1 **Chromosome organization by a conserved condensin-ParB system**
2 **in the actinobacterium *Corynebacterium glutamicum***
3

4 Kati Böhm[1], Giacomo Giacomelli[1], Andreas Schmidt[3], Axel Imhof[3], Romain Koszul[2],

5 Martial Marbouty[2]*, and Marc Bramkamp[1]*

6

7 [1] Ludwig-Maximilians-Universität München, Fakultät Biologie, Großhaderner Straße

8 2-4, 82152 Planegg-Martinsried, Germany

9 [2] Laboratoire Régulation Spatiale des Génomes, Institut Pasteur, UMR 3525, CNRS,

10 75015 Paris, France.

11 [3] Ludwig-Maximilians-Universität München, Zentrallabor für Proteinanalytik,

12 Biomedizinisches Centrum München, 82152 Planegg-Martinsried, Germany

13 * Corresponding authors:

14 Marc Bramkamp

15 Email: marc.bramkamp@lmu.de; Phone: +49-(0)89-218074611; Fax: +49-(0)89-

16 218074621

17 Martial Marbouty

18 Email: martial.marbouty@pasteur.fr; Phone: +33-144389365

19

Abstract

20

21 Higher-order chromosome folding and segregation is tightly regulated in all domains

22 of life. In bacteria, details on nucleoid organization regulatory mechanisms and function

23 remains poorly characterized, especially in non-model species. Here, we investigate

24 the role of DNA partitioning protein ParB and condensin complexes, two key players

25 in bacterial chromosome structuring, in the actinobacterium *Corynebacterium*

26 *glutamicum*. Chromosome conformation capture reveals SMC-mediated long-range

27 interactions around ten centromere-like *parS* sites clustered at the replication origin

28 (*oriC*). At least one *oriC*-proximal *parS* site is necessary for a reliable chromosome

29 segregation. Using a combination of chromatin immunoprecipitation and

30 photoactivated single molecule localization microscopy evidences the formation of

31 distinct ParB-nucleoprotein subclusters in dependence of *parS* numbers. We further

32 identified and functionally characterized two condensin paralogs. Whereas

33 SMC/ScpAB complexes are loaded via ParB at *parS* sites mediating chromosomal

34 inter-arm contacts like in *Bacillus subtilis*, the MukBEF-like SMC complex MksBEFG

35 does not contribute to chromosomal DNA-folding. Rather, the MksBEFG complex is

36 involved in plasmid maintenance and interacts with the polar *oriC*-tethering factor

37 DivIVA. These data complement current models of ParB-SMC/ScpAB crosstalk, while

38 showing that some condensin complexes evolved functions uncoupled from

39 chromosome folding.

Keywords

40

41 ParB, *parS*, SMC, MksB, condensin, actinobacteria, chromosome conformation

42 capture, chromatin immunoprecipitation, super-resolution microscopy, DNA

43 segregation

44

Introduction

Each organism must complete genome replication and separation in the course of one cell cycle prior to cell division in concert with transcriptional processes. To this end, chromosomes are highly organized structures in terms of segregation and overall folding patterns [1]. The functional organization of bacterial genomes, structured into the nucleoid, has been predominantly investigated in a limited number of model species, e.g. *E. coli*, *V. cholerae*, *B. subtilis* or *C. crescentus*, revealing diverse levels of compaction and segregation strategies[2-4].

ParABS systems and condensins are two (nearly) ubiquitous bacterial enzyme machineries that contribute to chromosome homeostasis. With a few exceptions amongst γ-proteobacteria, all branches of bacteria and several Archaea harbor *parS* sites that recruit partitioning protein ParB [5]. The ParABS system contains one or several *parS* sites usually in the vicinity to the chromosomal origin of replication (*oriC*). ParB proteins bind to these sequence-specific motives and form large nucleocomplexes by spreading and 3D-bridging between ParB dimers [6-9], resulting in large topological domains encompassing the *oriC,* that have been revealed by Hi-C for *B. subtilis* [10]. In an alternative model termed nucleation and caging, ParB-nucleation at *parS* is stabilized by dynamic ParB dimer-dimer interactions and weak interactions with non-specific DNA generating a scaffold for locally high ParB concentrations confined around *parS* [11]. The ParB segregation is driven by a ParA ATPase, which binds nonspecifically to the nucleoid and is released from DNA upon ATP hydrolysis triggered by transient ParB-interactions [12,13]. In the course of chromosome replication ParB-*oriC* complexes act in combination with ParA as Brownian ratchets along dynamic DNA loci: slow ParA-DNA rebinding rates generate ParA-gradients, which serve as tracks for directed movement of partition complexes away from their sisters [14-17]. Perturbation of the system by placing *parS* sites at ectopic, *oriC*-distal regions

3

71   can cause severe DNA segregation phenotypes [18,19]. To date, only few studies

72   investigated the impact of chromosomal *parS* localization on DNA-segregation and

73   folding [18-22].

74   In addition to ParABS systems, most bacteria harbor condensin complexes, members

75   of the structural maintenance of chromosomes (SMC) family of proteins found in all

76   kingdoms of life [23]. In standard model organisms, condensins are equally essential for

77   faithful chromosome segregation by compacting DNA into separate nucleoids [24-26]. The

78   SMC/ScpAB (structural maintenance of chromosomes) complex is well-studied in *B.*

79   *subtilis*, where it consists of two large SMC subunits and the kleisin ScpA associated

80   with dimeric accessory protein ScpB that assemble into a ring-like structure [27]. A recent

81   study suggests progressive extrusion of condensin-encircled DNA loops upon

82   conformational changes in the SMC subunit, which leads to a gradual size increase of

83   trapped DNA molecules [28]. The active process(es) driving DNA extrusion [29,30] allow(s)

84   for translocation along the chromosome with velocities of around 50 Kb/ min [31], and

85   depend(s) on the ATPase activity of SMC [32]. To be loaded on *parS* sites, SMC/ScpAB

86   complexes necessitate ParB [20,22,33,34]. They redistribute to distant chromosomal

87   regions, promoting topological changes, and notably the co-alignment of right and left

88   replichores [10,21,22,31,35]. In sharp contrast with SMC/ScpAB, the *E. coli* condensin

89   MukBEF does not promote the co-alignment of chromosomal arms [36,37], but facilitate

90   *cis* structuration by establishing long-range contacts between loci belonging to the

91   same replichores from stochastic chromosomal loci (except for the *ter* region

92   containing the replication terminus) [37,38]. Despite the importance of condensins in

93   chromosome organization, the role of SMC homologs besides the model species *B.*

94   *subtilis*, *C. crescentus* and *E. coli* remain largely unexplored. These species all contain

95   a single condensin complex, yet a broad range of bacteria possesses combinations of

96   SMC/ScpAB and MksBEFG (MukB-like SMC), for which functional characterizations

97   are non-existent to date [39]. Current work in bacteria as well as knowledge from

98   eukaryotic studies convey the general assumption that all SMCs are likely to play

99   role(s) in chromosome organization. In bacteria it is unknown, why some species

100   harbor more than one type of condensin, and whether and how they would work in

101   concert with each other and coordinate with systems such as ParABS.

102   In this work, we used a combination of high-resolution microscopy and genomic

103   chromosome conformation capture (3C/Hi-C) [35] to unveil the global organization of the

104   diploid *C. glutamicum* genome. *C. glutamicum* is a polar growing actinobacterium,

105   whose genome encodes both SMC/ScpAB and MksBEFG. In this species, the two

106   *oriC*s are continuously associated with the polar scaffold protein DivIVA, while newly

107   replicated sister *oriC*s segregate towards division septa via the ParABS system [40-42].

108   In contrast to *B. subtilis*, *C. glutamicum* ParAB are by themselves crucially important

109   drivers of reliable nucleoid separation prior to cell division, where ParAB deletions yield

110   in 20 % of anucleoid cells [43-45]. Here, analyses of chromosomal ParB-binding patterns

111   evince ten redundant *parS* sites, which mediate ParB subcluster formation at *oriC*. A

112   single *parS* site maintains ParB propagation over 32 Kb neighboring regions, and is

113   sufficient to promote the SMC-dependent alignment of the two chromosomal arms. Hi-

114   C also reveal SMC-dependent long-range contacts surrounding *oriC*. On the contrary,

115   we showed that the polarly positioned MksBEFG condensin acts exclusively on

116   plasmid-transmission to daughter cells, without influencing nucleoid architecture.

117

118 **Results**

119 **Chromosome segregation is governed by a cluster of ten *oriC*-proximal *parS***
120 **sites**

121 Previous studies on *C. glutamicum* chromosome partitioning have revealed two stable

122 ParB-*oriC* clusters at each cell pole, while newly replicated origins are segregated

123 towards a division septa formed at midcell [40]. In *B. subtilis*, *C. crescentus* and *P.*

124 *aeruginosa* ParAB-mediated chromosome segregation and folding depends on *parS*

125 sites [18,19,21]. In *C. glutamicum*, *parS* positions have not been characterized yet. Initially,

126 four to eight *parS* sites were predicted in *Corynebacterineae* [5]. However, we identified

127 ten *B. subtilis*-like 16 bp consensus sequences in *C. glutamicum* localized in one

128 cluster within a 35 Kb region distant 73 Kb from *oriC* using blast (1% of the 3.21 Mb

129 chromosome; Fig. 1A). Out of the ten *parS* sites, only the one located furthermost from

130 *oriC* (*parS1*) lies within a coding sequence (*trpCF*). All other *parS* sequences (labelled

131 *parS 2-10*) are located in intergenic regions. Degenerated *parS* sequences exhibiting

132 at least three base-pair mismatches were also identified further away from *oriC*, e.g. 5'

133 of *cg0146* or within the *fusA* and *cg1994* coding region. To test whether these putative

134 *parS* were responsible for the recruitment of ParB, α-mCherry-ChIP analyses were

135 performed with a strain harboring a mCherry-tagged version of the native ParB (note

136 that all mutant strains used in this study derive from clean allelic replacements and

137 have, unless otherwise noted, a wild type-like phenotype). Distinct and very

138 reproducible enrichment signals were obtained at ten *parS* sites close to *oriC* (*parS*1-

139 10 at 3.16 MB) (Fig. 1A), whereas the imperfect *parS* sequences identified with blast

140 clearly failed to recruit ParB. Additional smaller peaks were identified at highly

141 transcribed DNA regions, in particular at ribosomal genes, tRNA gene clusters and at

142 all of the rRNA operons (Fig. 1A). Magnification of the *oriC* region reveals three distinct

143 ParB propagation zones overlapping with *parS*1-4, *parS*5-8 and *parS*9-10, respectively

144   (Fig. 1B). Remarkably, those three regions seem to recruit decreasing amounts of

145   ParB, from *parS1-4* (most enriched) to *parS9-10* (less enriched). Since all *parS* are

146   identical in sequence, differences in ParB-recruitment might result from the number

147   and distance of *parS* sequences in the context of the overall nucleoid folding patterns

148   at the *oriC*-region.

149   **Higher-order organization of the *C. glutamicum* chromosome**

150   In *B. subtilis*, SMC-mediated chromosome folding initiates at ParB-*parS* clusters

151   surrounding the origin of replication, bridging the two replichores with each other [10,21].

