1   *Escherichia coli* **NusG links the lead ribosome with the transcription elongation**

2   **complex**

3

4   Robert S. Washburn[1], Philipp K. Zuber[2], Ming Sun[3,4], Yaser Hashem[5,6], Bingxin Shen[5,7],

5   Wen Li[5], Sho Harvey[8,9], Stefan H. Knauer[2,*], Joachim Frank[3,5,*], and Max E.

6   Gottesman[1,5,*]

7

8   [1]Department of Microbiology & Immunology, Columbia University Medical Center,

9   New York, NY 10032, USA.

10   [2]Biopolymers, University of Bayreuth, Universitaetsstrasse 30, 95447 Bayreuth,

11   Germany.

12   [3]Department of Biological Sciences, Columbia University, New York, NY 10027, USA.

13   [4]Current address: University of California at San Francisco, San Francisco, CA 94158,

14   USA.

15   [5]Department of Biochemistry and Molecular Biophysics, Columbia University Medical

16   Center, New York, NY 10032, USA.

17   [6]Current address: NSERM U1212, Institut Européen de Chimie et Biologie, University of

18   Bordeaux, Pessac 33607, France.

19   [7]Current address: Bristol-Myers Squibb Pharmaceutical Co., New Brunswick, NJ 08901,

20   USA.

21   [8]University of Michigan, Ann Arbor, MI 48109, USA.

22   [9]Current address: California Institute of Technology, Pasadena, CA 91125, USA.

23

24   *Corresponding Authors

Abstract

26   It has been known for more than 50 years that transcription and translation are physically

27   coupled in bacteria, but whether or not this coupling may be mediated by the two-domain

28   protein N-utilization substance (Nus) G in *Escherichia coli* is still heavily debated. Here,

29   we combine integrative structural biology and functional analyses to provide conclusive

30   evidence that NusG can physically link transcription with translation by contacting both

31   RNA polymerase and the ribosome. We present a cryo-electron microscopy structure of a

32   NusG:70S ribosome complex and nuclear magnetic resonance spectroscopy data

33   revealing simultaneous binding of NusG to RNAP and the intact 70S ribosome, providing

34   the first direct structural evidence for NusG-mediated coupling. Furthermore, *in vivo*

35   reporter assays show that recruitment of NusG occurs late in transcription and strongly

36   depends on translation. Thus, our data suggest that coupling occurs initially via direct

37   RNAP:ribosome contacts and is then mediated by NusG.

38

Introduction

Gene expression is a universal process in all cells and consists of transcription, i.e. the synthesis of RNA based on the DNA, and – if RNA is not the final gene product – translation, i.e. the messenger RNA (mRNA)-guided synthesis of a protein. Since the late 1960s it has been known that the rates of transcription and translation are synchronized in *Echerichia coli* so that mRNA is translated while being transcribed (Das et al., 1967; Mehdi and Yudkin, 1967; Miller et al., 1970; Proshkin et al., 2010; Vogel and Jensen, 1995, 1994). This process,called transcription:translation coupling, is possible due to the lack of a physical barrier between transcription and translation in bacteria (reviewed in (Conn et al., 2019)). Only recently, direct physical interactions between RNA polymerase (RNAP) and the ribosomes have been demonstrated (Demo et al., 2017; Fan et al., 2017; Kohler et al., 2017), consistent with earlier observations that transcriptional events may control translation activity and *vice versa*(Proshkin et al., 2010). As transcription and translation are closely connected to other central processes in a bacterial cell, such as DNA repair (Pani and Nudler, 2017) and protein folding (Thommen et al., 2017), transcription:translation coupling constitutes one of the key resulatory functions in bacterial gene expression.

However, there are also indications that transcription:translation coupling may involve a member of the family of N-utilization substance (Nus) G proteins, which serves as adapter connecting RNAP and the lead ribosome (Burmann et al., 2012, 2010; Saxena et al., 2018; Zuber et al., 2019). *E. coli* Nus G, member and eponym of the only universally conserved class of transcription factors (Werner, 2012), consists of two

3

62 domain, an N- and a C-terminal domain (NTD and CTD, resepectively) connected v*ia* a

63 flexible linker, which move independently (Burmann et al., 2011; Mooney et al., 2009).

64 NusG-NTD binds RNAP and accelerates transcription elongation (Burova et al., 1995;

65 Kang et al., 2018; Mooney et al., 2009). Structural studies demonstrate that NusG-CTD,

66 which is a five-stranded β-barrel with a Kyrpides-Ouzounis-Woese motif (Kyrpides et al.,

67 1996), is a versatile binding platform for different transcription factors. By binding to

68 protein S10, which is part of the 30S subunit of the ribosome, NusG may link

69 transcription and translation (Burmann et al., 2010). Saxena *et al* also demonstrated

70 specific 1:1 binding of NusG to 70S ribosomes both *in vitro* and *in vivo* (Saxena et al.,

71 2018).

72   S10 is identical with transcription factor NusE and forms a ribosome-free

73 complex with NusB, NusA and NusG which suppresses transcription termination

74 (Dudenhoeffer et al., 2019; Krupp et al., 2019; Said et al., 2017; Squires et al., 1993).

