The MerR-like protein BldC binds DNA direct repeats as cooperative multimers to regulate Streptomyces development

Maria A. Schumacher1, Chris D. den Hengst2, Matthew J. Bush2, T.B.K. Le2, Ngat T. Tran2, Govind Chandra2, Wenjie Zeng1, Brady Travis1, Richard G. Brennan1 & Mark J. Buttner2

Streptomycetes are notable for their complex life cycle and production of most clinically important antibiotics. A key factor that controls entry into development and the onset of antibiotic production is the 68-residue protein, BldC. BldC is a putative DNA-binding protein related to MerR regulators, but lacks coiled-coil dimerization and effector-binding domains characteristic of classical MerR proteins. Hence, the molecular function of the protein has been unclear. Here we show that BldC is indeed a DNA-binding protein and controls a regulon that includes other key developmental regulators. Intriguingly, BldC DNA-binding sites vary significantly in length. Our BldC-DNA structures explain this DNA-binding capability by revealing that BldC utilizes a DNA-binding mode distinct from MerR and other known regulators, involving asymmetric head-to-tail oligomerization on DNA direct repeats that results in dramatic DNA distortion. Notably, BldC-like proteins radiate throughout eubacteria, establishing BldC as the founding member of a new structural family of regulators.
Streptomyces are ubiquitous, primarily soil-dwelling filamentous bacteria that undergo a complex developmental transition from vegetative growth to the production of reproductive aerial hyphae, which differentiate into chains of exospores. Entry into development coincides with the biosynthesis of numerous secondary metabolites that serve as our most abundant source of clinically important antibiotics and provide other medically important drugs such as antitumor agents and immunosuppressants. As a consequence, there is considerable interest in understanding the mechanisms that control this developmental transition. Genetic studies identified the regulatory loci that control entry into development, which are called bld (bald) genes because null mutations in these loci prevent the formation of fuzzy aerial hyphae for one of two diametrically opposite reasons: either because they block differentiation (mutations in activators) or because they cause precocious hyper-sporulation, bypassing the formation of aerial hyphae (mutations in repressors). The three Bld regulators that fall into this latter class are BldD, BldO, and BldC. BldD is a transcriptional repressor that sits at the top of the developmental hierarchy and is regulated by cyclic-di-GMP (c-di-GMP). c-di-GMP mediates the dimerization of two BldD protomers, leading to DNA binding. This activation, specifi-
cally the conservon between the −10 and −35 promoter elements such that these motifs are misaligned for binding by the α factor of the RNA polymerase (RNAP) holoenzyme, rendering the promoters inactive for transcription. Structures for several canonical MerR proteins have been solved in their apo and DNA-bound activated forms and reveal a conserved mode of transcription activation. Specifically, classical MerR proteins bind operators in the spacer between the −10 and −35 regions and when the C-terminal effector-recognition domain binds its cognate ligand, the transcription factor untwists and shortens the DNA, realigning the −10 and −35 sequences to allow RNAP holoenzyme to bind and activate transcription.

While BldC appears to contain a MerR-like wHTH, it is to date, the only known MerR-like protein that consists entirely of a wHTH with no obvious effector or oligomerization domain. Interestingly, bioinformatic analysis shows that there are numerous BldC homologs in both Gram-negative and Gram-positive bacteria, suggesting that BldC represents a large family of putative DNA-binding proteins. Despite this, the roles of BldC and its homologs remain unknown. Using a battery of structural, biochemical, and in vivo approaches, we show here that BldC functions as a pleiotropic regulator of Streptomyces development by employing a unique mode of DNA binding for a transcription factor that involves asymmetric head-to-tail oligomerization on DNA direct repeats of varying number with concomitant distortion of the DNA. This mode of DNA binding defines a new family of transcription regulatory proteins that we designate the BldC family.

Results

Identification of BldC binding sites in vivo. The presence of a putative wHTH in BldC suggests it functions in DNA binding and transcription regulation. Therefore, to identify possible BldC binding sites in vivo, we performed chromatin immunoprecipitation-microarray (ChIP-chip) using a polyclonal BldC antibody. Figure 1a,b and Supplementary Data 1 present the results from two independent biological experiments. A congenic bldC null mutant was used as a control to eliminate any potential signals that might arise from cross-reaction of the antibody with other transcription factors. ~280 BldC-specific peaks were detected (P-value <= 0.05, P-value based on empirical Bayes moderated t-statistic test), scattered across the genome (Fig. 1a and Supplementary Data 1). 25 BldC target genes encode regulatory proteins themselves (Supplementary Table 1), implying a pleiotropic role for BldC in Streptomyces development. Indeed, promoter sites bound by BldC are found upstream of genes encoding many key transcriptional regulators of the Streptomyces developmental cascade including bldM, whiB, whiD, whiH, whiL, sigF, and bldC itself, in addition to others encoding proteins involved in chromosome segregation and condensation during sporulation such as smeA-sffA and hupS. Figure 1b and Supplementary Data 1 and genes that influence antibiotic production, specifically the conservon cvnA1 (ref. 27) and the serine-threonine protein kinase, afsK. In parallel work, we used qRT-PCR to examine the expression of 5 key BldC targets (whiL, smeA, sigF, and hupS) in Streptomyces venezuelae comparing the wild type (WT) and a congenic bldC mutant. In each case, the gene was expressed earlier in the bldC mutant than in the WT, indicating that the transcription of these genes is directly regulated by BldC. Both BldC and BldD inhibit entry into development, and previously we identified ~160 genes controlled by BldD using ChIP-chip. Comparison of genes bound by BldC or BldD showed only a small overlap of 15 genes/operons bound by both proteins (Supplementary Table 2). However, these include the key developmental genes whiB, whiD, bldM, smeA-sffA, bldC, and cvnA1.

BldC binding to the whil promoter. To glean insight into the DNA binding mode of BldC, we performed DNase I protection studies on the promoter regions of two important BldC-regulated loci: whil and the smeA-sffA operon. whil is among the key developmental targets of BldC (Fig. 1b and Supplementary Data 1), being essential for the late stages of sporulation and is involved in chromosome segregation and condensation during sporulation such as smeA-sffA and hupS. In parallel work, we used qRT-PCR to examine the expression of 5 key BldC targets (whiL, smeA, sigF, and hupS) in Streptomyces venezuelae comparing the wild type (WT) and a congenic bldC mutant. In each case, the gene was expressed earlier in the bldC mutant than in the WT, indicating that the transcription of these genes is directly regulated by BldC. Both BldC and BldD inhibit entry into development, and previously we identified ~160 genes controlled by BldD using ChIP-chip. Comparison of genes bound by BldC or BldD showed only a small overlap of 15 genes/operons bound by both proteins (Supplementary Table 2). However, these include the key developmental genes whiB, whiD, bldM, smeA-sffA, bldC, and cvnA1.