152   To characterize whether *C. glutamicum parS* sites play a similar role in the overall

153   organization of the chromosome, we applied a Hi-C like approach [10,46] to exponentially

154   growing wild type cells (Material and Methods). The genome-wide contact map,

155   displaying the average contact frequencies between all 5 Kb segments of wild type

156   chromosomes (Fig. 1C) displayed the following 3D features. First, a strong and broad

157   diagonal reflecting frequent local contacts between adjacent loci and observed in all

158   Hi-C experiments. Second, chromatin interaction domains (CIDs), i.e. regions making

159   increased contact frequencies within themselves and previously described in *C.*

160   *crescentus* and other species [10,21,35,37,47], (Fig. 1C, S1) (11 domains detected at a 200

161   Kb resolution). Third, a secondary diagonal perpendicular to the main one and

162   extending from the 35 Kb *parS* cluster (Fig. 1D – white dashed line) from the *ori* region

163   to the terminus. This structure shows that the two replichores are bridged over their

164   entire length, similarly to *B. subtilis* and, to some extent, *C. crescentus* [10,21,35].

165   Interestingly, this secondary diagonal also displays discrete long-range contact

166   enrichments (Fig 1C), which may reflect bridging of the two chromosomal arms at

167   specific locations. Finally, the contact map also displays a faint, cross-shaped signal

168   corresponding to contacts between the *ori* region and the rest of the chromosome (Fig.

7

169 1C – dark triangle on the sides of the contact matrix), a feature never described before.

170 These contacts might represent a replication signal reflecting the translocation of the

171 ParB-*oriC* complex along the nucleoid during segregation when *oriCs* reposition at

172 midcell. This signal is also maximal at the *parS* cluster and not at *oriC* locus. An

173 observation that reinforce the fact that the *parS* cluster is at the tip of *Corynebacterium*

174 chromosome fold and is one of the main actor of chromosome segregation.

A single *parS* site is sufficient to maintain a wild type ParB binding region and
175
176 **chromosome architecture**

177 Since all *parS* sites are in close proximity on the *C. glutamicum* chromosome, we

178 tested the importance of ParB-*parS* complex titration for the overall chromosome

179 organization. Cells with chromosomes carrying a single *parS* site grow and divide like

180 wild type cells (Fig. 2A, S2). However, the removal of all 10 *parS* sites resulted in a cell

181 length phenotype (Fig. S2) and 29% DNA-free mini-cells hinting to a nucleoid

182 segregation defect similar to the Δ*parB* phenotype (Fig. 2A, Tab. S1). We further

183 analyzed ParB localization in mutant strains carrying either one or none *parS* sites.

184 Firstly, cellular localization of fluorescent ParB-eYFP foci is similar to wild type,

185 positioning at cell poles and migrating to the newly formed septa (Fig. 2B, S3) [40].

186 Interestingly, the combination of a single *parS* site with ParB-eYFP resulted in 7%

187 anucleoid mini-cells (Fig. S2, Tab. S1), reflecting functional constraints of the ParB-

188 eYFP fusion in presence of only one *parS* site. Therefore, the high number of

189 chromosomal *parS* sites likely evolved to improve the robustness of the segregation

190 machinery. ParB ChIP-qPCR signals of locus *parS1* were similar in both wild type and

191 mutant strains (Fig. 2C). ParB spreading around the single *parS* site was characterized

192 through ChIP-seq analysis (Fig. 2D, S4), where ParB binding was maximum within 2

193 Kb windows on both sides of *parS*, while extending up to 16 Kb on either side. If each

194 *parS* site promotes ParB deposition in comparable ranges to *parS*1, abrupt enrichment

8

195   drops downstream of *parS4* and *parS8* would have been absent in wild type,

196   suggesting that DNA properties accountable for the regulation of ParB deposition are

197   independent of *parS* distributions along the nucleation zone. We next investigated the

198   role of *parS* sites and ParB in the overall chromosome folding by performing Hi-C in

199   mutants (Fig. 2E, 2F). The absence of ParB or all *parS* sites led to the disappearance

200   of the secondary diagonal. In addition, the cross-shaped pattern resulting from

201   contacts between the ori and the whole chromosome disappears in those mutants,

202   also illustrated by the ratio between wild type and mutant contact maps (Fig. 2F). This

203   result shows that *parS* sites and ParB are two major structural components of

204   chromosome organization and act to recruit downstream factors that fold the

205   chromosome emanating from the *parS* cluster, and bridge the two chromosomal arms

206   together down to the replication terminus region. The contact map of the strain deleted

207   for *parS2-10* but carrying *parS1* maintains a secondary diagonal, showing that a single

208   *parS* site is sufficient to ensure the loading of ParB and the overall folding of the

209   chromosome (Fig. 2E, 2F). However, some differences appeared between wild type

210   contact matrix and the one resulting from the strain harboring only one *parS* site. In

211   this mutant, the large domain surrounding *oriC* appears more defined than in wild type

212   suggesting that one *parS* site is not sufficient to fully restore *Corynebacterium*

213   chromosome folding possibly due to a slower replication and consequently a less

214   dynamic folding (Fig. 2E, 2F). The single *parS* site was repositioned at different

215   genomic regions. Cells harboring an ectopic *parS* site at 9.5°, 90°, 180° or 270°

216   positions were viable (Fig. S2, S3A, B). Unlike in cells harboring *parS*1 at its original

217   position, ParB-*parS* complexes distribute randomly along the longitudinal cell axis in

218   all of these mutants (Fig. S3B), leading to around 25% anucleate cells (Tab. S1).

219   Therefore, none of these *parS*-shifts restores controlled nucleoid segregation. The

220   number of ParB foci nevertheless correlates well with cell length (Fig. S3C), excluding

221   replication initiation deficiencies. ParB-binding to a *parS* sequence positioned at the

222   90° chromosomal position (locus *cg0904*, strain CBK037) was identified in a 9 Kb

223   range on either side of *parS* (Fig. S3D, S4), approximately half the ParB-propagation

224   distance determined for cells harboring one *parS* at its native locus. We also analyzed

225   mutant CBK037 (*parS* at 90° chromosomal position) using Hi-C (Fig. S5). The resulting

226   contact map presents a "bow shape" motif at the position of the aberrant *parS*

227   sequence demonstrating a recruitment of the ParB protein at this location as well as a

228   local folding of the chromosome. However, this loading appears insufficient to fold the

229   whole chromosome and to anchor the two chromosome arms over their entire length.

230   DNA topology, origin replication and overall chromosomal localization might hereby

231   determine *parS*-distant ParB-DNA interaction. Collectively, these results show a

232   redundancy of *parS* sites, yet their function is restricted to a confined *oriC*-proximal

233   region.

234   **ParB subclusters identified by PALM reflect protein enrichment at *parS* sites**

235   To directly characterize *oriC* domain compaction via ParB, we applied photoactivated

236   localization microscopy (PALM) to visualize individual ParB-PAmCherry molecules

237   with nanometer resolution. PALM revealed distinct ParB-dense regions at cell poles

238   and quarter positions regions, similar to foci observed via diffraction limited

239   epifluorescence microscopy (Fig. 2G). These ParB-enriched regions (macro-clusters)

240   display heterogeneous densities, with a variable number of higher density zones within

241   sub-clusters. Macro- and sub-clusters have been identified via the Optics algorithm[48,49]

242   (see Material and Methods) and analyzed in strains harboring a single, two or all the

243   *parS* sites (Fig. 2G, S6). We define a macrocluster as 32 events being localized within

244   a maximum distance of 50 nm for macroclusters and 35 nm for subclusters. Note that

245   high chromosome numbers promote inter-molecular *oriC*-colocalization in fast-growing

246 cells. For more accurate cluster estimations, PALM analysis was performed using

247 slow-growing cells resulting in significantly fewer ParB macro-clusters per cell (Fig.

248 S6B) [40]. Since segregation of *oriC* complexes might alter their DNA compaction, we

249 focused on the two largest macroclusters per cell, stably tethered at cell poles. The

250 amount of ParB contained within each macro-cluster in wild type is significantly higher

251 than in cells containing a single *parS* site (Fig. 2H), in agreement with the ParB

252 deposition observed via Chip-seq. A parallel between PALM and Chip-seq can also be

253 drawn with respect to the number of sub-clusters per macro-cluster, with a higher

254 number of sub-clusters in the wild type compared to the single *parS* site (Fig. 2H).

255 These differences were not observed when comparing cells harboring all, or two *parS*

256 sites (*parS*1,10) (Fig. S3, S4, S6). These observations could explain the differences

257 observed between contact matrices of wild type and Δ*parS* 2-10 strains and the higher

258 structuring of the *oriC* domain when only one *parS* site is present. We therefore

259 conclude that the architecture of *C. glutamicum* partition complex is dependent on *parS*

260 and ParB-*parS* nucleoprotein-complexes are visible as individual subclusters.

261 ### *C. glutamicum* harbors two paralogs of condensin complexes

262 In bacteria, the condensin paralog complexes SMC/ScpAB and, in *E. coli* and other

263 enterobacteria, MukBEF, are key players of chromosome folding [10,21,35,37]. MksBEF (for

264 MukBEF-like SMC) is another condensin occasionally found in bacteria genomes [39],

265 whose role(s) remain(s) obscure. A sequence homology search of the *C. glutamicum*

266 genome pointed at the presence of both SMC/ScpAB and MksBEF. The SMC/kleisin

267 is encoded by genes *cg2265* (*smc*), *cg1611* (*scpA*) and *cg1614* (*scpB*) (Fig. 3A), while

268 the Mks complex is encoded on a widely conserved operon[39] and comprises genes

269 *cg3103-cg3106* (*mksGBEF*) (Fig. 3A), including MksG which has being suggested to

270 act in complex with MksBEF [39].

271   To characterize condensin complex formation *in vivo*, mass spectrometry of pulldown

272   experiments using SMC and MksB as baits of whole cell lysates were performed.

273   Stability of SMC and MksB fluorescent fusions were confirmed by western blotting (Fig.

274   S7). Kleisin subunit ScpA and ScpB co-precipitated significantly with SMC compared

275   to the negative control, while subunits MksF and MksE, but not MksG, were

276   substantially enriched in the MksB pulldown experiments (Fig. S7). ParB, which

277   mediates SMC-loading onto DNA in *B. subtilis* and *S. pneumonia* [20,33,34], was not

278   immuno-precipitated with SMC in any of the experiments. Bacterial two-hybrid

279   analyses confirmed mass spectrometry results, pointing at the formation of

280   SMC/ScpAB and MksBEF complex (Fig. 3B, C). No significant interactions between

281   SMC/ScpAB and ParB were detected, and we observed ScpA-ScpA self-interaction

282   signals well above background. Moreover, MksG connects to the MksBEF complex via

283   interaction with MksF, while MksF and MksG subunits further interact with the *C.*

284   *glutamicum* polar scaffold protein DivIVA.

285   **SMC-mediated cohesion of chromosomal arms**

286   We aimed to characterize *C. glutamicum* condensin SMC/ScpAB. Mutation of the

287   SMC/ScpAB complex causes a conditionally lethal phenotype due to chromosome

288   mis-segregation in *B. subtilis* [25]. In contrast, a *smc* deletion in *C. glutamicum* did not

289   result in growth defects, DNA-segregation defects or aberrant cell length distributions

290   and morphologies compared to the wild type in minimal or complex media (Fig. S8,

291   Tab. S1). Nonetheless, the combination of genetic backgrounds *parB::parB-eYFP* and

292   Δ*smc* yield a minor fraction of anucleate cells (4-5 %) (Tab. S1), indicating that SMC

293   and ParB function in the same pathway and have a synthetic phenotype. Hence, a

294   functional interaction of SMC and ParB proteins regulating chromosome organization

295   is likely. In order to further determine cellular localization of SMC/ScpAB complexes, a

12

296   strain harboring a fluorescently tagged version of core subunit SMC was imaged,

297   revealing the formation of SMC clusters along the entire longitudinal axis of the cell

298   (Fig. 3D). Clusters of SMC and ParB investigated in a strain carrying both labelled

299   complexes (*parB::parB-mNeonGreen smc::smc-mCherry*) are often proximal but do

300   not necessarily co-localize, while the foci numbers correlate with cell length (Fig. 3E).