75 Finally, NusG-CTD binds to termination factor Rho and is required for most Rho activity

76 *in vivo* (Burmann et al., 2010; Lawson et al., 2018; Mitra et al., 2017).

77 Transcription:translation coupling prevents Rho factor from terminating transcription by

78 sequestering the NusG-CTD and by blocking Rho access to RNAP via untranslated

79 mRNA. Cryptic *E. coli* Rho-dependent terminators located within open reading frames

80 (orfs) are revealed when ribosomes are released by polar nonsense mutations (Cardinale

81 et al., 2008; Newton et al., 1965).

82   Nevertheless, there is evidence for intragenic uncoupling and Rho-dependent

83 transcription termination in the absence of nonsense mutations.   Washburn and

84 Gottesman (Washburn and Gottesman, 2011) and Dutta et al. (Dutta et al., 2011) found

85   that Rho resolves clashes between transcription and replication. Such conflicts are likely

86   to occur within, rather than at the end of, genes. Uncoupling would allow Rho to release

87   the stationary transcription elongation complexes (TECs).

88   Mutations in *nusE* or *nusG* that uncouple transcription from translation increase

89   sensitivity to chloramphenicol (Saxena et al., 2018). This antibiotic retards translation,

90   breaking the bond between the lead ribosome and TEC. Uncoupled TEC may backtrack

91   or terminate prematurely (Dutta et al., 2011).

92   In this report, we present a cryo-electron microscopy (cryo-EM) structure

93   showing NusG binding to the S10 subunit in a 70S ribosome. The NusG-CTD binding

94   site of S10 is also target of the ribosome-release factor, transfer-messenger (tmRNA),

95   raising the possibility that tmRNA might displace NusG at rare codons, thereby

96   uncoupling transcription from translation (Roche and Sauer, 1999). We also show by

97   solution-state nuclear magnetic resonance (NMR) spectroscopy that NusG, once bound to

98   RNAP, can interact with S10 or with a complete ribosome, setting the structural basis for

99   coupling.

100   NusG couples transcription with translation *in vivo*, as proposed earlier (Burmann

101   et al., 2010). Uncoupling of RNAP from the lead ribosome is enhanced when translation

102   is compromised. Importantly, we demonstrate that uncoupled RNAP can outpace

103   translation, leading to Rho-dependent transcription termination. This intragenic

104   termination explains the necessity for the apparent perfect synchronization between

105   transcription and translation (Proshkin et al., 2010).

106

107   **Results**

108   **Structural evidence of NusG binding to the ribosomal S10 subunit on a 70S**

109   **ribosome.**

110   We assembled a NusG:70S complex by incubating 70S ribosomes with an excess of

111   NusG and determined the structure of this complex by cryo-EM and single-particle

112   reconstruction. Overall, 188,127 particles were extracted from 1327 images and ~5% of

113   these particles showed an extra mass of density attached to the mass identified as protein

114   S10 (Fig. 1A,B). This additional density perfectly matches the size of NusG-CTD,

115   suggesting that NusG binds at the site predicted from the solution NMR structure of

116   NusG-CTD bound to the free ribosomal protein S10 in a 1:1 stoichiometry (Fig. 1A,B;

117   (Burmann et al., 2010)). The density map reconstructed from the class of NusG:70S

118   particles was refined to an average resolution of 6.8 Å. No density could be observed for

119   NusG-NTD, indicating that it is flexibly bound to the NusG-CTD and does not interact

120   with the ribosome.

121   During translation ribosomes may stall on incomplete mRNAs, i.e. they reach the

122   3' end of an mRNA without terminating, resulting in an unproductive translation

123   complex. Together with the small protein B (SmpB) tmRNA can bind to these stalled

124   ribosomes in order to rescue them and to tag the nascent polypeptide chain for

125   degradation in a process called trans-translation (Weis et al., 2010). Interestingly, the

126   NusG-CTD binding site overlaps with the region of S10 that is contacted by the tmRNA

127   when it is bound to a ribosome in its resume state (Fig. 1C; (Burmann et al., 2010; Fu et

128   al., 2010; Rae et al., 2019; Weis et al., 2010)). From this we conclude that NusG-CTD

129   and tmRNA share binding sites on S10, raising the possibility that, in addition to

6

130 releasing stalled ribosomes, tmRNA competes with NusG for ribosome binding, thus

131 preventing NusG from maintaining a linkage between the lead ribosome and RNAP. In

132 other words, tmRNA might be able to displace NusG and thereby facilitate uncoupled

133 transcription.

134

135 **Simultaneous binding of NusG to S10 and RNAP**

136 In the cryo-EM structure of *E. coli* NusG bound to a paused TEC (Kang et al., 2018) only

137 the density of NusG-NTD was observable, indicating that NusG-CTD moves freely and

138 does not interact with RNAP. Binding of NusG-CTD to S10 was observed both in a

139 binary system (Burmann et al., 2010) and a λN-dependent antitermination complex

140 (Krupp et al., 2019; Said et al., 2017).