While BldC appears to contain a MerR-like wHTH, it is to date, the only known MerR-like protein that consists entirely of a wHTH with no obvious effector or oligomerization domain. Interestingly, bioinformatic analysis shows that there are numerous BldC homologs in both Gram-negative and Gram-positive bacteria, suggesting that BldC represents a large family of
suggesting that BldC binding may distort DNA (Fig. 2). Notably, the enhanced cleavages also appear periodic in nature, falling on approximately the same face of the DNA and within the imperfect direct repeats (Fig. 2). To gain more insight into the specific sequence and minimal length requirement for high affinity DNA binding by BldC, we performed EMSA analyses employing a variety of DNA duplexes based on the whiI site that ranged from 16 to 30 bp (Supplementary Fig. 1). These analyses identified the 22-bp sequence, top strand 5′-TGTGTCCGAATTGCTCGGATTG-3′ (Supplementary Fig. 1). Two of these changes are naturally present in the whil promoter of the alternative model species S. venezuelae and the amino acid sequences of BldC from Streptomyces coelicolor and S. venezuelae are identical.

**BldC binding to the smeA-sffA promoter.** Another key target regulated by BldC is the smeA-sffA operon (Fig. 1b and Supplementary Data 1). The smeA-sffA operon encodes a DNA translocase (SffA) involved in chromosome segregation during sporulation that is specifically targeted to septa by the small membrane protein SmeA. Deletion of smeA-sffA results in a defect in spore chromosome segregation and has pleiotropic effects.

---

**Fig. 1** The S. coelicolor BldC regulon. a Chromosome-wide distribution of BldC binding sites identified by ChIP-chip analysis. DNA obtained from immunoprecipitation of BldC was labeled with Cy3 and hybridized to DNA microarrays together with a total DNA control labeled with Cy5. Data are plotted as Cy3/Cy5 ratios (y-axis), as a function of chromosome location (x-axis). b ChIP-chip data for nine selected BldC targets in WT S. coelicolor and the S. coelicolor ΔbldC mutant (blue and red dots, respectively). Plots span approximately 8 kb of DNA sequence. Gene names or identifiers (SCO numbers) are indicated below the arrows, which indicate gene orientation.
Strikingly, in contrast to the ~30 bp BldC-protected region in the \( \text{whiI} \) promoter, BldC protects a much larger region of the \( \text{smeA-ssfA} \) promoter, extending approximately from -60 to -10 relative to the transcription start site. A regular pattern of hypersensitive sites on the top strand is observed in the footprint (Fig. 3a,b). While the BldC-binding site in the \( \text{whiI} \) promoter contains two direct repeat elements, the BldC binding site within the \( \text{smeA-ssfA} \) promoter reveals four similar repeats, but arranged in the opposite orientation, 5' to 3' (see Fig. 3a,b). To explore BldC binding to the \( \text{smeA-ssfA} \)
a promoter further, a finer analysis was undertaken using hydroxyl radical footprinting. BldC protects four regularly spaced 5-bp tracts, with one helical turn of the DNA between one tract and the next (Fig. 3b,c), suggesting that four molecules of BldC bind the smeA-ssfA promoter. However, only one shift is observed in EMSA analyses of BldC binding to the smeA-ssfA site, suggesting that BldC binds this extended site cooperatively (Supplementary Fig. 2).

BldC-DNA structure reveals MerR-like structural family. The mechanism by which BldC binds sites with variable numbers of direct repeat elements is unclear, but it is likely to differ from that used by other MerR transcriptional regulators, which function as symmetric dimers that bind palindromic DNA sites. Consistent with this hypothesis, the BldC protein does not contain a canonical MerR family dimerization domain. Further, unlike classical MerR proteins, size exclusion chromatography (SEC) studies showed that apo BldC is a monomer (Supplementary Fig. 3). Thus, to deduce the molecular mechanism underlying DNA binding and recognition by BldC, we determined the structure of S. coelicolor BldC bound to a 22 bp double-stranded DNA based on the optimized whiI binding site (whiI opt). The structure was solved by single wavelength anomalous diffraction (SAD) using data collected from a selenomethionine-substituted BldC(L43M-L58M)-whiI opt DNA crystal (WT BldC-whiI opt produced the same crystals) and refined to Rwork/Rfree values of 22.3%/26.8% to 3.28 Å resolution (Supplementary Table 3 and Supplementary Fig. 4). Notably, the structure reveals that BldC forms a head to tail (head-tail) dimer on the DNA whereby each BldC protomer binds to a 9 bp direct repeat with the recognition helix of the HTH motif interacting in the major groove and the wing in the minor groove (Fig. 4a,b). Because the BldC dimer is asymmetric it binds head-tail with a specific directionality, which is 5′ to 3′ with the sequence, 5′-CAATTCGGACATTTACATGCATGTGAGC-3′ (Fig. 4d,e and Supplementary Fig. 1). Formation of the head-tail BldC dimer buries ~600 Å² of protein surface from solvent. The oligomer interface is largely hydrophobic with residues Phe21, Val23, and Trp31 packing against the side chains of Leu11’, Ile 40’, Thr42’, Leu43’, and Gly44’ (where ’ indicates the other subunit of the DNA-stabilized dimer) (Fig. 4c and Supplementary Fig. 5). Hydrogen bonds/electrostatic contacts between the side
chains of Arg22 and Glu16 help fasten the two subunits together (Fig. 4c and Supplementary Fig. 5). While the buried surface area of the DNA-bound BldC dimer is significantly distorted: the wing-bound, AT-rich minor groove has an average width of 9.3 Å (compared to 12.0 Å for B-DNA) whereas the minor groove outside this region is widened significantly (14.5 Å). The major groove in which the recognition helix docks is expanded to 19.4 Å compared 17.2 Å for B-DNA. Despite employing a DNA binding mode different from classical MerR proteins, the BldC wHTH indeed harboors a fold that is highly similar to these proteins. The BldC structure follows the typical MerR topology: α1 (BldC residues 14–21), α2...
Molecular basis of DNA binding specificity of BldC. Each BldC subunit makes a large number of phosphate and base contacts, allowing for high affinity BldC binding. The side chains of BldC residues Thr27, Thr29, and Lys33 and the amide nitrogens of Ala15 and Asp24 from the HTH anchor the wHTH on the DNA via phosphate contacts to both DNA strands. The Trp31 side chain also interacts with the phosphate backbone via its Nε and induces a kink in the DNA (Fig. 4d,e). The BldC wHTH is particularly rich in basic residues with Arg47 and Arg48 providing key contacts that aid in docking the wing into the minor groove. The combined DNA-binding region of the two interacting BldC subunits also presents a striking electropositive surface that complements the bound DNA (Fig. 4b).