301   Up to eight SMC-mCherry foci were counted per cell. On average, cells contained

302   fewer SMC-foci than ParB nucleoprotein complexes (Fig. S7). To further characterize

303   the role of SMC, we generated Hi-C contact maps of the mutant (Fig. 3F). Deletion of

304   *smc* abolishes the secondary diagonal in the maps (Fig. 3F). The combination of *smc*

305   and *parB* mutations mimics a *parB* phenotype, again resulting in the loss of contacts

306   between chromosomal arms and further in the loss of the segregation signal described

307   before (Fig. 3F). Therefore, an interplay of SMC/ScpAB with ParB is responsible for

308   replichore cohesion in *C. glutamicum*, similar to *B. subtilis* and *C. crescentus* each

309   harboring only one condensin complex [10,21,22,35]. Thus, ParB acts epistatic to SMC.

310   **ParB-dependent SMC-recruitment to chromosomal loading sites**

311   Since cellular SMC-mCherry signal hinted to distinct agglomeration clusters along the

312   *C. glutamicum* chromosome, we investigated its putative binding sites via ChIP-seq. A

313   small enrichment in SMC deposition was detected at and around the *parS*1-10 cluster

314   (Fig. 4A) that disappears upon *parB* or *parS* deletion (Fig. 4A, S4, S9). In addition,

315   comparably minor enrichment signals are present throughout the chromosome, which

316   partially coincide with genomic loci of high transcriptional activity. Distinct SMC-

317   mCherry foci are less frequent in the absence of ParB or *parS*. (Fig. S9). These findings

318   suggest that ParB promote condensin loading onto DNA at *oriC*-proximal *parS* sites.

319   In addition, ChIP-seq revealed that SMC concentrates at a 13 Kb region upstream

320   *parS*1 (Fig. 4A). SMC enrichment in this region was lost following a partial deletion of

321   this locus and its -reinsertion at another genomic position or following its substitution

322   by a random DNA sequence (Fig. S9). Therefore, the accumulation of SMC at the 13

323   Kb region in the vicinity of *parS* sites points at roadblocks that trap SMC, rather than

324   specific SMC-binding. This hypothesis is further supported by the study of the contact

325   map of wild type cells (Fig. 1C, 1D, S1, S10). Indeed, the SMC enrichment region is

326   clearly delimited by a strong border on its left (Fig. S1 – Directional Index at 100 Kb

327   resolution and Fig. S10 – red dashed line). In the absence of ParB or SMC (Fig. S10),

328   the strong border observed in HiC maps is shifted towards *parS* sites. Therefore, this

329   border originates from a combination of multiple processes.

330   SMC is also recruited to *parS* inserted in ectopic positions, e.g. the 90° *parS*-insertion

331   (Fig S4). Indeed, in the absence of SMC (Fig. S5), the bow shape motif is no longer

332   present at the ectopic *parS* site demonstrating that chromosomal arm cohesion is

333   SMC-dependent and that artificial loading of SMC at non-native positions is not

334   sufficient to fold the entire chromosome. We further assayed chromosomal SMC-

335   loading sites by making use of a well-characterized SMC ATP-hydrolysis mutant

336   E1084Q [32,50-52]. SMC^{E1084Q} mutant strongly accumulates at *parS* sites in *C.*

337   *glutamicum*, mimicking a ParB-enrichment pattern (Fig. 4A). Decreased ChIP-

338   enrichment signals throughout the rest of the chromosome hint to an impaired SMC-

339   migration along DNA (Fig. 4B). Conclusively, we confirm specific SMC-loading by ParB

340   to an *oriC*-proximal region on the *C. glutamicum* chromosome.

341   Interestingly, ChIP-analysis of a *C. glutamicum* ParB^{R175A} mutation, which leads to a

342   loss of dimer-dimer interactions in the corresponding *B. subtilis* ParB^{R79A} mutation [8],

343   results in increased SMC-binding at ParB^{R175A} propagation zones (Fig. 4A). Changes

344   in *in vitro* dsDNA-binding affinities compared to wild type ParB could not be verified

345   (Fig. S3), neither enhanced binding affinity for SMC/ScpAB by bacterial two-hybrid

14

346   analyses (Fig. 3B). The mutation results in large fractions of DNA-free cells and growth

347   rates and ParB^{R175A} cluster formation are particularly affected in cells harboring a single

348   *parS* site (Fig. S2, Fig. S3). ChIP-data indicate broadened and less distinct enrichment

349   signals compared to wild type ParB in presence of all or one *parS* sites (Fig. 4A, S3,

350   S4). Therefore, ParB^{R175A} is still capable of building up weak nucleoprotein complexes

351   around *parS* sites. Hi-C data of the corresponding mutant show the same tendency

352   with a conservation of the overall chromosome architecture with the presence of a

353   secondary diagonal and the conservation of the origin domain folding (Fig. 4C, S10).

354   However, the signal ranging from the secondary diagonal is weak compared to the wild

355   type one as shown by the ratio matrix (Fig. 4C). Consequently, SMC-translocation

356   along DNA appears only partially impaired in this mutant (Fig. 4A, S4). The ParB^{R175A}

357   mutation either locks the translocation ability of SMC/ScpAB by a direct interaction or

358   alterations of ParB^{R175A} nucleoprotein complex properties lead to SMC-trapping along

359   DNA-loops at *parS*. Altogether, these analyses confirm that the *C. glutamicum*

360   SMC/ScpAB complex is a *Bacillus*-like condensin that loads and redistributes to distant

361   chromosomal regions via an explicit ParB-crosstalk at *parS*.

MksB impacts on plasmid maintenance in *C. glutamicum*

363   To test whether both *C. glutamicum* condensins SMC and MksB are redundant in

364   function, we generated mutants lacking the condensin core subunit Δ*mksB* or both

365   Δ*smc* Δ*mksB*. Similar to Δ*smc*, no growth and morphology phenotypes could be

366   detected for both mutants (Fig. S8, Tab. S1). A triple mutation Δ*parB* Δ*smc* Δ*mksB* did

367   not attenuate the Δ*parB* phenotype, excluding redundancy of condensin functions in

368   chromosome segregation (Fig. S8). Further, *oriC*-ParB foci numbers (Fig. 3D) as well

369   as their spatiotemporal localization (Fig. S8, time-lapse microscopy not shown) remain

370   largely unaffected upon deletion of *smc* and *mksB*. MksB fluorescence was mainly

15

371   detected at the cell poles (Fig. 3D), further supporting an interaction with the polar

372   protein DivIVA. Moreover, we applied Hi-C to characterize the role of MksB in genome

373   folding in the different mutants (Fig. 3F). In contrast to *smc*, deletion of *mksB* had no

374   effect on chromosome organization, as shown by the nearly white ratio map between

375   the wild type and the mutant (Fig. 3F). Moreover, Δ*smc* and Δ*smc*Δ*mksB* contact maps

376   were nearly identical (Fig. 3F), showing that MksB and SMC are most likely not

377   involved in the same process(es). ChIP-seq of MksB failed to detect specific loading

378   sites along the *C. glutamicum* chromosome (Fig. S11), supporting the hypothesis that

379   MksB, unlike other bacterial condensins studied so far, plays no direct or indirect role

380   in *C. glutamicum* chromosomes organization. Therefore, we analyzed its impact on the

381   maintenance of extrachromosomal DNA. The MksBEFG complex appears involved in

382   plasmid maintenance, as shown by the qPCR copy number analysis of two low-copy

383   number (pBHK18 and pWK0) and two high-copy number (pJC1 and pEK0) plasmids.

384   In Δ*mksB* mutants both low-copy number plasmids were enriched up to ten-fold

385   compared to wild type, when grown in the absence of selection marker (Fig. S11). On

386   the contrary, the amount of high copy number vectors per cell was hardly affected. We

387   confirmed these findings by plasmid extractions from *C. glutamicum* cells lacking MksB

388   that yielded exceptionally large quantities of pBHK18 and pWK0, turning them into high

389   copy number plasmids under these conditions (Fig. S11). By contrast, amounts of

390   pJC1 and pEK0 did not differ notably compared to control strains. These analyses

391   show a MksB-dependent decrease in plasmid level, specifically of low copy number

392   plasmids.

393   Together, our data show that the two condensins in *C. glutamicum* evolved very

394   different functions: whereas SMC/ScpAB act with ParB to promote replichore pairing

395   and   origin   domain   organization,   MksBEFG   does   not   organize   chromosome

16

396   architecture and seems involved in plasmid maintenance through a mechanism that

397   remains to be characterized.

398

Discussion

399   Condensins are widely conserved enzyme machineries, which have been implicated

400   in chromosome organization of pro- and eukaryotes [53]. For long it was considered that

401   bacterial genomes encode one condensin complex that would either be of the

402   Smc/ScpAB type as found in *B. subtilis* and *C. crescentus* or the MukBEF complex

403   encoded in *E. coli* and related proteobacteria [23]. However, recent reports suggested

404   the existence of two or even multiple condensin systems in a single species [39]. A

405   prominent example is *P. aeruginosa*, where the MksBEF complex was suggested to

406   act in chromosome organization due to a synthetic DNA segregation phenotype in

407   combination with SMC/ScpAB [39]. Yet, the underlying mechanisms and the precise

408   function of these two condesnin systems remained largely untested. We report here

409   that the Gram positive actinobacterium *C. glutamicum* also contains SMC/ScpAB and

410   the Muk-like MksBEFG complexes. We set out to address the individual functions of

411   the two condensin systems. Surprisingly our data provide clear evidence that the class

412   of MksBEFG proteins do not work as chromosomal interactors, thus the function of

413   bacterial condensins in promoting DNA segregation to daughter cells is not generally

414   conserved. A recent bioinformatics study predicted a role for MksBEFG complexes

415   (termed Wadjet system) in plasmid-related defense, where heterologous complex

416   expression conveyed protection against the uptake of a high copy number plasmid [54].

417   However, function of MksBEFG in its native host had not been addressed before. We

418   could show that the Mks system is indeed involved in the control of plasmid copy

419   numbers and that there is no involvement of this system in chromosome organization.

420   This is a fascinating similarity to specific eukaryotic condensins such as Rad50, being

421   the closest eukaryotic relative to MukB/MksB [23,39]. It was recently shown that Rad50-

422   CARD9 complexes sense foreign cytoplasmic DNA in mammalian cells. This includes

423   antiviral effects of Rad50 that are important for innate immune responses. Thus, it

18

425   takes no wonder that several DNA-viruses have evolved strategies to inhibit Rad50

426   signaling [55]. Also the more distantly related eukaryotic SMC5/6 complex had been

427   shown to act in a defense mechanism against circular hepatitis B virus DNA,

428   resembling the specific effect of prokaryotic MksBEFG on plasmids [56]. Together, our

429   data lend support to the notion that condensins function in innate immunity is an

430   ancient mechanism and condensins complexes that diverged early in evolution have

431   specialized functions beyond chromosome organization. However notably, we provide

432   evidence that the MksBEFG complex is the only known condensin amongst pro- and

433   eukaryotes that exclusively targets non-chromosomal DNA. For MksBEF systems it

434   has been proposed that a fourth subunit, MksG is important for function in plasmid

435   maintenance [54]. We could verify that MksG is part of the MksBEF complex of *C.*

436   *glutamicum*. The direct interaction of a Mks complex with a polar scaffold protein like

437   the *C. glutamicum* DivIVA has not been described before, but may be very well in

438   accord of its function in plasmid control. Polar plasmid localization has been described

439   in other systems before [57]. A challenging question for the future will be to determine

440   how the MksBEFG system can identify foreign DNA and how it is loaded onto the

441   plasmids. In vitro analysis of the related MukB has revealed that the protein

442   preferentially binds to single stranded DNA, although also autonomous loading on

443   double stranded DNA was observed [58]. Thus, one might speculate that MksB is loaded

444   onto plasmid DNA once replication is initiated.