141 Since the NusG-CTD:S10 interaction is a prerequisite for NusG-mediated

142 transcription:translation coupling, we probed this contact when NusG was bound to

143 RNAP - but not in an antitermination context - by solution-state NMR spectroscopy. We

144 employed NusG samples where [^{1}H,^{13}C]-labeled methyl groups of Ile, Leu, and Val

145 residues in perdeuterated proteins served as NMR-active probes ([ILV]-NusG) to

146 increase sensitivity, allowing us to study large systems.

147 In the methyl-transverse relaxation optimized spectroscopy (methyl-TROSY)

148 spectrum of free [ILV]-NusG (Fig. 2A), signals of the NusG-NTD and NusG-CTD

149 perfectly superimpose with the signals of the isolated [ILV]-labeled protein domains,

150 suggesting that the domains move independently, confirming a previous report stating

151 that there are no intramolecular domain interactions (Burmann et al., 2011). Upon

152 addition of RNAP in a two-fold molar excess, [ILV]-NusG signals were significantly

7

153   decreased in the one-dimensional methyl-TROSY spectrum (Fig. 2B, inset), indicating

154   [ILV]-NusG:RNAP complex formation. Binding of RNAP increases the molecular mass

155   of [ILV]-NusG dramatically, resulting in enhanced relaxation, which ultimately leads to

156   drastic line broadening and a decrease in signal intensity. Interestingly, the two-

157   dimensional spectra revealed a non-uniform signal decrease (Fig. 2B), which is caused by

158   a combination of several effects. First, there is a general loss of signal intensity due to the

159   increase in molecular mass upon complex formation, as discussed above. Second, upon

160   binding, methyl groups of Ile, Leu, and Val residues located in the binding surface come

161   into close proximity of RNAP protons. Dipole-dipole interactions contribute to relaxation

162   processes so that the signal intensity of these methyl groups is decreased more strongly

163   than that of methyl groups located elsewhere in [ILV]-NusG. Finally, signal intensities

164   may be affected by chemical exchange processes. We analyzed the signal intensity of

165   [ILV]-NusG signals in the presence of RNAP quantitatively by calculating relative signal

166   intensities, i.e. the ratio of the remaining signal intensity of [ILV]-NusG in the presence

167   of RNAP to the signal intensity of free [ILV]-NusG (Figure 2-figure supplement 1).

168   The average relative intensity of NusG-NTD signals was significantly lower than

169   that of the linker or the NusG-CTD, suggesting that NusG-NTD binds to RNAP whereas

170   NusG-CTD remains flexible and moves independently, able to interact with other

171   partners, as indicated by the NusG:TEC structure (Kang et al., 2018). The signal intensity

172   of all Ile, Leu, and Val residues in the RNAP binding site of NusG was completely

173   extinguished, confirming that NusG-NTD binds to RNAP at its known binding site

174   (Drögemüller et al., 2015; Kang et al., 2018; Krupp et al., 2019; Said et al., 2017).

175   To test if NusG-CTD can bind to S10 while being tethered to RNAP *via* NusG-NTD,

176   we titrated the [ILV]-NusG:RNAP complex with S10ᴀ (Fig. 2C). In order to increase

177   stability, we used this S10 variant lacking the ribosome binding loop in complex with

178   NusB (Luo et al., 2008). Chemical shift changes of [ILV]-NusG-CTD signals upon

179   titration of [ILV]-NusG:RNAP with S10ᴀ:NusB were determined (Fig. 2D) and affected

180   residues were mapped onto the three-dimensional structure of NusG-CTD (Fig. 2E).

181   Strongly affected residues are found to be located in β-strands 3 and 4 as well as in the

182   connecting loop, in agreement with the binding site observed in the binary NusG-

183   CTD:S10ᴀ complex (Burmann et al., 2010). The loop between β-strands 1 and 2 is also

184   part of the NusG-CTD:S10ᴀ binding site, but as it does not contain any Ile, Leu, or Val

185   residues, no NMR-active probes are available in this region; nevertheless, affected

186   residues can be found in β-strand 1, directly preceding this loop. This suggests that the

187   CTD:S10ᴀ binding surface in the RNAP:NusG:S10ᴀ:NusB complex is identical to the one

188   determined in the binary system. Importantly, the NusG-NTD signals do not change

189   when S10 is added to the NusG:RNAP complex, indicating that S10 binding does not

190   release the bound RNAP.

191   We conclude that the S10 interaction site of NusG-CTD is accessible in the

192   NusG:RNAP complex and thus can promote ribosome binding and formation of a

193   ribosome:NusG:RNAP complex.

194

195   To look for a ribosome:NusG-RNAP complex, we repeated the experiment using

196   intact 70S ribosomes instead of S10ᴀ:NusB (Figure 3). In a first test, we titrated [ILV]-

197   NusG with 70S ribosomes (Figure 3A). As in the [ILV]-NusG:RNAP experiment, signal

198   intensity of [ILV]-NusG methyl groups was significantly, but not uniformly, decreased.