The DNA-binding specificity of BldC is mediated by residues from both its wing and its HTH. Residues from the wing recognize the AT-rich narrowed minor groove. The His46 side chain fits within the narrowed groove and makes numerous hydrogen bonds and van der Waals contacts with sugar and phosphate moieties (Fig. 4d). The importance of the AT-rich motif in BldC binding sites was assessed by fluorescence polarization (FP) binding assays. These studies showed that replacement of the AATT motif with GGCC abrogated DNA binding (Supplementary Fig. 7). Major groove base contacts are provided by BldC residues Arg30 and Lys26, both located on the recognition helix of the HTH. The relatively low resolution of the structure prevents a highly detailed description of protein-DNA contacts and the possible role of waters in the protein-DNA interface. However, despite the low resolution, the electron density for the Arg30 side chain is clear and reveals that it provides highly specific, bidentate hydrogen bonds to Gua2' and Gua1' (Fig. 4d,e and Supplementary Fig. 8). Simultaneously, the Arg30 guanidinium groups stack with the bases of Thy1' and Thy10', which are located 5' of the contacted guanine (Fig. 4d,e and Supplementary Fig. 9). This type of specific protein-DNA contact has been called a 5'-pyrimidine-guanine-3' (5'-YpG-3') interaction and it arises from the inherent flexibility of pyrimidine-guanine steps. A survey of the protein database by Glover and coworkers showed that diverse classes of DNA-binding motifs utilize this DNA recognition element. This type of contact mediates specific binding to two bases whereby an arginine side chain interacts with the major groove face of the 5'-guanine nucleobase and also contacts the unstacked preceding pyrimidine. The side chain of BldC residue Lys26 also makes base contacts. However, the electron density for the Lys26 side chain is less well resolved than that of the Arg30 side chain suggesting it is conformationally flexible. Moreover, while the Lys26 side chain is within hydrogen bonding distance of major groove guanines from bps 7, 8, and 17 (Fig. 4e), it could also interact with thymine O4 atoms in other sequences, suggesting that it may not dictate complete readout specificity of major groove bases. Indeed, binding studies revealed that replacement of the guanines at bp 8 and bp 17 with thymine (Fig. 4e) had little effect on binding (app \( K_d \) of WT = 20 nM compared to 50 nM ± 1.1 nM for the T mutant), while replacement of the guanines with cytosines resulted in a nearly 6-fold reduction in binding (app \( K_d \) = 115 nM ± 10 nM) (Supplementary Fig. 7). Hence, the flexible contacts provided by this residue may contribute to the range of sequences that BldC can bind within select promoters.

Thus, the analysis of the BldC-whil opt structure indicates that there are two main DNA elements specifically recognized by BldC: an AT-rich region and a C-G, 4 base pairs downstream from the AT-rich region, which is recognized by Arg30. The consensus direct repeat for binding by each BldC protomer is, therefore, AATTXXXXX(C), where the C represents the C-G bp containing the guanine recognized by Arg30. Because the BldC dimer binds direct repeats in a head-tail manner, its interaction with the DNA is also directional. As a result, BldC binds whil opt and the WT whil promoter in a reverse orientation, i.e., the BldC dimer is arranged head-tail, 5' to 3' on the bottom strand of the DNA (Figs. 2, 4a).

Probing the BldC-DNA structural model. The BldC-DNA structure revealed several unexpected findings for a MerR-like protein, in particular a direct repeat DNA binding mode involving the formation of a head-tail dimer. Thus, to further test the structural model, we mutated several residues observed to be important for DNA binding in the structure as well as those involved in dimer formation and carried out FP DNA binding assays. The WT protein, with or without the his6-tag, and the selenomethionine BldC(L43M-L58S) protein bound the 22 bp whil opt DNA with essentially the same \( K_d \) (~20 nM) (Fig. 4f). An E16R mutation, which would disrupt the dimer stabilizing contact with Arg22 (Fig. 4c), resulted in a ~10-fold reduction in DNA binding while a G44E substitution, which places a large, negative, charged residue at the hydrophobic interface of the dimer and hence is predicted to disrupt the head-tail dimer, led to complete loss of DNA binding (Fig. 4f). Individual substitution of residues His46 and Arg30, which make key contacts to the minor and major grooves, respectively, with a glutamate (H46E) or an alanine (R30A), also essentially abrogated DNA binding (Fig. 4f). Thus, these combined data support the unusual mode of DNA binding observed in the BldC-DNA structure. Importantly, circular dichroism (CD) analyses showed that all mutant BldC proteins were properly folded and exhibited the same CD spectrum as WT BldC (Supplementary Fig. 10).

BldC-smea-ssfA structure shows extended protein-DNA filament. The head-tail mode of DNA binding observed in the BldC-whil opt structure (Fig. 4) could permit the formation of continuous protein–DNA complexes through extension at one or both ends of the asymmetric dimer, which could explain the mechanism which BldC utilizes to interact with longer DNA binding sites such as that of the smea-ssfA promoter. The BldC protected region in the smea-ssfA promoter region contains four AATTXXXXX(C) motifs similar those in the whil site (5'-AATTCGGTC-3', 5'-GATTTCCCC-3', 5'-CATTTGCA-3', and 5'-CTTTTATC-3'), where the C-G bp recognized by Arg30 is underlined) (Fig. 3). But, in the smea-ssfA promoter these repeats proceed in a 5' to 3' direction and the last repeat does not conform to the consensus (Fig. 2). However, conservation of the exact consensus sequence does not appear to be critical for specific binding, in particular for the AT-rich region, as it is the narrowing of the minor groove caused by the AT-rich nature of that sequence that is important. This so-called indirect readout based on the shape of the minor groove as well as the pliability in contacts afforded by the Lys36 side chain in the major groove likely allows BldC to bind DNA sites that harbor differences from the consensus. As a result, BldC might be able to interact with sites that contain repeats that diverge from the consensus in longer DNA sites provided there are consensus sites to enable the
docking of subunits that can then enable binding by other promoters. This binding mode predicts that multiple BldC subunits might bind a DNA site such as the smeA-ssfA operator. To address this issue, we determined the $K_d$ and stoichiometry of BldC binding to the 22mer whiI opt site, to an extended 36-mer opt site that contains four perfect direct repeats and to the smeA-ssfA site. BldC bound these DNA sites with apparent affinities ($K_d$) of 20, 20, and 60 nM, respectively, and displayed binding stoichiometries (BldC subunits per DNA duplex) of 2:1, 4:1, and 4:1, respectively (Fig. 5a). These data, therefore, support the hypothesis that multiple BldC subunits can bind cooperatively to extended DNA sites.