445   We further describe here that the second condensin SMC/ScpAB is indeed the major

446   player of replichore cohesion and chromosome organization in *C. glutamicum*. Like in

447   *B. subtilis*, SMC is loaded onto the chromosome by a ParB/*parS* loading complex

448   before spreading to the entire chromosome. The mild DNA-partitioning defects of a

449   *smc* deletion in combination with a ParB-eYFP modification (Tab. S1) strongly suggest

450   a supportive role of SMC/ScpAB in the process of nucleoid separation, yet the *smc*

451   phenotype appears to be entirely compensated by ParB. Therefore, our data

452   demonstrate that the conserved role for SMC in chromosome organization [10,20-22,34,35]

453   is also maintained in presence of a second condensin complex. Moreover, bacterial

454   two-hybrid analyses of SMC/ScpAB subunits evidence a self-interaction of *C.*

455   *glutamicum* kleisin ScpA (Fig. 3), that has not been described in other organisms

456   before. Based on this result, we speculate that *C. glutamicum* SMC/ScpAB might form

457   dimers via kleisin subunits similar to *E. coli* MukBEF complex [59,60]. These data point to

458   a handcuffing model, where two SMC/ScpAB complexes are physically coupled

459   together and translocate in pairs along the chromosome, similar as suggested for *B.*

460   *subtilis* [31]. We further describe a new phenotype for a ParB^{R175A} point mutation in *C.*

461   *glutamicum*, that blocks SMC-release from its loading site. Building on this, we observe

462   a weak interaction signal of ParB^{R175A} with ScpB in bacterial two-hybrid analyses.

463   Alternatively, SMC/ScpAB remains indirectly entrapped in higher-order ParB^{R175A}-

464   nucleocomplexes, which possess altered DNA-folding properties. In either case, this

465   mutant underlines the crosstalk between SMC/ScpAB and ParB nucleoprotein

466   complexes in bacterial nucleoid organization.

467   Analysis of ParB complexes using 2D PALM reveals ParB-dense regions within

468   clusters that correlate to the number of ParB enrichment zones along adjacent *parS*

469   sites. In line with a current study on ParB cluster-assembly in *V. cholerae* [61], we

470   suggest that these subclusters derive from independent nucleation and caging events,

471   which merge into one ParB-macrocomplex per *oriC* in *C. glutamicum*. Presence of a

472   single *parS* site leads to formation of almost globular ParB densities. Using Hi-C

473   approaches, we further show that *parS* sites as well as ParB are major players of

474   chromosome folding in *C. glutamicum* as previously shown in other models [10,21,22,35].

475   *C. glutamicum* chromosome adopts a global folding with a strong cohesion between

476   the two chromosomal arms as expected from a bacterium harboring a longitudinal

477   chromosomal organization similar to *B. subtilis* and to a lesser extent *C. crescentus*

478   (Fig. 1). Our analysis also suggests the existence of a chromosomal domain at *parS*

479   sites in *C. glutamicum* as previously observed in *B. subtilis*, but with important

480   differences: *parS* sites in *C. glutamicum* are only found on one side of the *oriC* locus

481   and appeared to be at the edge of the nucleoid structure as observed in *C. crescentus*.

482   A hairpin structure as it was observed in *B. subtilis* is absent in *C. glutamicum*[10].

483   Contact maps of a strain with an ectopic *parS* site feature a bow-shaped structure

484   reflecting an asymmetry in arm interaction, which has been shown before in *B. subtilis*

485   and *C. crescentus* [21,22]. Zipping of the chromosome is not complete and the ectopic

486   *parS* site does not reorient the entire chromosome. Therefore, additional factors are

487   involved in chromosome localization that supplement polar ParB-*parS* binding to

488   DivIVA.

489   Importantly, we describe ParB/*parS*-dependent DNA segregation signals showing that

490   the chromosomal *ori-ter* configuration allows for ParAB-driven *oriC*-segregation trailing

491   along the nucleoid. Based on our data we propose the following model shown in Figure

492   5: Organisms with polarly localized *oriC*s and a longitudinal chromosome organization

493   rely on ParAB for *oriC* segregation, since they can use the DNA-scaffold as a track. By

494   contrast, species with a central replication factory cannot efficiently use ParAB. *B.*

495   *subtilis* is an exception, since here a longitudinal chromosome orientation is present

496   during sporulation and hence, *parAB* (*spo0J*/*soj*) phenotypes are only obvious during

497   spore formation. Consequently, SMC/ScpAB-mediated replichore cohesion is

498   generally dispensable for *oriC* segregation in bacteria with a strict longitudinal

499   chromosome arrangement that allows for efficient ParABS-driven chromosome

500   partitioning.

Acknowledgements and Funding

Authors contribution

Conceptualization: Kati Böhm, Martial Marbouty, Romain Koszul, Marc Bramkamp

Methodology:

- Strain constructions and analyses, epifluorescence microscopy, ChIP, bacterial two-hybrid screening, *in vitro* protein assays: Kati Böhm
- Chromosome Conformation Capture: Martial Marbouty
- PALM: Giacomo Giacomelli
- Mass spectrometry: Andreas Schmidt, Axel Imhof

Writing manuscript: Kati Böhm, Martial Marbouty, Andreas Schmidt (Material and Methods), Giacomo Giacomelli (Material and Methods), Romain Koszul, Marc Bramkamp

Materials and Methods

Bacterial strains, plasmids and oligonucleotides

Primers, plasmids and strains used in this study are listed in Tables S2 and S3. Detailed information on strain construction and growth conditions are provided in the **Supplementary Information**.

Plasmid extraction from *C. glutamicum* cells

C. glutamicum cells were grown in 10 ml BHI medium to exponential growth phases in presence of selection antibiotic, following incubation with 20 mg/ml lysozyme in P1 buffer (NucleoSpin® Plasmid Kit, Macherey-Nagel) overnight at 30 °C. Subsequently, plasmids were extracted by using the plasmid kit according to manufacturer's instruction.

Protein identification via immunoprecipitation and mass spectrometry

Immunoprecipitation of SMC and MksB interaction partners was performed with strains CBK012 and CBK015, further including strain CBK052 as negative control. Lysate of exponentially grown cells was used for immunoprecipitation via magnetic RFP-Trap® agarose beads. For proteomic analysis samples were further processed and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) to identify and quantify proteins in all samples. A detailed description of immunoprecipitation and proteomic analysis is provided in the **Supplementary Information**.

Bacterial two-hybrid screening

Protein interactions obtained by mass spectrometry were confirmed via bacterial two-hybrid assays [62], using compatible vectors expressing adenylate cyclase subunits T25 and T18 (pKT25/ pKNT25 and pUT18/ pUT18C). *E. coli* BTH101 co-transformed with respective vectors were plated on indicator medium LB/ X-Gal (5-bromo-4-chloro-3-

24

547   indolyl-β-D-galactopyranoside, 40 µg/ml) supplemented with IPTG (0.5 mM) and

548   antibiotics kanamycin (50 µg/ml), carbenicillin (100 µg/ml) and streptomycin (100

549   µg/ml) and incubated at 30°C for 24 h. Interacting hybrid proteins were identified by

550   blue-white screening (not shown) and β-galactosidase assays in a 96 well plate format

551   as previously described [63]. Cotransformants harboring empty plasmids or pUT18C-zip/

552   pKT25-zip plasmids served as positive and negative controls. Miller units of negative

553   controls served as reference and were set to zero; Miller units of any other sample

554   were normalized accordingly. All C- and N-terminal combinations of hybrid proteins

555   were assayed and positive signals were confirmed through at least three replicates.

556   **Fluorescence microscopy**

557   Microscopy images were acquired on an Axio-Imager M1 fluorescence microscope

558   (Carl Zeiss) and a Delta Vision Elite microscope (GE Healthcare, Applied Precision).

559   Experimental procedures and technical configurations are specified in the

560   **Supplementary Information**. Data analysis was performed via FIJI and MicrobeJ

561   software [64,65].

562   **Chromatin Immunoprecipitation (ChIP) combined with sequencing**

563   A detailed description of this method is available in the **Supplementary Information**.

564   Briefly, cells were crosslinked (1 % formaldehyde) for 30 min at room temperature and

565   lysed. DNA was sheared by sonication, incubated with α-mCherry antibody for 2 h at

566   4 ℃, washed subsequently and crosslinks were reverted at 65 ℃ o. N. DNA purification

567   was followed by library preparation and sequencing using an Illumina MiSeq system.

568   Reads were aligned to the *C. glutamicum* ATCC 13032 genome sequence

569   (GeneBankID: BX927147.1). Further data analysis was performed using online tools

570   [66].

571

Real-time PCR

DNA amplification was performed using a 2x qPCR Mastermix (KAPA SYBR®FAST, Peqlab) according to manufacturer's instruction, where reaction volumes of 10 µl contained 200 nM oligonucleotides and 4 µl of diluted DNA, respectively. Samples were measured in technical duplicates via an iQ5 multicolor real-time PCR detection system (Bio-Rad) and CT-values were determined via the Bio Rad-IQ™5 software version 2.1. Primer efficiencies were estimated by calibration dilution curves and slope calculation [67]; data were analyzed via the $2^{-\Delta CT}$ method [68] accounting for dilution factors and sample volumes used for DNA purification. qPCR data of ChIP samples were normalized according to the ParB-mCherry2 signal obtained at locus *parS1* in the wild type background, serving as reference in each experiment.

Electrophoretic mobility shift assay

DNA-ParB binding was assayed using purified protein (for protein purification, see **Supplementary Information**) and double-stranded DNA fragments of approximately 1100 bp length with or without two *parS* sites. Fragments were generated by PCRs of a *C. glutamicum* genomic locus surrounding *parS9* and *parS10* using primer pairs parS9mut-HindIII-up-F/ parS10mut-EcoRI-D-R. ParB concentrations of 0.05 to 25 µM were incubated with 100 ng DNA for 30 min at 30°C, following sample separation in native gels (3-12 % polyacrylamide, ServaGel™). DNA was stained using SYBR® Green I (Invitrogen).

Photoactivated localization microscopy (PALM)

C. glutamicum cells were fixed with 3 % of formaldehyde prior to super resolution imaging using a Zeiss ELYRA P.1 microscope with laser lines HR diode 50 mW 405 nm and HR DPSS 200 mW 561 nm and an Andor EM-CCD iXon DU 897 camera. Cellular ParB-PAmCherry signals were further analyzed using Fiji software[64] and

26

597   identification of distinct protein clusters was carried out by applying the OPTICS

598   algorithm in R [48,49]. A detailed description of sample preparation, parameters for

599   detection of ParB-PAmCherry signals, PALM image calculation and analysis is

600   available in the **Supplementary Information**.

601   **Chromosome Conformation Capture Libraries**

602   Chromosome Conformation Capture (3C/Hi-C) libraries were generated as previously

603   described by Val *et al*. [47] with minor changes, as specified in the **Supplementary**

604   **Information**.