199   In the presence of a twofold molar excess of ribosomes some NusG-NTD signals

200   remained visible, whereas most NusG-CTD signals were nearly completely extinguished.

201   Quantitative analysis of the [ILV]-NusG methyl group signal intensity in the presence of

202   0.5 equivalents of 70S ribosomes clearly shows that the relative intensity of NusG-CTD

203   signals was in a narrow range < 2 %, whereas the relative intensity of NusG-NTD signals

204   covered values from 0-4 %, and was higher on average (Figure 3B). Relative intensities

205   of zero of NusG-NTD signals can be attributed to the fact that these signals are weak

206   even in free NusG, and can thus not be quantified upon ribosome binding. Due to the

207   flexibility of the linker, signals corresponding to amino acids in this region had the

208   highest relative signal intensities. From these results we conclude that NusG binds to the

209   ribosome via its CTD, in agreement with our cryo-EM structure (Figure 1). Due to the

210   drastic increase in molecular mass we were unable to determine a binding site from these

211   experiments, but nevertheless, the pattern of intensity changes of NusG-CTD signals was

212   similar to that resulting form the titration of RNAP-bound NusG with S10, i.e. the most

213   drastic decrease of signal intensity can be observed for residues 160-170, which are part

214   of β-strands 3 and 4 and the intervening loop. Consequently, we conclude that the

215   ribosome binding site is identical with the binding site for isolated S10.

216   Next, we formed a complex of [ILV]-NusG and RNAP (molar ratio 1:2). The 2D

217   methyl-TROSY spectrum of the complex revealed a decrease of signal intensities typical

218   for NusG binding to RNAP (see Fig. 2C), i.e. primarily NusG-CTD signals remained

219   visible. When we then added one equivalent of 70S ribosomes nearly all [ILV]-NusG

220   signals were diminished (e.g. the signal corresponding to I164, which is in the loop

10

221 responsible for ribosome binding). Strikingly, the spectrum differs from the spectrum of

222 [ILV]-NusG in the presence of 70S ribosome (Fig. 3A). These results can be explained

223 by three scenarios: (i) NusG-NTD is bound to RNAP, NusG-CTD is bound to a

224 ribosome, and the ribosome directly interacts with RNAP, (ii) NusG-NTD is bound to

225 RNAP, NusG-CTD is bound to the ribosome, but the ribosome does not interact with

226 RNAP, (iii) NusG-NTD is bound to RNAP, the ribosome directly interacts with RNAP,

227 and NusG-CTD is free, but is in the vicinity of the ribosome. To exclude the last scenario

228 we repeated the experiment using a NusG variant, NusG^{F165A}, in which F165, essential

229 for ribosome binding (Burmann et al., 2010; Knowlton et al., 2003), is substituted by an

230 Ala. Having ensured that the amino acid substitution does not influence the structure of

231 NusG (Fig. 3-figure supplement 1A) we tested in a control experiment [ILV]-NusG^{F165A}

232 binding to S10$^{\Delta}$. Indeed, we could detect no interaction (Fig. 3-figure supplement 1B,C).

233 When we added 70S ribosomes to a preformed [ILV]-NusG^{F165A}:RNAP complex (molar

234 ratio 2:1), the spectrum corresponding to the [ILV]-NusG^{F165A}:RNAP complex did not

235 significantly change and, in particular, NusG-CTD signals remained visible, suggesting

236 that the ribosome was not bound. However the general decrease in signal intensity

237 indicated a direct RNAP:ribosome interaction. Thus, we conclude that NusG can serve as

238 physical linker between ribosome and RNAP, although it remains elusive if a direct

239 interaction between RNAP and a ribosome occurs in this NusG-coupled complex.

240

241 **Translation promotes NusG attachment to TEC.**

242 Chromatin Immuno Precipitation (ChIP) analysis showed that NusG binds to TEC well

243 after transcription and translation initiation (Mooney et al., 2009).   Thus, we asked

11

244 whether translation was, in fact, required for attachment of NusG to TEC. To approach

245 this question, we examined the effects of translation on NusG-mediated Rho-dependent

246 termination within the *lac* operon (Fig. 4A, Table 1) as NusG recruitment to the TEC is

247 necessary for efficient Rho-dependent termination. Rho-dependent termination occurs

248 within *lacZ* both *in vitro* (Burns and Richardson, 1995) and, upon the introduction of *lacZ*

249 nonsense mutations, *in vivo* (Adhya and Gottesman, 1978; Newton et al., 1965). Polarity

250 was measured using a probe to *lacA*, comparing mRNA levels with or without treatment

251 with the Rho inhibitor bicyclomycin (BCM). Wildtype (WT) cells revealed no detectable

252 termination (Table 1, Fig. 4A-I), which may be attributed to (i) sequestering of NusG-

253 CTD by the ribosome, (ii) binding of the ribosome to the nascent RNA, or (iii) both. In

254 all scenarios, however, the presence of the translating ribosome prevents Rho binding.