To deduce the molecular basis for this extended binding we next determined the structure of BldC in complex with an 18mer smeA-ssfA DNA fragment that contained the first two repeats of the smeA-ssfA BldC binding site, 5'-GCAATTCCGGTC-GAATTTC-3' (Fig. 3b). The DNA was constructed to generate pseudo-continuous packing of the DNA, which was indeed generated in the crystals (see Methods). The structure was solved by molecular replacement (MR) and refined to $R_{work}$/$R_{free}$ values...
The binding of multiple BldC subunits along the DNA also leads to severe DNA distortion and shortening: the BldC subunits reduce the length of the DNA by a third when compared to canonical B-DNA (Fig. 6b). Consequently, the DNA has an overall positive writhe.

Discussion

Our combined data reveal that while BldC harbors a MerR-like whtH motif, it binds DNA in a manner distinct from classical MerR family proteins. Interestingly, recent structural studies on other non-MerR proteins have revealed whtH domains with structural similarity to the MerR whtH. For example, structural analyses of the master regulators of nitrogen metabolism in B. subtilis, TnrA and GlnR, revealed that both proteins have a MerR-like whtH motif, however both TnrA and GlnR lack the coiled-coil domain of classical MerR proteins but instead contain N-terminal extensions and unstructured C-terminal regions. Instead of using the canonical coiled-coil, TnrA and GlnR dimerize upon DNA binding through residues in their N-terminal regions and use their C-terminal regions to interact with glutamine synthetase (GS). In common with canonical MerR proteins, however, both TnrA and GlnR bind as homodimers to palindromic DNA sites. Structural homology searches also uncovered architectural DNA binding proteins that do not function in transcription, including RacA and Xis, which harbor whtH structures similar to MerR proteins.

Interestingly, BldC appears similar to classical MerR proteins in that it makes few base specific contacts. Indeed, the structures of BmrR, MtaN, and CueR bound to DNA revealed that only one to two residues make base contacts per subunit. This is consistent with MerR proteins employing indirect readout as part of their DNA binding mechanisms. Notably, unlike canonical MerR regulators, the MerR-like protein RacA was observed to make YpG contacts using an arginine that is equivalent to the BldC residue Arg30. However, aside from this similarity, RacA and BldC utilize very different DNA binding mechanisms. In particular, the interaction of RacA with its GC-rich DNA site does not lead to global DNA bending but causes a significant widening of its entire minor groove. By contrast, BldC binding induces significant global bending in its bound DNA and a narrowing of the AT-rich minor grooves. Also, unlike RacA, which binds DNA as a symmetric dimer, BldC binds direct repeats in a head-tail manner. To our knowledge, BldC is the only MerR-like transcription regulator that binds direct repeats and employs only its whtH domain for DNA contact and oligomerization. Bacterial transcription regulators that bind DNA direct repeats appear less common than those that bind palindromic DNA sites as symmetric dimers. Well characterized examples of bacterial regulators, other than BldC, that bind direct repeats include some members of the large family of two component response regulators (RR). RR harbor N-terminal receiver domains flexibly attached to DNA-binding domains. More than 95% of characterized RR utilize a HTH DNA-binding domain while the remaining 5% of RR employ a LytTR domain for DNA binding. Structures have been obtained for both these types of DNA-binding domain, revealing that some RR, such as PhoP, PmrA, and AgrA, bind DNA direct repeats. Studies on the RR AcrA revealed that, like BldC, it can bind multiple tandem repeats, however a structure is not yet available of the full length AcrA bound to cognate DNA elements.

Finally, bioinformatic analyses reveal that small, BldC-like proteins radiate throughout the domain of bacteria, being present in ~500 species, including actinomycete pathogens (Mycobacterium tuberculosis and Corynebacterium diphtheriae), commercially important actinomycetes (Corynebacterium glutamicum), and...
other Gram-positive (e.g., *Bacillus thuringiensis*) and Gram-negative bacteria (e.g., *Sinorhizobium meliloti* and *Porphyromonas gingivalis*) (Supplementary Fig. 11a). Sequence analysis shows strong conservation of key DNA-binding and head-tail oligomerization residues in BldC-like proteins in the actinomy- cetes. More distant members of the BldC family also contain hydrophobic residues in regions involved in head-tail dimer formation (Supplementary Fig. 11a). However, key residues involved in BldC head-tail dimer formation are not conserved in classical MerR family proteins. In particular, Leu43, which interacts with Trp31 and Phe21 to form the key hydrophobic core of the BldC head-tail dimer, is replaced by a charged residue (glutamic or aspartic acid) in ZntR, CueR, MtaN, and BmrR, or by a polar residue (serine, tyrosine or histidine) in SoxR, MerR, and TipAL (Supplementary Fig. 11b). Modeling indicates that such charged or polar residues would prohibit formation of a head-tail dimer as they would be placed proximal to hydrophobic side chains. These findings suggest that, unlike classical MerR proteins, the other BldC-like proteins identified bioinformatically are likely to have the same DNA binding mode observed for BldC itself, and that they can likewise form extended protein-nucleic acid superstructures. In that regard, it is interesting to note that some BldC-like orthologues in the database are annotated as possible resolvases/integrases, consistent with our finding that BldC shares strong structural similarity to Xis. These findings suggest formerly overlooked, possible evolutionary connections between MerR transcription regulators and integrases and DNA architectural proteins.

**Methods**

**Cell culture, strains, plasmids, and oligonucleotides.** Strains and plasmids used in this study are listed in Supplementary Table 1. Oligonucleotides used are listed in Supplementary Table 2. Oligonucleotides given in the text. Streptomyces coelicolor strains were grown on mannostol soy flour (MS, also known as SFM) or RS solid media42, 43. Strains used for ChIP-chip analysis were grown in a 1:1 mixture of yeast extract and tryptic soy-broth (TSB) liquid media42 at 30°C.