605   **Contact map generation**

606   Contact maps were generated as previously described [37]. Reads were aligned

607   independently (forward and reverse) using Bowtie 2 in local and very sensitive mode

608   and were assigned to a restriction fragment. Non-informative events (self-circularized

609   DNA fragments, or uncut fragments) were discarded by taking into account the pair-

610   reads relative directions and the distribution of the different configurations as described

611   in Cournac *et al* [69]. We then bin the genomes into regular units of 5 Kb to generate

612   contact maps and normalized them using the sequential component normalization

613   procedure (SCN) [69].

614   **Contact map comparison**

615   Ratio between contact maps was computed for each point of the map by dividing the

616   amount of normalized contacts in one condition by the amount of normalized contacts

617   in the other condition and by plotting the Log2 of the ratio. The color code reflects a

618   decrease or increase of contacts in one condition compared to the other (blue or red

619   signal, respectively). No change is represented by a white signal. Ratio maps of

620   replicates are illustrated in Figure S12.

Identification of domains frontiers using directional index

622 To quantify the degree of directional preference, we applied on correlation matrices

623 the same procedure as in Marbouty, et al. [10]. For each 5 Kb bin, we extracted the vector

624 of interactions from the correlation matrix between the studied bin and bins at regular

625 5 Kb intervals, up to 250 Kb in left and right directions. The two vectors were then

626 compared with a paired t-test to assess their statistical significant difference (p=0.05).

627 The directional preferences for the bin along the chromosome are represented as a

628 bar plot with positive and negative t-values shown as red and green bars.

Statistical analysis

630 Correlation coefficients, linear regressions and analyses of nearest neighbor distance

631 distributions were calculated using Excel, Graph Pad Prism (GraphPad Software) and

632 R [49].

Data availability

634 Proteomic data are available via ProteomeXchange with the project identifier

635 PXD008916 [70]. Reviewer account details: username: reviewer52143@ebi.ac.uk,

636 password: zyJpq85V

637 Genome-wide sequencing reads of ChIP-seq and chromosome conformation capture

638 assays generated in this study are available in the Sequence Read Archive (SRA)

639 under accession numbers PRJNA529385 and PRJNA525583.

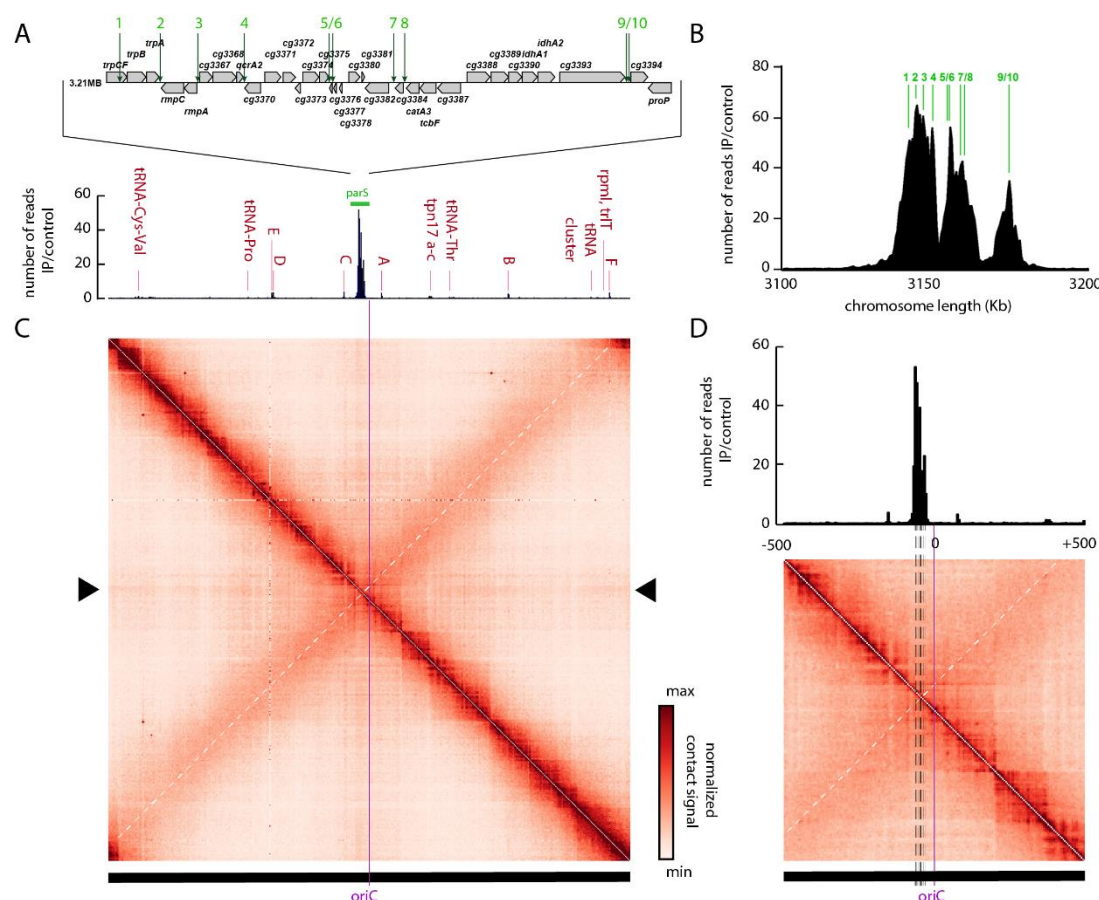641 **Figure legends**

642

643

Figure 1. Chromosome organization hub at *oriC* domain in *C. glutamicum*.

A) Top: Genomic region including ten *parS* sites of *C. glutamicum* with 16 bp consensus sequences. Below: ChIP-seq data on ParB-mCherry DNA binding protein confirm *parS* sites shown above. Exponentially growing *C. glutamicum parB::parB-mCherry* cells (CBK006) were used for *in vivo* anti-mCherry ChIP-seq experiments. Shown is the ratio of ChIP signal relative to the input (fold enrichment IP/control) in 5 Kb bins in linear scale along the chromosome with an x-axis centered at *oriC*. Red labels indicate minor enrichment signals at highly transcribed regions, like rRNA operons (letters A-F). **B)** ParB-ChIP-seq enrichment encompassing 3.1-3.2 Mb genomic region; *parS* sites 1-10 are indicated (green lines). **C)** Normalized genomic contact map derived from asynchronously grown cells (fast growth μ ≥ 0.6 h^{-1},

29

655   exponential phase). X- and Y-axes indicate chromosomal coordinates binned in 5 Kb;

656   *oriC*-centered (purple bar). Color scales, indicated beside the contact map, reflect

657   contact frequency between two genomic loci from white to red (rare to frequent

658   contacts). White dashed line on the contact matrix indicate the mean signal of the

659   secondary diagonal and black triangles on the side of the contact matrix indicate the

660   "cross like" signal. **D)** Structural chromosome organization of the *oriC* region.

661   Magnification of contacts within 500 Kb surrounding *oriC*; *oriC* is indicated as a purple

662   line and *parS* sites are indicated by dashed lines. ParB enrichment zones at *parS* are

663   shown above the contact map (ChIP signal relative to the input in 5 Kb bins). White

664   dashed line on the contact matrix indicate the mean signal of the secondary diagonal.
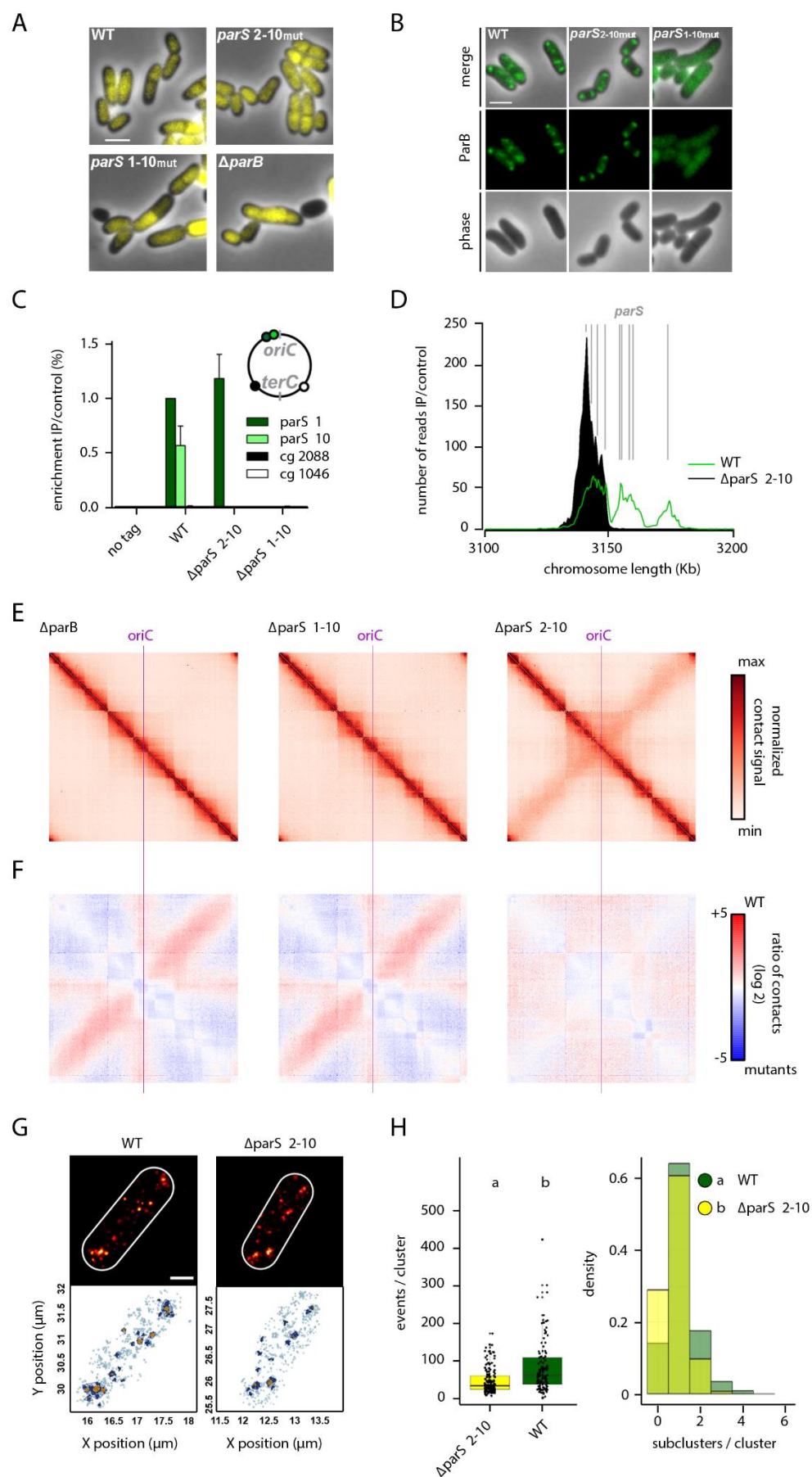
665

666

31

667 **Figure 2. A single *parS* site mediates substantial ParB propagation and wild**
668 **type-like chromosome folding.**

669 **A)** One *parS* site is necessary and sufficient for wild type-like morphology and nucleoid

670 segregation. Shown are phase contrast images of exponentially grown cells harboring

671 either all (WT), one (*parS*$_{2\text{-}10\text{mut}}$, CBK023) or none (*parS*$_{1\text{-}10\text{mut}}$, CBK024) *parS* site(s)

672 or lacking *parB* (Δ*parB*, CDC003). DNA is stained with Hoechst (yellow). Scale bar, 2

673 μm. **B)** One *parS* sequence is sufficient to form wild type-like ParB clusters *in vivo*.