255 We interfered with translation initiation by mutating the ribosome-binding site, i.e. the

256 Shine-Dalgarno (SD) sequence (Fig. 4A-II), or translation elongation by introducing six

257 successive rare arginine codons at two different locations in *lacZ* (Fig. 4A III, IV).

258 Introduction of two G to A mutations in the *lacZ* SD sequence prohibits translation

259 initiation of *lacZ* (Fig. 4A-II). *lacA* mRNA measurements gave a read-through of 21%,

260 indicating that Rho-dependent termination occurs, but was inefficient in the absence of

261 translation of *lacZ* mRNA. Introduction of the six in-frame rare arginine residues at the

262 +4 position of *lacZ* (Fig. 4A-III, Table 1) allowed 29% read-through, i.e. Rho-dependent

263 termination is present, but still inefficient if translation of *lacZ* mRNA is interfered with

264 at early elongation. In contrast, introduction of the rare arginine residues 200 nt from the

265 start site of transcription (Fig. 4A-IV, Table 1) resulted in high polarity, yielding < 1 %

266   read-through. As efficient Rho-dependent termination requires NusG our results suggests

267   that NusG binding to TEC occurs late and is dependent on translation.

268   To confirm the hypothesis that NusG failed to attach to TEC in the absence of

269   translation, we asked if a complex comprising Nus factors A, B, and E (Nus complex)

270   assembled at a λ *nutL* site was able to recruit NusG so that it associates with TEC.

271   Accordingly, we introduced the λ *nutL* site just upstream of the flawed *lacZ* SD sequence

272   and measured *lacA* mRNA (Figure 4B, Table 1). Indeed, Rho-dependent termination was

273   highly efficient, indicating that NusG had been recruited to TEC. Thus,

274   counterintuitively, the Nus complex, which normally suppresses transcription termination

275   in ribosomal (*rrn*) operons and, together with λN, on the phage λ chromosome,

276   stimulates termination in this case.

277   We finally demonstrated that reduced termination efficiency in the mutant with

278   the non-functional SD sequence was due to the failure of NusG recruitment to the TEC.

279   In this assay we monitored Rho-dependent termination in a fusion construct that carries

280   λ *cro,* the λ *nutR* site, the Rho-dependent λ *tR1* terminator and a *lacZ* reporter, with *lacZ*

281   expression being heat-inducible (Fig. 5). Termination at the λ *tR1* site is poor when *cro* is

282   translated, as seen with the *cro ms27* fusion (Table 2A, Fig. 5A-I; in the presence of an

283   intact SD sequence we used *cro ms27*, where codon 27 carries a missense mutation so

284   that the resulting protein is non-functional). The 3' end of *cro* is adjacent to the λ *tR1*

285   terminator, limiting the amount of free RNA available for Rho attachment if *cro* mRNA

286   is translated. When λ *cro* carried a SD mutation translation initiation was ablated, but

287   nevertheless there was significant termination at λ *tR1* (Table 2A, Figure 5A-II, compare

288   read-through values with and without BCM). Formation of the Nus complex at λ *nutR*

289   allows NusG recruitment and efficient termination. In the absence of NusB, the complex

290   does not assemble, and there is extensive read-through at λ *tR1*.

291   The *boxA69* mutation also reduces Nus complex formation at λ *nutR*, and like the

292   *nusB-* mutation, enhances read-through of λ *tR1* (Table 2B, Fig. 5B). In this experiment,

293   we suppressed termination at λ *tR1* with λ N antitermination factor instead of BCM.

294   Finally, we showed that expression of *nusG-NTD*, which competes with NusG for

295   binding to RNAP, enhances read-through (Table 2C, Fig. 5C). Taken together, these

296   results strongly support the idea that NusG can be supplied by the Nus complex

297   assembled at λ *nutR* in the absence of translation, inducing Rho-dependent termination at

298   λ *tR1*.

299

Discussion

300

301 We determined a cryo-EM structure of a NusG: 70S complex showing binding of one

302 molecule NusG per ribosome, consistent with previous results (Saxena et al., 2018).

303 NusG binds to the S10 protein on the 30S subunit via its CTD as indicated by the study of

304 isolated NusG-CTD and S10 (Burmann et al., 2010); density for NusG-NTD was not

305 observable, suggesting that it remains flexible. We must attribute the low occupancy of

306 the NusG CTD on the 70S ribosome in the cryo-EM experiment to weak binding

307 adversely affected by the conditions of sample preparation. Notably, although tmRNA

308 contacts the ribosome at various sites, the binding of NusG-CTD and tmRNA on S10 is

309 mutually exclusive. This suggests a model in which uncoupling at rare codons, at which

310 tmRNA releases ribosomes, is promoted by tmRNA-induced release of NusG (Roche and

311 Sauer, 1999). The freed NusG:TEC complex exposes the NusG-CTD, and is then subject

312 to Rho-dependent transcription termination.