**Chromatin immunoprecipitation-microarray studies.** To carry out the ChIP-chip experiments, *S. coelicolor* M600 was grown in duplicate for 15 h in YEME/TSB liquid medium and a single culture of the congenic *bldC* null mutant strain J2166 was grown in the same way as the control. Formaldehyde was added to cultures at a final concentration of 1% (v/v) and incubation was continued for 30 min. Glycine was then added to a final concentration of 125 mM to stop the cross-linking. Cultures were left at room temperature (RT) for 5 min before the mycelium was harvested and washed twice in PBS buffer pH 7.4. Each mycelial pellet was resuspended in 0.5 ml buffer (10 mM Tris HCl pH 8.0, 50 mM NaCl) containing 15 mM MgCl₂, 1 mM dithiothreitol, and 1 µg poly(dI-dC) (Roche Applied Science) and incubated at 25°C for 1 h. Subsequently, 0.5 ml IP lysis buffer (100 mM Tris HCl pH 8, 250 mM NaCl, 0.5% Triton X-100, 0.1% SDS) containing protease inhibitor (Roche Applied Science) was added and samples were incubated for 10 min on ice. Samples were sonicated for 7 cycles of 15 s each at 100 W to shear the chromosomal DNA into fragments ranging from 3000–1000 bp in size. Samples were centrifuged twice at 30,000 × g at 4°C for 15 min to clear the cell extract, after which 10 µl of cell extract was set aside for total DNA extraction. The remainder (990 µl) was incubated with 45 µl protein A-agarose (Sigma) for 1 h on a rotating wheel to clear from non-specifically binding proteins. Samples were centrifuged for 15 min at 4°C at 14,000 × g and the supernatant was saved. Supernatants were incubated with 75 µl crude anti-BldC antiserum (76 mg/ml yeast tRNA). Samples were incubated for a further 4 min at RT before quenching the reaction with 9 µl stop solution, phenol chloroform extracted and analyzed in the same way as the DNase I footprints. Hybridization to microarrays was carried out using a glass slide spotted by Perkin Elmer (Waltham, MA). DNA labeling and hybridization to DNA microarrays were carried out by using a BioPrime kit (Invitrogen). Briefly, for these experiments 800 ng of the total and immuno-precipitated DNA was labeled with Cy5-dCTP and Cy3-dCTP, respectively. Using an Agilent Technologies hybridization oven, the labeled DNA was hybridized to high-density DNA microarrays representing the genome of *S. coelicolor*, which were designed and manufactured by Oxford Gene Technology (OGT). Following washing, the arrays were read using an Agilent Technologies scanner and the Cy5 and Cy3 signals were quantified using Agilent’s Feature Extraction software.

From the data files received from OGT, columns relevant for further analysis were extracted using a bespoke Perl script. This resulted in three files (two for the WT replicates and one for the control) containing the following columns: name of the probe; start of the probe on the *S. coelicolor* genome; end of the probe on the *S. coelicolor* genome; median green (Cy3) signal intensity; median red (Cy5) signal intensity; the probe sequence. Each of the three files obtained above was read into R and the Cy3 and the Cy5 signals were transformed to their log2 values and normalized using the normalizeQuantiles function of the Bioconductor package limma. Log ratios of the Cy3-Cy5 signals were then calculated and normalized to the average signal intensity using the loess function of R. The three sets of log ratio signals obtained were put in a single data frame in R. After removing the design and contrasts matrices appropriate for the data, the limma functions limFit, contrasts.fit and eBayes were applied in sequence. Finally, the function topTable was used to generate a table of probes listing enrichment ratios of the WT relative to the bldC mutant control, ordered by P-values (determined by empirical Bayes moderated t-statistic test) adjusted using the method of Benjamini and Hochberg. Signals around probes with P-values less than or equal to 0.05 were then plotted and inspected manually to determine whether they were part of a signal peak or not.

**Purification of BldC.** Plasmid pH6838, a pET15b (Novagen) derivative expressing his-tagged full-length BldC, was introduced into E. coli BL21(DE3) and grown in LB containing 100 µg/ml of ampicillin, to an OD600 of 0.4. IPTG induction (1 mM) was performed at 30°C in the presence of 0.3 mM IPTG, a microfluidizer was used to disrupt cells. The lysate (in Buffer A: 25 mM Tris HCl pH 7.5, 300 mM NaCl, 5% glycerol) was loaded onto a nickel-nitrilotriacetic acid (Ni-NTA) column. The his-tagged BldC was eluted using increasing concentrations of imidazole in Buffer A. For protein used for crystallization trials, an additional SEC purification step was included. The protein was >95% pure after this step.

**In vitro DNA binding assays.** DNase I footprinting experiments were carried out similar to previous studies44 and according to the method supplied with the Sure Track footprinting kit (GE Healthcare). Specifically, DNA fragments were prepared by PCR using the oligonucleotides pairs 6029_F1 and 6029_R1 and 1415_F2 and 1415_R2 to generate DNA probes. These DNA probes were hybridized to DNA microarrays representing the genome of S. coelicolor and smeA-sfA promoters, respectively. Oligonucleotides were first end-labeled with T4 polynucleotide kinase (GE Healthcare) and [γ-32P]-ATP as described by the manufacturer. Binding reactions were performed in a total volume of 40 µl containing 10 mM Tris HCl pH 7.8, 150 mM NaCl, 2 mM dithiothreitol, 1 µg poly(dI-dC) (Roche), and 10% (v/v) glycerol in the presence of approximately 110,000 cpm of the DNA probe. Following DNase I treatment and purification, products were separated on denaturing 6% (w/v) polyacrylamide gels and visualized using a FLA-7000 phosphorimagier (Fujiﬁlm).

**Hydrosol radical footprinting was performed using a protocol similar to that described previously45.** Speciﬁcally, the binding reactions and probes were identical to those used for the DNase I footprinting experiments, but in a total volume of 25 µl and in the absence of glycerol. After 20 min at RT, Fe(NH₄)₂(SO₄)₂ and EDTA were added to ﬁnal concentrations of 10 µM and 20 µM, respectively, followed by the addition of sodium ascorbate to 1 µM and H₂O₂ to 0.015% (v/v). Reactions were incubated for a further 1–4 min at RT before quenching the reaction with 9 µl stop solution (27.77 mM thiourea, 27.77 mM EDTA, 1 M Na acetate, containing 25 µg yeast RNA). Samples were precipitated with three volumes of ethanol, re-dissolved in stop solution, phenol–chloroform extracted and analyzed in the same way as the DNase I footprints.