674 Microscopy analysis of *parB::parB-mCherry* (shown in green) in wild type (CBK006),

675 *parS*$_{2\text{-}10\text{mut}}$ (CBK027) and *parS*$_{1\text{-}10\text{mut}}$ backgrounds (CBK028). Absence of *parS* leads

676 to diffuse cellular ParB localizations. Scale bar, 2 μm. **C)** ChIP-qPCR for strains

677 described before, normalized to wild type *parS*1 signal; standard deviations derive from

678 biological triplicates. **D)** ChIP-seq of *C. glutamicum parB::parB-mCherry parS*$_{2\text{-}10\text{mut}}$

679 (black) at a 3.1-3.2 Mb chromosomal range. Wild type-like propagation (green) of ParB

680 protein around *parS1-4*; 0.5 Kb bin size. Location of *parS* sites present in wild type or

681 mutant sequences are indicated (gray lines). **E)** Normalized contact maps of Δ*parB*,

682 *parS*$_{1\text{-}10\text{mut}}$ and *parS*$_{2\text{-}10\text{mut}}$ mutants centered at *oriC* (CDC003, CBK024, CBK023).

683 Color codes as in Fig. 1 were applied. **F)** Differential maps correspond to the log2 of

684 the ratio (wild type norm/ mutant norm); color scales indicate contact enrichment in

685 mutant (blue) or wild type (red) (white indicates no differences between the two

686 conditions). **G)** Single-molecule localization microscopy of representative wild type and

687 *parS*$_{2\text{-}10\text{mut}}$ cells (CBK029, CBK031). Top: Gaussian rendering of ParB-PAmCherry

688 signals (0.71 PSF, 1 px = 10 nm), below: color-coded representation of ParB-

689 PAmCherry events within corresponding cells [48]; all events (light blue), macroclusters

690 (dark blue) and subclusters (yellow) are indicated. Scale bar, 0.5 μm. For detailed

691 parameters see Material and Methods and Fig. S6. **H)** Comparison of ParB-

692 PAmCherry cluster properties in strains described above. Top: Events per

693   macrocluster, medians are indicated as solid lines and whiskers mark 1.5 IQR (Inter

694   Quartile Ranges). Note that only the two biggest clusters per cell were taken into

695   account for analyses (clusters$_{wild\ type}$: n = 130, clusters$_{parS2\text{-}10}$: n = 143). Below:

696   Subcluster numbers per macrocluster shown as overlay bar chart for both strains. Both

697   subcluster numbers per macrocluster (Kruskal-Wallis Rank Sum Test: chi-squared =

698   12.284, df = 1, p<0.05) and macroclusters size (Kruskal-Wallis Rank Sum Test: chi-

699   squared = 27.582, df = 1, p<0.05) differ significantly between both strains.
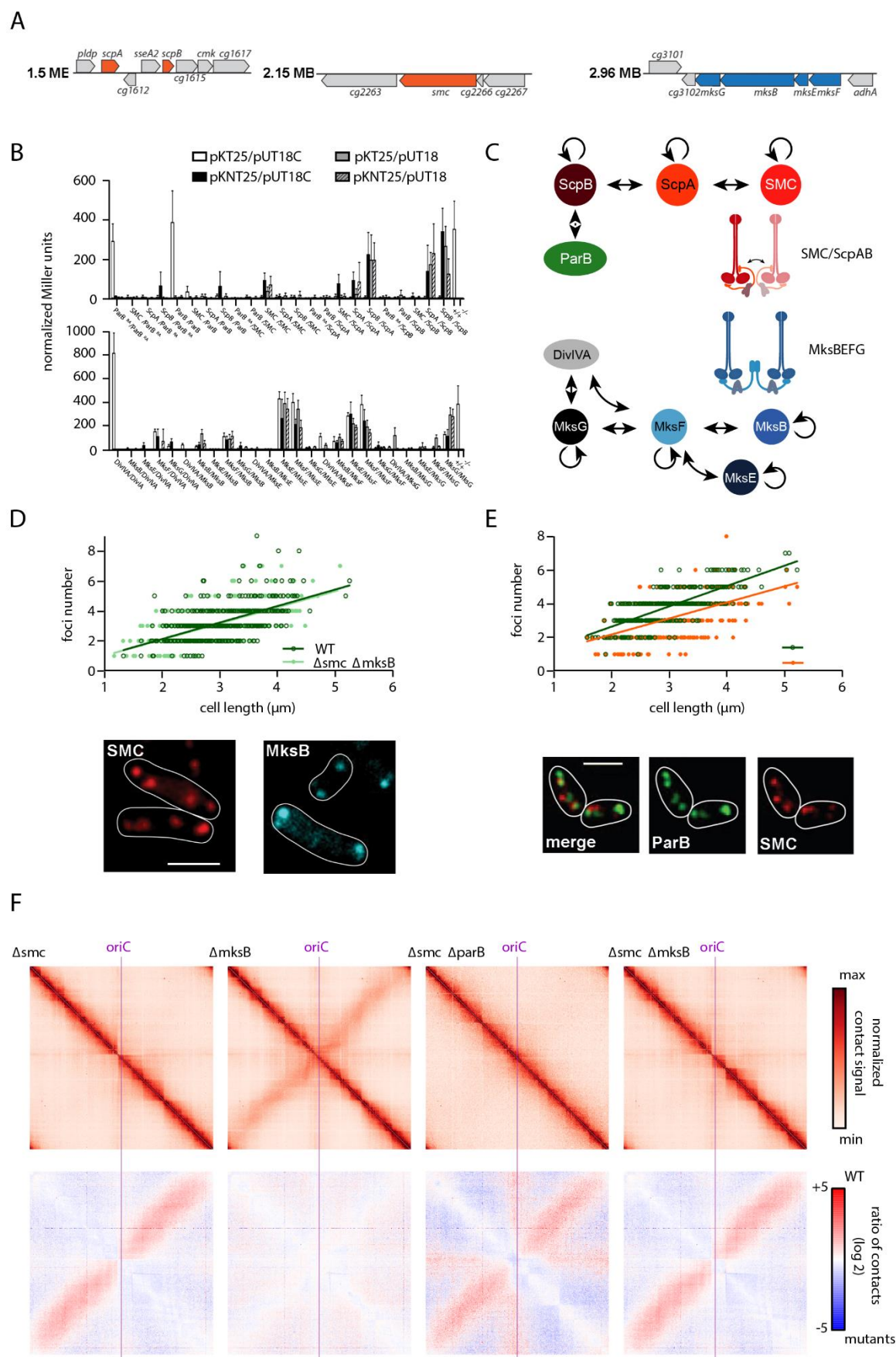
700

701
**Figure 3. Functional characterization of two SMC-like complexes in *C.
glutamicum*.**

34

704   **A)** Sections of the *C. glutamicum* genome map indicating localizations of condensin

705   subunit genes. **B)** Confirmation of protein-protein interactions via bacterial two-hybrid

706   screen. Interactions were quantified by β-galactosidase assays in all combinations of

707   hybrid proteins: C/C- (18C/T25), N/C- (18/T25), C/N- (18C/NT25), and N/N- (18/NT25)

708   terminal fusions of adenylate cyclase fragments, ParB^{RA}: ParB mutant R175A. **C)**

709   Illustration of SMC/ScpAB and MksBEFG subunit interactions based on bacterial two-

710   hybrid data, cartoons indicate condensin complex formations. **D)** Top: Dependence of

711   ParB foci numbers on cell length in *C. glutamicum* wild type (WT) and Δ*smc* Δ*mksB*

712   (ΔΔ, CBK011) cells grown in BHI (n>350). Linear regression lines are shown

713   r(WT)=0.57, r(ΔΔ)=0.62; slopes and intercepts are equal (ANCOVA, $F_{(1, 770)}$=0.059,

714   p >.05; ANCOVA, $F_{(1, 771)}$=0.60, p <.05). Below: Cellular localization of condensin

715   subunits in *C. glutamicum smc::smc-mCherry* and *mksB::mksB-mCherry* cells

716   (CBK012, CBK015). Microscopy images exemplify cellular mCherry fluorescence of

717   SMC (left) and MksB (right); white lines indicate cell outlines. Scale bar, 2 μm. **E)** Top:

718   SMC and ParB foci numbers positively correlate with cell length in double labeled strain

719   *smc::smc-mCherry parB::parB-mNeonGreen* (CBK013), r(ParB)=0.74, r(SMC)=0.53;

720   (n>350). Below: Subcellular localization of ParB and SMC is exemplified in

721   representative cells shown in overlays between mNeonGreen and mCherry

722   fluorescence and in separate channels. Scale bar, 2 μm. **F)** Normalized contact maps

723   of Δ*smc*, Δ*mksB*, Δ*parB*/Δ*smc* and Δ*smc*/Δ*mksB* mutants (CDC026, CBK001,

724   CBK002, CBK004), displayed as in Fig. 1. Corresponding differential maps indicating

725   the log of the ratio (wild type norm/ mutant norm) are presented as in Fig. 2.
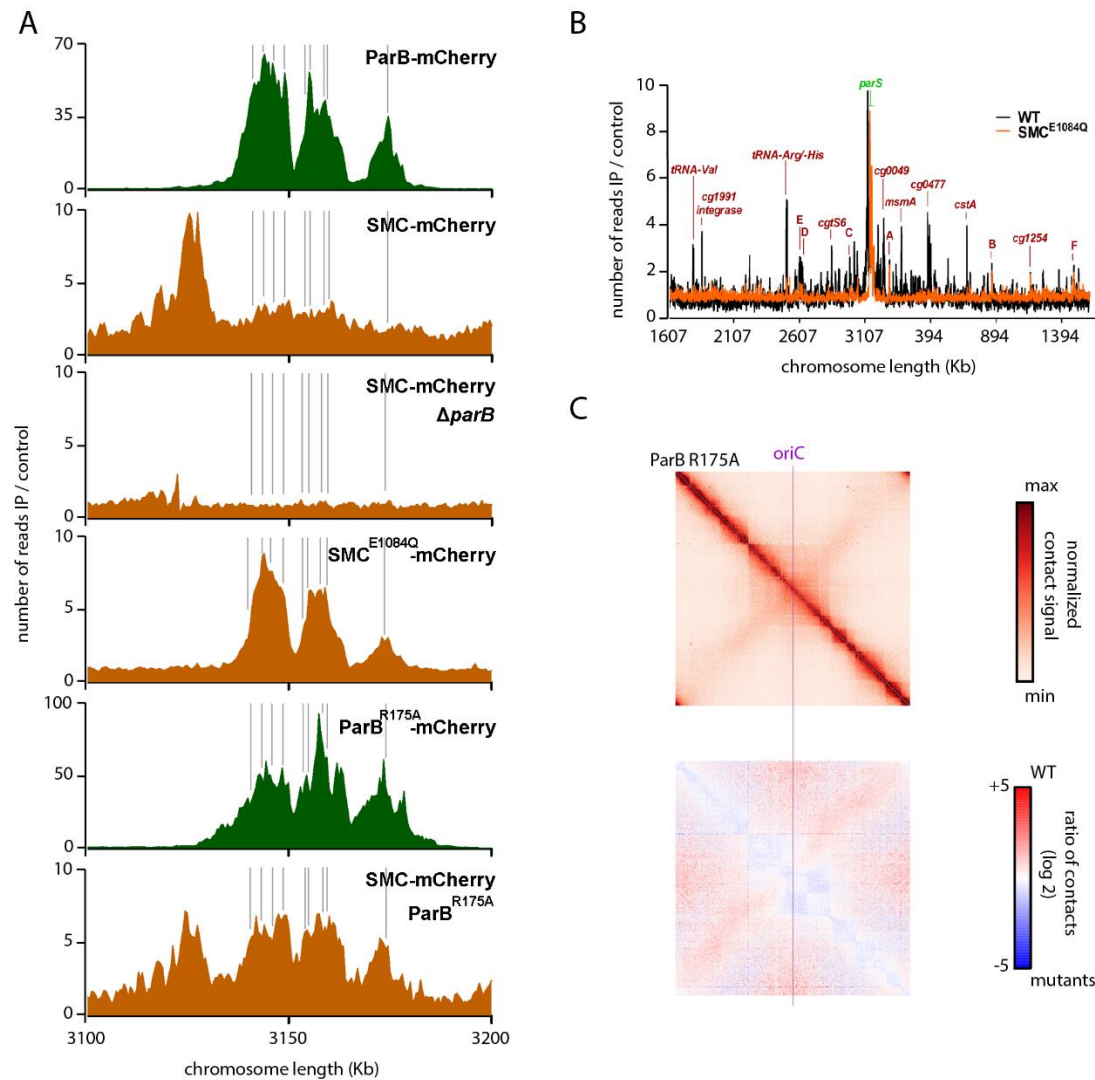
726

Figure 4. Chromosomal SMC-loading is mediated by ParB at *parS* sites.

