacts. The usage of one flexible residue to make the key specific contacts and a large number of residues that can dock adaptively to interact with the phosphate backbone explains how RacA can bind DNA both specifically and non-specifically. Finally, the presence of the flexible linker between the DNA-binding and C-domain allows the domains to act as separate entities with the distinct functions. Thus, in conclusion, our studies unveil that RacA employs a unique DNA binding modality in which the same residues can be adjusted to bind DNA specifically or non-specifically and a modular domain architecture that permits two separate functions in sporulation.

ACCESSION NUMBERS

Structure factor amplitudes and coordinates have been deposited in the PDB under the accession codes 5I41 and 5I44.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank the staff at ALS beamline 8.3.1, especially Jane Tanamachi for assistance with beam line matters and scheduling.

FUNDING

Beamline 8.3.1 at the Advanced Light Source is operated by the University of California Office of the President, Multicampus Research Programs and Initiatives [grant MR-15-328599] and Program for Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation. Additional support comes from National Institutes of Health [GM105404, GM073210, GM082250, GM094625]; National Science Foundation [1330685]; Plexxikon Inc. and the M.D. Anderson Cancer Center.
Conflict of interest statement. None declared.

REFERENCES

1. Stragier,P. and Losick,R. (1996) Molecular genetics of sporulation in *Bacillus subtilis*. *Annu. Rev. Genet.*, **30**, 297–341.
2. Pavlendova,N., Muchova,K. and Barak,I. (2007) Chromosome segregation in *Bacillus subtilis*. *Folia Microbiol.*, **52**, 563–572.
3. Levin,P.A. and Brossman,A.D. (1998) Cell cycle and sporulation in *Bacillus subtilis*. *Curr. Opin. Microbiol.*, **1**, 630–635.
4. Errington,J. (2001) Septation and chromosome segregation during sporulation in *Bacillus subtilis*. *Curr. Opin. Microbiol.*, **1**, 117–126.
5. Piggot,P.J. and Hilbert,D.W. (2004) Sporulation of *Bacillus subtilis*. *Curr. Opin. Microbiol.*, **7**, 579–586.
6. Horvitz,H.R. and Herskowitz,I. (1992) Mechanisms of asymmetric cell division: two Bs or not two Bs, that is the question. *Cell*, **68**, 237–255.
7. Way,J.C. (1996) The mechanism of bacterial asymmetric cell division. *Bioessays*, **18**, 99–101.
8. Roubinet,C. and Cabernard,C. (2014) Control of asymmetric cell division. *Curr. Opin. Cell Biol.*, **31**, 84–91.
9. Ryter,A., Schaeffer,P. and Ionesco,H. (1966) Classification cytologique, par leur stade de blocage, des mutants de sporulation de *Bacillus subtilis* Marburg. *Ann. Inst. Pasteur (Paris)*, **110**, 305–315.
10. Glaser,P., Sharpe,M.E., Raether,B., Perego,M., Ohlsen,K. and Errington,J. (1997) Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes Dev.*, **11**, 1160–1168.
11. Webb,C.D., Teleman,A., Gordon,S., Straight,A., Belmont,A., Lin,D.C., Grossman,A.D., Wright,A. and Losick,R. (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell*, **88**, 667–674.
12. Lin,D.C. and Grossman,A.D. (1998) Identification and characterization of a bacterial chromosome partitioning site. *Cell*, **92**, 675–685.
13. Wu,L.J. and Errington,J. (1997) Septal localization of the SpoIIIE chromosome partitioning protein in *Bacillus subtilis*. *EMBO J.*, **16**, 2161–2169.
14. Wu,L.J. and Errington,J. (2003) RacA and the Soj-Spo0J system combine to effect polar chromosome segregation in sporulating *Bacillus subtilis*. *Mol. Microbiol.*, **49**, 1463–1475.
15. Ben-Yehuda,S., Rudner,D.Z. and Losick,R. (2003) Assembly of the SpoIIIE DNA translocase depends on chromosome trapping in *Bacillus subtilis*. *Curr. Biol.*, **13**, 2196–2200.
16. Ben-Yehuda,S., Rudner,D.Z. and Losick,R. (2003) RacA, a bacterial protein that anchors chromosomes to the cell poles. *Science*, **299**, 532–536.
17. Ben-Yehuda,S., Fujita,M., Liu,X.S., Gorbatyuk,B., Skoko,D., Yan,J., Marko,J.F., Liu,J.S., Eichenberger,P., Rudner,D.Z. *et al.* (2005) Defining a centromere-like element in *Bacillus subtilis* by identifying the binding sites for the chromosome-anchoring protein RacA. *Mol. Cell*, **17**, 773–782.
18. Brünger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
19. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Echols,N., Headd,J.J., Hung,L.W., Jain,S., Kapral,G.J., Grosse Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Phyton-based system for macromolecular structure solution. *Acta Crystallogr. D*, **66**, 213–221.