**Electrophoretic mobility shift assays (EMSA).** EMSA experiments were carried out using a DNA probe containing the BldC-binding site from the smeA-sfA promoter, radiolabeled on the forward strand. The probe was generated by PCR using the oligonucleotides 1415_F2 and 1415_R2. For EMSA experiments, increasing concentrations of BldC protein was added from 0.004 µM to 1.2 µM. The reaction samples were incubated for 20 min at RT and then run on 5% polyacrylamide gels.

**Crytalization and determination of BldC-DNA structures.** The his-tag was removed from BldC prior to crystallization trials using a thrombin cleavage kit (Qiagen). Crystals of BldC bound to an optimized version of the whiA site were obtained by hanging drop vapor diffusion. The DNA (termed whiA) contained TT and AA overlaps (top strand: 5’TTCATTCGCA-GAACATTTACGGAAGCA-3’). The DNA concentration was 5 µM. The DNA was precipitated by adding 0.1 vol of EtOH to the DNA solution. The precipitated DNA was resuspended in 10 mM Tris HCl pH 8.0, 50 mM NaCl, 2 mM dithiothreitol, 10% (v/v) glycerol in the presence of approximately 110,000 cpm of the DNA probe. Following DNase I treatment and purification, products were separated on denaturing 6% (w/v) polyacrylamide gels and visualized using a FLA-7000 phosphorimagier (Fujiﬁlm).

Protein crystals were grown at 4°C using the hanging drop vapor diffusion method by using drops composed of 1 µl of a protein solution containing up to 10 mg/ml His-BldC, 1 µl of a reservoir solution containing 0.1 M sodium acetate pH 4.5, 10% (v/v) glycerol, and 20% (w/v) PEG 400. The crystals diffracted to 5 Å.
however, the diffusion limit was extended by dragging a crystal to the edge of the
drop with a loop and allowing it to dehydrate before placing it in the cryo-stream. This
saturated diffusion condition beyond 3.5 μl 20% PEG 6000—i.e., crystals take the
hexagonal space group, P6122. SAD X-ray intensity data were collected for a
selenomethionine BldC(L43M-SLM)-whit opt crystal to 3.28 Å resolution at the
selenium peak at the advanced light source (ALS) beamline 8.3.1. The data were
processed in MOSFLM and scaled using SCALA. Autooof was used to determine
heavy atom parameters and perform density modification 46. Due to the high
content (65%) density modification produced an experimental map of excellent
quality, which was used to manually build the model. There are two BldC subunits
and one DNA duplex in the crystallographic asymmetric unit (ASU). Once con-
structed, the model was subjected to multiple rounds of refinement using Phenix
rebuilding using G37 and validation with MolProbity 47, resulting in final Rwork/Rfree
values of 22.3%/26.8% to 3.28 Å resolution (Supplementary Table 3). The analysis
of the DNA structure was carried out using w3DNA 48.

The BldC-whit opt structure revealed that each BldC subunit binds a 9 bp DNA
site arranged as direct repeats. Hence, to obtain a BldC-smea-safs promoter/repressor
complex, WT BldC was mixed with a DNA site that contained the first two BldC
binding sites from the smea-safs promoter anticipating that pseudcontinuous
DNA packing would generate the smea-safs site. Crystals of the complex were
obtained by mixing 0.5 mM BldC (subunit) with 0.5 mM smea-safs DNA and
concentrating the mixture 3-fold using a 30 kDa microconcentrator. The concentra-
tion of DNA was crystallized by mixing it 1:1 with 35% PEG 400, 0.1 M CatCl2, 0.1 M Hepes 7.5. Crystals were produced at RT and grew to maximum size
within a week. The crystals take the tetragonal space group, P4122. X-ray intensity
experiments with WT BldC and BldC mutants. For each FP experiment, WT BldC
polarization (FP) experiments were performed using a PanVera Beacon 2000 FP
instrument. Fluoresceinated oligonucleotides were used for DNA binding
experiments shown in Supplementary Fig. 7 are: 5′-F-TTCAATTCGTACAATTCGTACA-3′.

Fluorescence polarization-based DNA binding experiments. Fluorescence
polarization (FP) experiments were performed using a PanVera Beacon 2000 FP
system. 5′-Fluoresceinated oligonucleotides were used for DNA binding
experiments with WT BldC and BldC mutants. For each FP experiment, WT BldC
or BldC mutant protein were titrated into 0.995 mL of reaction buffer (25 mM Tris,
PH 7.5, 150 mM NaCl) containing 1 μM fluoresceinated oligonucleotide. Fluor-
esceinated oligonucleotides used in the experiments include the whit opt (top
strand: 5′-F-TTCAATTCGTACAATTCGTACA-3′ where 5′-F denotes the 5′
fluorescein label), 36mer optimized site (top strand: 5′-F-ATTCGGA-
CAATTCGGACAAATTCGGACAGAC-3′) and the smea-safs promoter site
(top strand: 5′-F-TAATCCGGATATCCGGATAATCCGGATATCACCT-
TATAGA-3′). The sequences of the three whit opt sites used for FP
experiments shown in Supplementary Fig. 7 are: 5′-F-
TTGCGGGGGCAGGGCGCGGGCGA-3′ (top strand), 5′-F-CTAATTCGGGA-
CAATTAATCCGGACAAATTCGGACAGAC-3′ (top strand), 5′-F-TAATCGGG-
GAATTCGGACAAATTCGGACAGAC-3′ (top strand) and 5′-F-TAATCCGGATATCCGGATAATCCGGATATCACCT-
TATAGA-3′ (top strand). The binding studies on BldC ± the his-tag, selenomethionine BldC
(L43M-SLM), and BldC mutants with the F-whit opt DNA showed that the his-
tag has no significant effect on the L43M-SLM mutants and sele-
nomethionine substitution. FP experiments examining binding of the 36mer and
smea-safs sites were conducted with WT BldC and all data were fit using
Kaleigraph.