A) SMC-enrichment at *parS* sites (gray) is ParB-dependent. ChIP-seq of ParB-mCherry (green, CBK006, CBK047) and SMC-mCherry (orange, CBK012, CBK014, CBK051, CBK049) in strain backgrounds as indicated. Depicted are chromosomal ranges of 3.1-3.2 Mb, bin size 0.5 Kb. **B)** Whole-genome ChIP-seq data of strains harboring SMC-mCherry wild type (gray, CBK012) or E1084Q mutant (orange, CBK051). SMC enrichment at *parS* sites and at other loci (red letters), in particular tRNA gene clusters and at rRNA genes (A-F) is illustrated in 0.5 Kb bins in linear scale along the chromosome with an x-axis centered at *oriC*. **C)** Normalized contact map of

36

737   mutant strains *parB::parB*[R175A] (CBK047) and the corresponding differential map

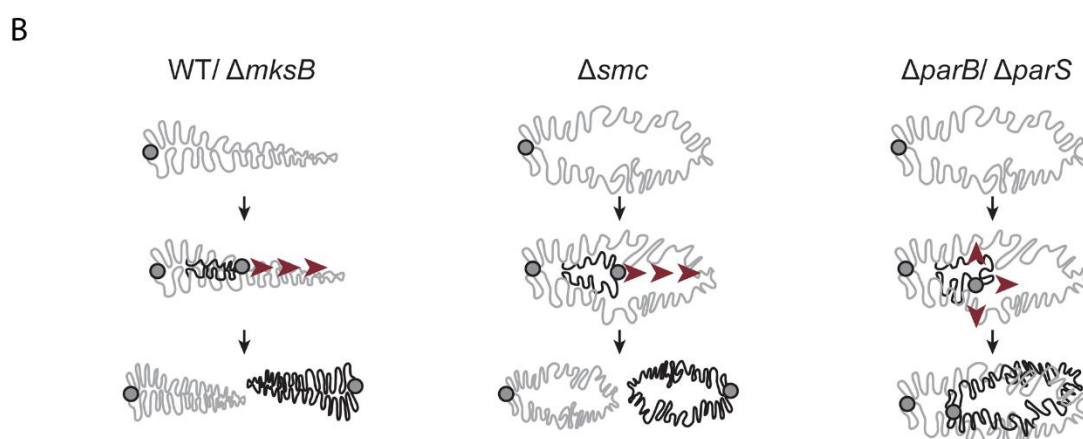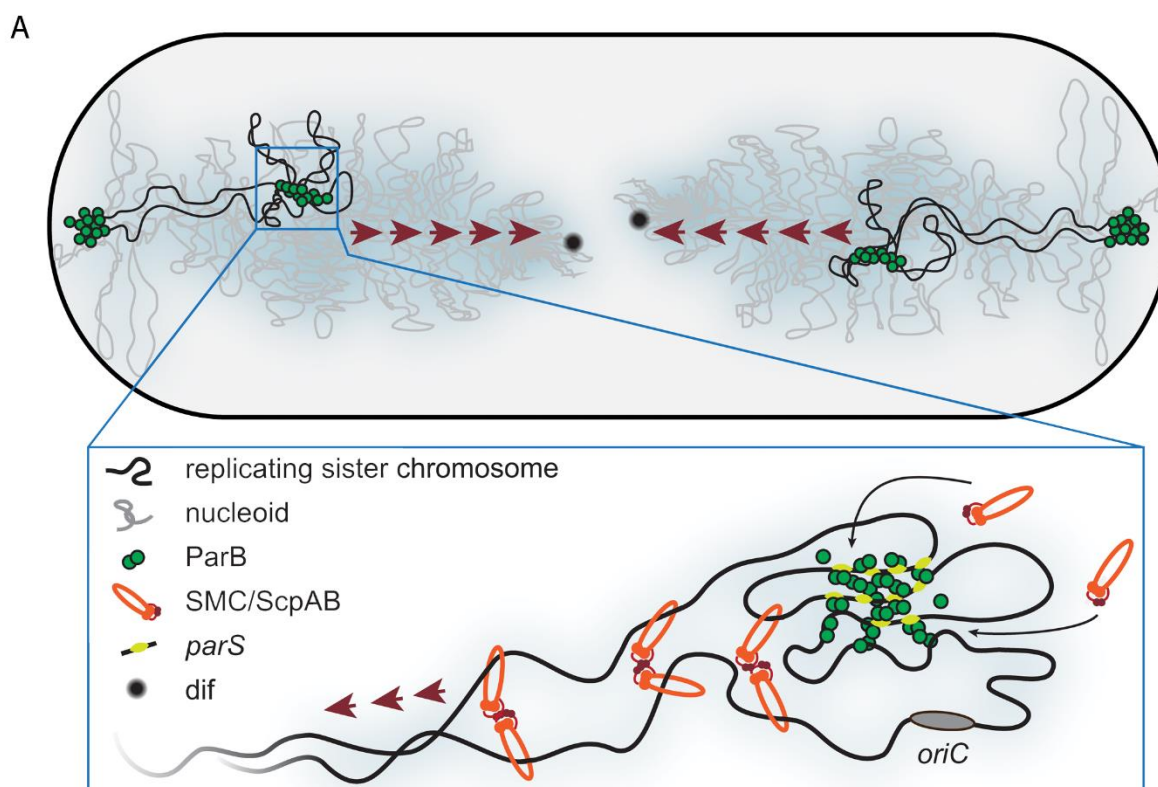738   indicating the log of the ratio (wild type norm/ mutant norm) as in Fig. 2.

739

Figure 5. Model of ParB- and SMC-mediated chromosome organization in *C. glutamicum*.

A) Top: Illustration of chromosome segregation in a diploid *C. glutamicum* cell. Newly

replicated ParB-*oriC*s complexes translocate from cell poles towards *terC*s at septal

positions via a ParABS system, where ParB nucleocomplexes use ParA-bound sister

chromosomal loci as transient tethers for translocation across the nucleoid. Below:

Condensins SMC/ScpAB are loaded ParB-dependently at *parS* sites and relocate to

distant chromosomal regions (red arrows) causing inter-linkage of chromosomal arms;

38

748   here illustration is based on a two-translocator model. **B)** Cartoon showing mutant

749   phenotypes of overall chromosome folding and segregation compared to wild type

750   (WT). Replication *oriC* (grey mark), chromosome (grey line) and replicating sister DNA

751   (black line) are indicated; red arrows indicate direction of DNA segregation.

752

References

1   Sazer, S. & Schiessel, H. The biology and polymer physics underlying large-scale chromosome organization. *Traffic (Copenhagen, Denmark)* **19**, 87-104, doi:doi:10.1111/tra.12539 (2018).

2   Badrinarayanan, A., Le, T. B. K. & Laub, M. T. Bacterial chromosome organization and segregation. *Annual review of cell and developmental biology* **31**, 171-199, doi:10.1146/annurev-cellbio-100814-125211 (2015).

3   Gruber, S. Multilayer chromosome organization through DNA bending, bridging and extrusion. *Current opinion in microbiology* **22**, 102-110, doi:10.1016/j.mib.2014.09.018 (2014).

4   Kleckner, N. *et al.* The bacterial nucleoid: nature, dynamics and sister segregation. *Current opinion in microbiology* **22**, 127-137, doi:10.1016/j.mib.2014.10.001 (2014).

5   Livny, J., Yamaichi, Y. & Waldor, M. K. Distribution of centromere-like *parS* sites in bacteria: insights from comparative genomics. *Journal of bacteriology* **189**, 8693-8703, doi:10.1128/jb.01239-07 (2007).

6   Rodionov, O., Lobocka, M. & Yarmolinsky, M. Silencing of genes flanking the P1 plasmid centromere. *Science* **283**, 546-549 (1999).

7   Murray, H., Ferreira, H. & Errington, J. The bacterial chromosome segregation protein Spo0J spreads along DNA from *parS* nucleation sites. *Molecular microbiology* **61**, 1352-1361, doi:10.1111/j.1365-2958.2006.05316.x (2006).

8   Graham, T. G. W. *et al.* ParB spreading requires DNA bridging. *Genes & development* **28**, 1228-1238, doi:10.1101/gad.242206.114 (2014).

9   Broedersz, C. P. *et al.* Condensation and localization of the partitioning protein ParB on the bacterial chromosome. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8809-8814, doi:10.1073/pnas.1402529111 (2014).

10   Marbouty, M. *et al.* Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. *Molecular cell* **59**, 588-602, doi:10.1016/j.molcel.2015.07.020 (2015).

11   Sanchez, A. *et al.* Stochastic self-assembly of ParB proteins builds the bacterial DNA segregation apparatus. *Cell systems* **1**, 163-173, doi:10.1016/j.cels.2015.07.013 (2015).

12   Leonard, T. A., Butler, P. J. & Löwe, J. Bacterial chromosome segregation: structure and DNA binding of the Soj dimer – a conserved biological switch. *The EMBO journal* **24**, 270-282, doi:10.1038/sj.emboj.7600530 (2005).

13   Zhang, H. & Schumacher, M. A. Structures of partition protein ParA with nonspecific DNA and ParB effector reveal molecular insights into principles governing Walker-box DNA segregation. *Genes & development* **31**, 481-492, doi:10.1101/gad.296319.117 (2017).

14   Vecchiarelli, A. G., Hwang, L. C. & Mizuuchi, K. Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E1390-1397, doi:10.1073/pnas.1302745110 (2013).

15   Vecchiarelli, A. G., Neuman, K. C. & Mizuuchi, K. A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4880-4885, doi:10.1073/pnas.1401025111 (2014).