20. Schumacher,M.A., Chinnam,N.B., Cuthbert,B., Tonthat,N.K. and Whitfill,T. (2015) Structures of regulatory machinery reveal novel molecular mechanisms controlling *B. subtilis* nitrogen homeostasis. *Genes Dev.*, **29**, 451–464.
21. Brown,N.L., Stoyanov,J.V., Kidd,S.P. and Nobman,J.L. (2003) The MerR family of transcription regulators. *FEMS Microbiol. Rev.*, **27**, 145–163.
22. Heldwein,E.E. and Brennan,R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature*, **409**, 378–382.
23. Lamoureux,J.S., Maynes,J.T. and Glover,J.N. (2004) Recognition of 5′-YpG-3′ sequences by coupled stacking/hydrogen bonding interactions with amino acid residues. *J. Mol. Biol.*, **335**, 399–408.
24. Lamoureux,J.S., Stuart,D., Tsang,R., Wu,C. and Glover,J.N. (2002) Structure of the sporulation-specific transcription factor Ndt80 bound to DNA. *EMBO J.*, **21**, 5721–5732.
25. Zheng,G., Lu,X.J. and Olson,W.K. (2009) Web 3DNA- a web server for the analysis, reconstruction and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **37**, W240–W246.
26. Schumacher,M.A. (2012) Bacterial plasmid partition machinery: a minimalist approach to survival. *Curr. Opin. Struct. Biol.*, **22**, 72–79.
27. Leonard,T.A., Butler,P.J. and Löwe,J. (2004) Structural analysis of the chromosome segregation protein Spo0J from *Thermus thermophilus*. *Mol. Microbiol.*, **53**, 419–432.
28. Lobocka,M. and Yarmolinsky,M. (1996) P1 plasmid partition: a mutation analysis of ParB. *J. Mol. Biol.*, **259**, 366–382.
29. Zhou,NE, Kay,CM and Hodges,RS (1994) The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Eng.*, **7**, 1365–1372.
30. Sutherland,T.A., Church,J.S., Hu,X., Huson,M.G., Kaplan,D.L. and Weisman,S. (2011) Single honeybee silk protein mimics properties of multi-protein silk. *PLoS One*, e16489.
31. Van Baarle,S., Celik,I.N., Kaval,K.G., Bramkamp,M., Hamoen,L.W. and Halbedel,S. (2013) Protein-protein interaction domains of *Bacillus subtilis* DivIVA. *J. Bacteriol.*, **195**, 1012–1021.
32. Lenarcic,R., Halbedel,S., Visser,L., Shaw,M., Wu,L.J, Errington,J., Marenduzzo,D. and Haomen,L.W. (2009) Localisation of DivIVA by targeting to negatively curved membranes. *EMBO J.*, **28**, 2272–2282.
33. Oliva,M.A., Halbedel,S., Freund,S.M., Dutow,P., Leonard,T.A., Veprintsev,D.B., Hamoen,L.W. and Löwe,L. (2010) Features critical for membrane binding revealed by DivIVA crystal structure. *EMBO J.*, **29**, 1988–2001.
34. Arold,S.T., Leonard,P.G., Parkinson,G.N. and Ladbury,J.E. (2010) H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 15728–15732.
35. Dupaigne,P., Tonthat,T.K., Espeli,O., Whitfill,T., Boccard,F. and Schumacher,M.A. (2012) Molecular basis for protein-mediated DNA-bridging mechanism that functions in condensation of the *E. coli* chromosome. *Mol. Cell*, **48**, 560–571.