Stoichiometry determination experiments. To determine the binding stoichio-
metry of BldC to the whit opt, 36mer opt, and smea-safs site, the buffer and
conditions were identical to those used in the FP binding affinity determination
experiments except that unlabeled DNA was added above the Kd (by using a solution containing 1 mM F-DNA and an amount of unlabeled DNA necessary to achieve a concentration 10-fold above Kd) thereby
ensuring stoichiometric binding. For each oligonucleotide, BldC was titrated into
the binding solution and the graph of the resulting data shows a linear increase in
the observed signal intensity in an excellent fitting. The concentration of the
DNA was increased to 10-fold above Kd (by using a solution containing 1 mM F-DNA and an amount of unlabeled DNA necessary to achieve a concentration 10-fold above Kd) thereby

Size exclusion chromatography. SEC was used to probe the molecular weight of
BldC using a HiLoad 26/600 Superdex 75 prep grade column. Experiments
were performed in a buffer containing 200 mM NaCl, 5% glycerol, 20 mM Tris HCl
pH 7.5 and 1 mM β-mercaptoethanol (BME). For SEC analysis BldC was included in
the buffer at a concentration of 10 μM.

CD spectroscopy. Far-UV CD spectra of WT and mutant BldC proteins were
recorded on an AVIV 435 CD Spectrophotometer in a 1 mm sample cell. Mea-
surements were taken from 200 to 260 nm with a wavelength step of 1.0 nm and a
1 averaged time. Each spectrum is the average of 5 scans. Protein concentrations
ranged from 0.3 to 1.5 mg/ml (the final spectra were normalized for concentration) and the buffer composition was 20 mM NaH2PO4 (pH 7.5), 300 mM NaCl, 5% glycerol, and 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP).

Data availability. ChIP-chip data supporting the findings of the study have
been deposited at the ArrayExpress database under accession number E-GEOD-28908
[https://www.ebi.ac.uk/arrayexpress/experiment/E-GEOD-28908]. Structure factor
amplitudes and coordinates for the BldC-whit opt and BldC-smea-safs
structures have been deposited in the RCSB Protein Data Bank under the accession
codes 6AMK and 6AMA. All other relevant data supporting the findings of the
study are available from the corresponding authors upon reasonable request.

Received: 24 October 2017 Accepted: 22 February 2018
Published online: 19 March 2018

References
1. Flårdh, K. & Buttner, M. J. Streptomyces morphogenetics: dissecting
differentiation in a filamentous bacterium. Nat. Rev. Microbiol. 7, 36–49 (2009).
2. McCormick, J. R. & Flårdh, K. Signals and regulatory mechanisms that govern
Streptomyces development. Fems. Microbiol. Rev. 36, 206–231 (2012).
3. Jakimowicz, D. & van Wezel, G. P. Cell division and DNA segregation
in Streptomyces: how to build a septum in the middle of nowhere? Mol.
Microbiol. 85, 393–404 (2012).
4. Bush, M. J., Tschowri, N., Schlimpert, S., Flårdh, K. & Buttner, M. J. c-di-GMP
signaling and the regulation of developmental transitions in Streptomyces.
Nat. Rev. Microbiol. 13, 749–760 (2015).
5. Chater, K. F. Recent advances in understanding Streptomyces. F1000Res 5,
2795 (2016).
6. Hopwood, D. A. Streptomyces in Nature and Medicine: the Antibiotic Makers
(Oxford University Press, New York, NY, 2012)
7. Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of
antibiotic biosynthesis in Streptomyces. Microbiol. Mol. Biol. Rev. 77, 112–143
(2013).
8. van Wezel, G. P. & McDowall, K. J. The regulation of the secondary
metabolism of Streptomyces: new links and experimental advances. Nat.
Prod. Rep. 28, 1311–1333 (2011).
9. Tschowri, N. et al. Tetrameric c-di-GMP mediates effective transcription
factor dimerization to control Streptomyces development. Cell 158, 1136–1147
(2014).
10. Bush, M. J., Chandra, G., Findlay, K. C. & Buttner, M. J. Multi-layered
inhibition of Streptomyces development: BldO is a dedicated repressor of
whiB. Mol. Microbiol. 104, 700–711 (2017).
11. Flårdh, K. & McCormick, J. R. The Streptomyces O-B one connection: a force
within layered repression of a key developmental decision. Mol. Microbiol.
104, 695–699 (2017).
12. Bush, M. J., Al-Bassam, M. M., Chandra, G., Findlay, K. C. & Buttner, M. J.
BldC delays entry into development to produce a sustained period of
vegetative growth in Streptomyces venezuelae. Preprint at BioRxiv https://doi.org/10.1101/194126 (2018).
13. Schumacher, M. A. et al. The Streptomyces master regulator BldD binds c-di-
GMP sequentially to create a functional BldD2-(c-di-GMP)4 complex. Nucleic
Acids Res. 45, 6923–6933 (2017).
14. Kim, I. K. et al. Crystal structure of the DNA-binding domain of BldD, a
central regulator of aerial mycelium formation in Streptomyces coelicolor A3
(2). Mol. Microbiol. 60, 1179–1193 (2006).
15. den Hengst, C. D. et al. Genes essential for morphological development
and antibiotic production in Streptomyces coelicolor are targets of BldD during
vegetative growth. Mol. Microbiol. 78, 361–379 (2010).

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-03576-3 | www.nature.com/naturecommunications
11
16. Hunt, A. C., Servin-Gonzalez, L., Kedemen, G. H. & Buttner, M. J. The DNA-binding domains of the MerR family of transcription regulators. *J. Bacteriol.* 187, 716–728 (2005).

17. Heldwein, E. E. & Brennan, R. G. Crystal structure of the transcriptional activator BmrA bound to DNA and a drug. *Nature* 409, 378–382 (2001).

18. Brown, N. L., Stoyanov, J. V., Kidd, S. P. & Hobman, J. L. The MerR family of transcription regulators. *FEMS Microbiol. Rev.* 27, 145–163 (2003).

19. Newberry, K. J. & Brennan, R. G. The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N. terminus. *J. Biol. Chem.* 279, 20356–20362 (2004).

20. Changelia, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301, 1383–1387 (2003).

21. Watanabe, S., Kita, A., Kobayashi, K. & Miki, K. Crystal structure of the [Fe2S2] oxidative-stress sensor SoxR bound to DNA. *Proc. Natl Acad. Sci. USA* 105, 4121–4126 (2008).

22. Chang, C. C., Lin, L. Y., Zou, X. W., Huang, C. C. & Chan, N. L. Structural basis of the mercury(II)-mediated conformational switching of the dual function transcription regulator MerR. *Nucleic Acids Res.* 43, 7612–7623 (2015).

23. Samach, H. et al. Structural and dynamics characterization of the MerR family metallocorepressor CueR in its repression and activation states. *Structure* 25, 988–996 (2017).