313 Simultaneous binding of NusG to S10 and RNAP has been demonstrated by

314 solution-state NMR studies, confirming the S10 binding site on NusG-CTD as identified

315 in a binary NusG-CTD:S10 system (Fig. 2) (Burmann et al., 2010). Moreover, we show

316 that NusG can bind RNAP and 70S ribosome simultaneously; this is the first direct

317 structural evidence for NusG-mediated transcription:translation coupling. The flexibility

318 of the linker between the NusG-NTD and the NusG-CTD permits these interactions.

319 The operon-specific *E. coli* NusG paralog, RfaH, likewise simultaneously binds S10 and

320 RNAP in the context of a paused TEC (Burmann et al., 2012; Zuber et al., 2019). RfaH,

321 which also comprises an NTD and a flexibly-connected CTD (Belogurov et al., 2007;

322 Burmann et al., 2012), uses the same binding sites as NusG to interact with RNAP and

15

323   S10 (Burmann et al., 2012, 2010; Kang et al., 2018; Sevostyanova et al., 2011; Zuber et

324   al., 2019). However, RfaH, unlike NusG, complexes with TEC early after transcription

325   initiation, when TEC pauses at an operon polarity suppressor (*ops*) site, a representative

326   of the *E. coli* consensus pause sequence (Larson et al., 2014; Vvedenskaya et al., 2014).

327   Located in the untranslated leader region of RfaH-controlled operons, *ops* is responsible

328   for sequence-specific recruitment of RfaH (Zuber et al., 2018). Importantly, RfaH-

329   dependent operons lack a SD sequence. To initiate translation, RfaH recruits a ribosome

330   to these mRNAs, making coupling essential for translation activation and efficient gene

331   expression. The binding modes of RfaH and NusG to RNAP and S10 are very similar,

332   indicating that coupling as observed for RfaH can also be mediated by NusG and vice

333   versa. However, once recruited, RfaH excludes NusG, thus preventing intra-operon Rho-

334   dependent transcription termination in RfaH-controlled operons (see (Artsimovitch and

335   Knauer, 2019)).

336   We have confirmed the results of Mooney et al that NusG binds to TEC only after

337   significant RNA synthesis (Mooney et al., 2009). As postulated by these authors, binding

338   depends on active translation of the mRNA. Thus efficient Rho-dependent transcription

339   termination, which requires the attachment of Rho to the NusG-CTD, does not occur at

340   the end of an untranslated gene. We have shown that the failure of NusG to bind TEC is

341   responsible for the absence of termination. Thus, placing a λ *nut* site at the start of the

342   gene recruits NusG and restores termination. At present, it is not understood why NusG

343   appears to be delivered to TEC by ribosomes *in vivo*, whereas it binds directly to RNAP

344   in a purified system lacking ribosomes. A possible explanation would be that NusG binds

345   to RNAP discontinuously in an on-and-off mode in the untranslated leader region and

346   that the NusG:RNAP interaction is only stabilized when the ribosome is coupled upon

347   translation initiation. We should recall that NusG has two binding sites in the coupled

348   system, which significantly increases its affinity.

349   In 1970, Miller *et al.* performed electron miscroscopy analyses of lysed *E. coli*

350   cells (Miller et al., 1970). They demonstrated that all mRNA molecules are connected to

351   the *E. coli* genome, and that the ribosome at the newly synthesized end of a polyribosome

352   is almost always immediately adjacent to the putative RNAP molecule. They concluded

353   that translation is completely coupled with transcription. Since NusG attaches to the TEC

354   downstream to the translation initiation site, the elongating complex must initially consist

355   of a ribosome bound directly to TEC, in agreement with the cryo-EM structures of an

356   expressome (Kohler et al., 2017) and an RNAP:30S complex (Demo et al., 2017), as well

357   as *in vitro* data (Fan et al., 2017). At some distance downstream, NusG recognizes and

358   enters the complex. A cryo-EM analysis of a ribosome:RNAP complex reveals a structure

359   with sufficient flexibility to accept NusG between the ribosome and RNAP (Demo et al.,

360   2017). These authors suggest that NusG could enter the complex late, in agreement with

361   our data. Thus we hypothesize that two coupling modes exist, a direct coupling between

362   the ribosome and TEC during translation initiation and early elongation and a NusG-

363   mediated coupling mode later in translation. The question of whether the 70S ribosome

364   still directly contacts TEC in the NusG-mediated system remains elusive. The Kohler

365   structure (Kohler et al., 2017) does not allow NusG binding to TEC and the 70S

366   ribosome, thus the relative orientation of 70S ribosome to TEC might be different in the

367   direct and the NusG-mediated system. The latter may thus require a reorientation of the

368   TEC and the 70S ribosome and confer the system more flexibility, necessary to keep

17

369   transcription and translation synchronized, even if these processes are differently

370   regulated or occur at different rates.

371

18

372   **Methods**

373

374   **Strain Construction.** Standard bacteriological techniques used in strain construction

375   (e.g., transformation, transduction and media preparation) are as described in Silhavy et

376   al. (Silhavy et al., 1984). Standard molecular biology techniques were as described in

377   Sambrook and Russell (Sambrook and Russel, 2001). N10780 was constructed by P1

378   transduction of *rpoC-his:kanR nusGF165A* from NB885 into MDS42. N11158 was

379   constructed by P1 transduction of Δ*ssrA::camR* from RSW943 into MDS42. N11816

380   was constructed by P1 transduction of Δ*relA::kanR* from RLG847 into N11158.