802 16   Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C. & Liu, J. Directed and
803   persistent movement arises from mechanochemistry of the ParA/ParB system.
804   *Proceedings of the National Academy of Sciences of the United States of*
805   *America* **112**, E7055-E7064, doi:10.1073/pnas.1505147112 (2015).
806 17   Surovtsev, I. V., Campos, M. & Jacobs-Wagner, C. DNA-relay mechanism is
807   sufficient to explain ParA-dependent intracellular transport and patterning of
808   single and multiple cargos. *Proceedings of the National Academy of Sciences*
809   *of the United States of America* **113**, E7268-E7276,
810   doi:10.1073/pnas.1616118113 (2016).
811 18   Lagage, V., Boccard, F. & Vallet-Gely, I. Regional control of chromosome
812   segregation in *Pseudomonas aeruginosa*. *PLoS genetics* **12**, e1006428,
813   doi:10.1371/journal.pgen.1006428 (2016).
814 19   Tran, N. T. *et al.* Permissive zones for the centromere-binding protein ParB on
815   the *Caulobacter crescentus* chromosome. *Nucleic acids research* **46**, 1196-
816   1209, doi:10.1093/nar/gkx1192 (2018).
817 20   Sullivan, N. L., Marquis, K. A. & Rudner, D. Z. Recruitment of SMC by ParB-
818   *parS* organizes the origin region and promotes efficient chromosome
819   segregation. *Cell* **137**, 697-707, doi:10.1016/j.cell.2009.04.044 (2009).
820 21   Wang, X. *et al.* Condensin promotes the juxtaposition of DNA flanking its
821   loading site in *Bacillus subtilis*. *Genes & development* **29**, 1661-1675,
822   doi:10.1101/gad.265876.115 (2015).
823 22   Tran, N. T., Laub, M. T. & Le, T. B. K. SMC progressively aligns chromosomal
824   arms in *Caulobacter crescentus* but is antagonized by convergent
825   transcription. *Cell reports* **20**, 2057-2071, doi:10.1016/j.celrep.2017.08.026
826   (2017).
827 23   Cobbe, N. & Heck, M. M. The evolution of SMC proteins: phylogenetic
828   analysis and structural implications. *Molecular biology and evolution* **21**, 332-
829   347, doi:10.1093/molbev/msh023 (2004).
830 24   Niki, H., Jaffe, A., Imamura, R., Ogura, T. & Hiraga, S. The new gene *mukB*
831   codes for a 177 kd protein with coiled-coil domains involved in chromosome
832   partitioning of *E. coli*. *The EMBO journal* **10**, 183-193 (1991).
833 25   Britton, R. A., Lin, D. C. & Grossman, A. D. Characterization of a prokaryotic
834   SMC protein involved in chromosome partitioning. *Genes & development* **12**,
835   1254-1259 (1998).
836 26   Jensen, R. B. & Shapiro, L. The *Caulobacter crescentus smc* gene is required
837   for cell cycle progression and chromosome segregation. *Proceedings of the*
838   *National Academy of Sciences of the United States of America* **96**, 10661-
839   10666 (1999).
840 27   Bürmann, F. *et al.* An asymmetric SMC–kleisin bridge in prokaryotic
841   condensin. *Nature Structural &Amp; Molecular Biology* **20**, 371,
842   doi:10.1038/nsmb.2488 (2013).
843 28   Diebold-Durand, M. L. *et al.* Structure of full-length SMC and rearrangements
844   required for chromosome organization. *Molecular cell* **67**, 334-347.e335,
845   doi:10.1016/j.molcel.2017.06.010 (2017).
846 29   Terakawa, T. *et al.* The condensin complex is a mechanochemical motor that
847   translocates along DNA. *Science* **358**, 672-676, doi:10.1126/science.aan6516
848   (2017).
849 30   Ganji, M. *et al.* Real-time imaging of DNA loop extrusion by condensin.
850   *Science*, doi:10.1126/science.aar7831 (2018).

851 31 Wang, X., Brandão, H. B., Le, T. B. K., Laub, M. T. & Rudner, D. Z. *Bacillus*
852   *subtilis* SMC complexes juxtapose chromosome arms as they travel from
853   origin to terminus. *Science* **355**, 524-527, doi:10.1126/science.aai8982 (2017).
854 32 Minnen, A. *et al.* Control of Smc coiled coil architecture by the ATPase heads
855   facilitates targeting to chromosomal ParB/*parS* and release onto flanking DNA.
856   *Cell Reports* **14**, 2003-2016, doi:10.1016/j.celrep.2016.01.066 (2016).
857 33 Minnen, A., Attaiech, L., Thon, M., Gruber, S. & Veening, J. W. SMC is
858   recruited to oriC by ParB and promotes chromosome segregation in
859   *Streptococcus pneumoniae*. *Molecular microbiology* **81**, 676-688,
860   doi:10.1111/j.1365-2958.2011.07722.x (2011).
861 34 Gruber, S. & Errington, J. Recruitment of condensin to replication origin
862   regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell*
863   **137**, 685-696, doi:10.1016/j.cell.2009.02.035 (2009).
864 35 Le, T. B., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-resolution mapping
865   of the spatial organization of a bacterial chromosome. *Science* **342**, 731-734,
866   doi:10.1126/science.1242059 (2013).
867 36 Wang, X., Liu, X., Possoz, C. & Sherratt, D. J. The two *Escherichia coli*
868   chromosome arms locate to separate cell halves. *Genes & development* **20**,
869   1727-1731, doi:10.1101/gad.388406 (2006).
870 37 Lioy, V. S. *et al.* Multiscale structuring of the *E. coli* chromosome by nucleoid-
871   associated and condensin proteins. *Cell* **172**, 771-783.e718,
872   doi:10.1016/j.cell.2017.12.027 (2018).
873 38 Nolivos, S. *et al.* MatP regulates the coordinated action of topoisomerase IV
874   and MukBEF in chromosome segregation. *Nature communications* **7**, 10466-
875   10466, doi:10.1038/ncomms10466 (2016).
876 39 Petrushenko, Z. M., She, W. & Rybenkov, V. V. A new family of bacterial
877   condensins. *Molecular microbiology* **81**, 881-896, doi:10.1111/j.1365-
878   2958.2011.07763.x (2011).
879 40 Böhm, K. *et al.* Novel chromosome organization pattern in *Actinomycetales*—
880   overlapping replication cycles combined with diploidy. *mBio* **8**,
881   doi:10.1128/mBio.00511-17 (2017).
882 41 Donovan, C., Schwaiger, A., Krämer, R. & Bramkamp, M. Subcellular
883   localization and characterization of the ParAB system from *Corynebacterium*
884   *glutamicum*. *Journal of bacteriology* **192**, 3441-3451, doi:10.1128/jb.00214-10
885   (2010).
886 42 Donovan, C., Sieger, B., Krämer, R. & Bramkamp, M. A synthetic *Escherichia*
887   *coli* system identifies a conserved origin tethering factor in Actinobacteria.
888   *Molecular microbiology* **84**, 105-116, doi:10.1111/j.1365-2958.2012.08011.x
889   (2012).
890 43 Ireton, K., Gunther, N. W. & Grossman, A. D. *spo0J* is required for normal
891   chromosome segregation as well as the initiation of sporulation in *Bacillus*
892   *subtilis*. *Journal of bacteriology* **176**, 5320-5329 (1994).
893 44 Lee, P. S. & Grossman, A. D. The chromosome partitioning proteins Soj
894   (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning,
895   separation of replicated sister origins, and regulation of replication initiation in
896   *Bacillus subtilis*. *Molecular microbiology* **60**, 853-869, doi:10.1111/j.1365-
897   2958.2006.05140.x (2006).
898 45 Donovan, C., Schauss, A., Krämer, R. & Bramkamp, M. Chromosome
899   segregation impacts on cell growth and division site selection in
900   *Corynebacterium glutamicum*. *PloS one* **8**, e55078,
901   doi:10.1371/journal.pone.0055078 (2013).

46 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306-1311, doi:10.1126/science.1067799 (2002).

47 Val, M.-E. *et al.* A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Science Advances* **2**, doi:10.1126/sciadv.1501914 (2016).

48 Mihael Ankerst, M. M. B., Hans-Peter Kriegel, Joerg Sander. in *ACM SIGMOD international conference on Management of data.* 49-60 (ACM Press).

49 Team, R. C. (URL: http://www.R-project.org/, 2014).

50 Schwartz, M. A. & Shapiro, L. An SMC ATPase mutant disrupts chromosome segregation in *Caulobacter*. *Molecular microbiology* **82**, 1359-1374, doi:10.1111/j.1365-2958.2011.07836.x (2011).

51 Hirano, M. & Hirano, T. Positive and negative regulation of SMC–DNA interactions by ATP and accessory proteins. *The EMBO journal* **23**, 2664-2673, doi:10.1038/sj.emboj.7600264 (2004).

52 Hirano, M. & Hirano, T. Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions. *Molecular cell* **21**, 175-186, doi:10.1016/j.molcel.2005.11.026 (2006).

53 Hassler, M., Shaltiel, I. A. & Haering, C. H. Towards a unified model of SMC complex function. *Current biology : CB* **28**, R1266-r1281, doi:10.1016/j.cub.2018.08.034 (2018).

54 Doron, S. *et al.* Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, doi:10.1126/science.aar4120 (2018).

55 Roth, S. *et al.* Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1β production. *Nature Immunology* **15**, 538, doi:10.1038/ni.2888

https://www.nature.com/articles/ni.2888#supplementary-information (2014).

56 Aragón, L. The Smc5/6 complex: new and old functions of the enigmatic long-distance relative. *Annual review of genetics* **52**, 89-107, doi:10.1146/annurev-genet-120417-031353 (2018).

57 Yao, S., Helinski, D. R. & Toukdarian, A. Localization of the naturally occurring plasmid ColE1 at the cell pole. *Journal of bacteriology* **189**, 1946-1953, doi:10.1128/jb.01451-06 (2007).

58 Niki, H. & Yano, K. In vitro topological loading of bacterial condensin MukB on DNA, preferentially single-stranded DNA rather than double-stranded DNA. *Scientific reports* **6**, 29469, doi:10.1038/srep29469 (2016).

59 Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M. C. & Sherratt, D. J. In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science (New York, N.Y.)* **338**, 528-531, doi:10.1126/science.1227126 (2012).

60 Woo, J. S. *et al.* Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell* **136**, 85-96, doi:10.1016/j.cell.2008.10.050 (2009).

61 Debaugny, R. E. *et al.* A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids. *Molecular Systems Biology* **14**, doi:10.15252/msb.20188516 (2018).

62 Karimova, G., Pidoux, J., Ullmann, A. & Ladant, D. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5752-5756 (1998).

950   63   Mehla, J., Caufield, J. H., Sakhawalkar, N. & Uetz, P. A comparison of two-
951   hybrid approaches for detecting protein-protein interactions. *Methods in*
952   *enzymology* **586**, 333-358, doi:10.1016/bs.mie.2016.10.020 (2017).
953   64   Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis.
954   *Nat Meth* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
955   65   Ducret, A., Quardokus, E. M. & Brun, Y. V. MicrobeJ, a tool for high
956   throughput bacterial cell detection and quantitative analysis. *Nature*
957   *Microbiology* **1**, 16077, doi:10.1038/nmicrobiol.2016.77 (2016).
958   66   Ramírez, F. *et al.* deepTools2: a next generation web server for deep-
959   sequencing data analysis. *Nucleic acids research* **44**, W160-W165,
960   doi:10.1093/nar/gkw257 (2016).
961   67   Rasmussen, R. in *Rapid Cycle Real-Time PCR Methods and Applications*
962   (ed Wittwer C Meuer S, Nakagawara K (ed.)) 21-34 (Springer Berlin
963   Heidelberg, 2001).
964   68   Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data
965   using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.
966   *Methods* **25**, 402-408, doi:10.1006/meth.2001.1262 (2001).
967   69   Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R. & Mozziconacci, J.
968   Normalization of a chromosomal contact map. *BMC Genomics* **13**, 436,
969   doi:10.1186/1471-2164-13-436 (2012).
970   70   Deutsch, E. W. *et al.* The ProteomeXchange consortium in 2017: supporting
971   the cultural change in proteomics public data deposition. *Nucleic acids*
972   *research* **45**, D1100-d1106, doi:10.1093/nar/gkw936 (2017).
973