24. Wang, D. et al. Structural analysis of the Hg(II)-regulatory protein TN501 metallothionein from *Bacillus subtilis*. *Sci. Rep.* 7, 33391 (2016).

25. Ausmees, N. et al. SmexA, a small membrane protein with multiple functions in *Streptomyces sporulation* including targeting of a SpoIIE/FtsK-like protein to cell division septa. *Mol. Microbiol.* 65, 1458–1473 (2007).

26. Salerno, P. et al. One of the two genes encoding nucleoid-associated HU proteins in *Streptomyces coelicolor* is developmentally regulated and specifically involved in spore maturation. *J. Bacteriol.* 191, 649–650 (2009).

27. Komatsu, M. et al. Proteins encoded by the conserved of *Streptomyces coelicolor* A3(2) comprise a membrane-associated heterocomplex that resembles eukaryotic G protein-coupled regulatory system. *Mol. Microbiol.* 62, 1534–1546 (2006).

28. Hempel, A. M. et al. The Sec/Thr kinase AfsK regulates polar growth and anchoring by the sporulation kinetochore-like RacA protein. *Proc. Natl Acad. Sci. USA* 109, E2371–E2379 (2012).

29. Ainsa, J. A., Parry, H. D. & Chater, K. F. A response regulator-like protein that functions at an intermediate stage of sporulation in *Streptomyces coelicolor* A3 (2). *Mol. Microbiol.* 34, 607–619 (1999).

30. Al-Bassam, M. M., Bibb, M. J., Bush, M. J., Chandra, G. & Buttner, M. J. Response regulator heterodimer formation controls a key stage in *Streptomyces* development. *PLoS Genet.* 10, e1004554 (2014).

31. Lamoureux, J. S., Maynes, J. T. & Glover, J. N. Recognition of 5′-YpG-3′ sequences by coupled stacking/hydrogen bonding interactions with amino acid residues. *J. Mol. Biol.* 335, 399–408 (2004).

32. Lamoureux, J. S., Stuart, D., Tsang, R., Wu, C. & Glover, J. N. Structure of the sporulation-specific transcription factor Ndt80 bound to DNA. *Embo. J.* 21, 5721–5732 (2002).

33. Rohs, R. et al. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* 79, 233–269 (2010).

34. Schumacher, M. A., Chinnam, N. B., Cuthbert, B., Tonthat, N. K. & Whitfield, T. Structures of regulatory machinery reveal novel molecular mechanism controlling B. subtilis nitrogen homeostasis. *Genes Dev.* 29, 451–465 (2015).

35. Ben-Yehuda, S., Rudner, D. Z. & Lisick, R. RaC, a bacterial protein that anchors chromosomes to the cell poles. *Science* 299, 532–536 (2003).

36. Schumacher, M. A., Lee, J. & Zeng, W. Molecular insights into DNA binding and anchoring by the Bacillus subtilis sporulation kinetochore-like RaC protein. *Nucleic Acids Res.* 44, 5438–5449 (2016).

37. Abbami, M. A. et al. Structure of the cooperative Xis-DNA complex reveals a micromoleculeprotein filament that regulates phage lambda intasome assembly. *Proc. Natl Acad. Sci. USA* 7, 2109–2114 (2007).

38. Gao, R. & Stock, A. M. Biophysical insights from structures of two-component proteins. *Annu. Rev. Microbiol.* 63, 133–154 (2009).

39. Sidote, D. J., Barbieri, C. M. & Stock, A. M. The structure of the Staphylococcus aureus AgrA LyrTR domain bound to DNA reveals a beta fold and a novel mode of binding. *Structure* 16, 727–735 (2008).

40. He, X., Wang, L. & Wang, S. Structural basis of DNA sequence recognition by the response regulator PhoP in *Mycobacterium tuberculosis*. *Sci. Rep.* 15, 24442 (2016).

41. Park, D. M. & Kiley, P. J. The influence of repressor DNA binding site architecture on transcriptional control. *mbio* 26, e01684014 (2016).

42. Kieser, T., Bibb, M. J., Buttner M. J., Chater, K. F. & Hopwood, D. A. Practical *Streptomyces* genetics (*The John Innes Foundation*, Norwich, 2000).

43. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300 (1995).

44. den Hengst, C. D. et al. Probing direct interactions between CodY and the oppD promoter of *Lactococcus lactis*. *J. Bacteriol.* 187, 512–521 (2005).

45. Hoonen, L. W., Van Werkhoven, A. F., Bijlsma, J. J., Dubnau, D. & Venema, G. The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes Dev.* 12, 1539–1550 (1998).

46. Adams, P. D. et al. PHENIX: a comprehensive Pbython-based system for macromolecular structure solution. *Acta Crystallogr.* D 66, 213–221 (2010).

47. Jones, A. T., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* A 47, 110–119 (1991).

48. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.* D 66, 12–21 (2010).

49. Zheng, G., Lu, X. J. & Olson, W. K. Web 3DNA – a web server for the analysis, reconstruction and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 37, W240–W246 (2009).

50. Bunkoczi, G. et al. Phaser, MRage: automated molecular replacement. *Acta Crystallogr.* D 69, 2276–2286 (2013).

Acknowledgements

We are grateful to the Oxford Gene Technology for expert handling of the ChIP samples, and Ray Dixon, Nick Tucker and Tony Maxwell for helpful discussion. This work was funded by BBSRC grant BB/H006125/1 (to M.J.B.), by the BBSRC Institute Strategic Programme Grant BB/J004561/1 to the John Innes Centre, and by the National Institutes of Health grant GM115547 (to M.A.S.). We would like to acknowledge the Advanced Light Source beamline 8.3.1 for data collection with a special thanks to Jane Tanamachi. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, the National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Beamline 8.3.1 at the ALS is operated by the University of California Office of the President, Multicampus Research Programs, and Initiatives grant MR-15-328599 and the Program for Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation.

Author contributions

M.A.S., C.D.H., R.G.B. and M.J.But designed the research. M.A.S., R.G.B., M.J.Bus and M.J.But wrote the manuscript. C.D.H., M.I.Bus., T.B.K.L. and N.T.T. performed the ChIP-chip, footprinting, EMMA, and operator optimization experiments. G.C. analyzed the raw ChIP-chip data. W.Z. purified proteins for crystallization and performed SEC. B. T. carried out CD analyses. M.A.S. performed the crystallization and structure determination and FB binding assays.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-03576-3.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.