381   RSW1008 was constructed by P1 transduction of Δ*ssrA::camR* from RSW943 into

382   N4837. RSW1010 was constructed by P1 transduction of *rpoC-his:kanR nusGF165A*

383   from NB885 into N4837. RSW1012 was constructed by P1 transduction of Δ*ssrA::camR*

384   from RSW943 into RSW1010. RSW1175 was constructed by P1 transduction of

385   Δ*relA::kanR* and Δ*spoT::camR* from RLG847 into MDS42. RSW1245 was generated

386   using recombineering (Sharan et al., 2009) to introduce six rare arginine codons (atg-acc-

387   atg-AGG-AGA-CGA-AGG-AGA-CGA-att-acg-gat) into the 5'end of *lacZ* in MDS42

388   changing the amino acid sequence of the aminoterminus from MTMITD to

389   MTMRRRRRRITD with six inefficiently translated arginine codons. RSW1225 was

390   produced using recombineering to introduce two G to A mutations in the ribosome

391   binding site of *lacZ* in MDS42. This resulted in a change from

392   …TTCACACA<u>GG</u>AAACAGCTatgaccatg… to …TTCACACA<u>CC</u>

393   AAACAGCTatgaccatg…inactivating the ribosome binding site. RSW1225 is *lac⁻*.

394

395   **Cloning.** The plasmid encoding NusG-F165A was generated by site-directed

396   mutagenesis according to the QuikChange Site-Directed Mutagenesis Kit (Stratagene),

397   using vector pET11A_nusG (Burmann et al., 2011) as template and primers Fw_NusG-

398   F165A (5' GTG TCT GTT TCT ATC GCG GGT CGT GCG ACC CCG 3') and

399   Rv_NusG-F165A (5' CGG GGT CGC ACG ACC CGC GAT AGA AAC AGA CAC 3';

400   both primers were obtained from metabion, Martinsried, Germany).

401

402   **Protein production and isotopic labeling.** NusG and NusG-NTD were produced as

403   described (Burmann et al., 2011) as were NusG-CTD (Burmann et al., 2010) and RNAP

404   and S10$_\Delta$:NusB used for NMR spectroscopy (Zuber et al., 2019). Production of NusG-

405   F165A was analogous to NusG (Burmann et al., 2011). For unlabeled proteins, bacteria

406   were grown in lysogeny broth (LB) medium. [^{1}H,^{13}C]-labeling of methyl groups of Ile,

407   Leu, and Val residues in perdeuterated proteins was accomplished by growing bacteria in

408   M9 medium (Meyer and Schlegel, 1983; Sambrook and Russel, 2001) prepared with

409   increasing amounts of D$_2$O (0 % (v/v), 50 % (v/v), 100 % (v/v); Eurisotop, Saint-Aubin,

410   France) and (^{15}NH$_4$)$_2$SO$_4$ (CortecNet, Voisins-Le-Bretonneux, France) and d$_7$-glucose

411   (Cambridge Isotope Laboratories, Inc., Tewksburgy, USA) as sole nitrogen and carbon

412   sources, respectively. Amino acid precursors (60 mg/l 2-keto-3-d$_3$-4-^{13}C-butyrate and 100

413   mg/l 2-keto-3-methyl-d$_3$-3-d$_1$-4-^{13}C-butyrate; Eurisotop, Saint-Aubin, France) were

414   added 1 h prior to induction. Expression and purification protocols were identical to those

415   of the non-labeled proteins.

416   Intact 70S ribosomes were produced as follows. *E. coli* strain MRE600 cells grown in LB

417   medium were harvested, lysed by passing through the French Press 3x at ~800 PSI, and

418   clarified by a short centrifugation (20,000 rpm, 40min) in opening buffer (20 mM Tris-

419   HCl pH=7.5, 100mM NH_4Cl, 10.5 mM Mg acetate, 0.5 mM EDTA, with half a protease

420   inhibitor cocktail tablet (Roche, EDTA-free), and 1mM TCEP added just before use). The

421   lysate was loaded onto the top of 5 mL sucrose cushion (20 mM Tris-HCl, pH=7.5, 500

422   mM NH_4Cl, 10.5 mM Mg acetate, 0.5 mM EDTA, 1.1 M sucrose, and 1 mM TCEP

423   added before use) and centrifuged for 24 h at 28,000 rpm in a 70Ti rotor (Beckman

424   Coulter, Inc.). The pellets were suspended in washing buffer (20 mM Tris-HCl, pH=7.5;

425   500mM NH_4Cl, 10.5mM Mg acetate, 0.5mM EDTA and 1 mM TCEP added before use),

426   and centrifuged through a 10–35% sucrose gradient for 19 h at 16,000 rpm in a SW28

427   rotor (Beckman Coulter, Inc.). Fractions containing the 70S peaks were pooled and kept

428   at -80℃ for further use.

429   Ribosomes for NMR experiments were obtained from New England Biolabs.

430

431   **Electron Microscopy**

432   Purified 70S ribosomes were incubated with full-length NusG at a ratio of 1:7 for 40 min

433   at room temperature, prior to blotting and plunge-freezing as previously described

434   (Grassucci et al., 2007). Data were collected on a TF30 Polara electron microscope (FEI,

435   Portland, Oregon) at 300kV using a K2 Summit direct electron detector camera (Gatan,

436   Pleasanton, CA). Images were recorded using the automated data collection system

437   Leginon (Suloway et al., 2005) in counting mode, and taken at the nominal magnification

438   of 32,000x, corresponding to a calibrated pixel size of 1.66Å.

Image processing

439

440 A total of 188,127 particles were automatically extracted from 1327 images using

441 Arachnid (Langlois et al., 2014). RELION (Scheres, 2012) 3D classification was used to

442 resolve the heterogeneity of the particle images, and auto-refinement to further improve

443 resolution for each class. The final refinement for the NusG-bound 70S class containing

444 17,122 particles yielded an average resolution of ~6.8Å (FSC=0.143; following "gold

445 standard" protocol).

446

Model building

447

448 The starting model was assembled from the X-ray structure of the *E. coli* 30S ribosomal

449 subunit (PDB ID 4GD2) and the NMR solution structure of the NusG-CTD (PDB 2KVQ

450 chain G). This starting model was first docked into the segmented maps of our 70S

451 density map as a rigid body using UCSF Chimera (Pettersen et al., 2004). Then it was

452 fitted into the segmented map using the Molecular Dynamic Flexibly Fitting (MDFF)

453 method (Trabuco et al., 2008) and run using the NAMD program (Phillips et al., 2005) for

454 0.5 ns of simulation time, followed by 5,000 steps of energy minimization.

455

456 **NMR spectroscopy.** NMR experiments were conducted on Bruker Ascend Aeon 900 and

457 1000 MHz spectrometers equipped with cryogenically cooled, inverse triple resonance

458 probes at 298 K. NMR data was converted and processed using in-house software. 2D

459 correlation spectra were visualized and analyzed with NMRViewJ (One Moon Scientific,

460 Inc., Westfield, NJ, USA), 1D spectra were plotted using MatLab (The MathWorks, Inc.,

461   Version 9.2.0.538062). Resonance assignments for NusG methyl groups were taken from

462   a previous study (Mooney et al., 2009).

463   [ILV]-NusG-CTD was in 10 mM K-phosphate (pH 7.5), 50 mM KCl, 1 mM MgCl$_2$, 99.9

464   % (v/v) D$_2$O, [ILV]-NusG-NTD in 50 mM Na-phosphate (pH 7.5), 50 mM KCl, 0.3 mM

465   EDTA, 5 % (v/v) d$_7$-glycerol, 0.01 % (w/v) NaN$_3$, 99.9 % D$_2$O. For the titration of [ILV]-

466   NusG with RNAP and S10^D:NusB, all proteins were in 50 mM Na-phosphate (pH 7.5),

467   50 mM KCl, 0.3 mM EDTA, 99.9 % (v/v) D$_2$O and 5 mM MgCl$_2$ and 2 mM DTT were

468   added to the NMR sample to increase the-long-term stability of RNAP. For all interaction

469   studies involving ribosomes and for the titration of [ILV]-NusG-F165A with S10^A:NusB,

470   all components were in 20 mM HEPES-KOH (pH 7.6), 10 mM Mg-acetate, 30 mM KCl,

471   7 mM β-mercaptoethanol, 10 % D$_2$O. The titration of [ILV]-NusG-F165A with

472   S10^A:NusB was conducted in a 5 mm tube with an initial sample volume of 550 µl. All

473   other measurements were carried out in 3 mm NMR tubes with an (initial) volume of 200

474   µl.

475   1D and 2D titration experiments were evaluated quantitatively by analyzing either

476   changes in signal intensity or changes in chemical shifts. If chemical shift changes were

477   in the fast regime on the chemical shift the normalized chemical shift perturbation

478   ($\Delta\delta_{norm}$) was calculated according to equation 1.

479

480
$$\Delta\delta_{norm} = \sqrt{\left(\Delta\delta\ ^1H\right)^2 + \left[0.25\left(\Delta\delta^{13}C\right)\right]^2} \quad (1)$$

481   with $\Delta\delta$ being the resonance frequency difference between the initial and final state of the

482   titration (i.e. [ILV]-NusG:RNAP:S10^D:NusB = 1:2:0:0 *vs*. 1:2:2:2) in ppm.

483

484   If the system was in slow or intermediate chemical exchange the signal intensities were

485   analyzed quantitatively. First, the intensity of each 1D spectrum or methyl group signal,

486   respectively, was normalized by the concentration of the [ILV]-labeled protein, the

487   receiver gain, the number of scans, and the length of the 90° ^{1}H pulse. Then the relative

488   intensity, i.e. the ratio of the normalized signal intensity of [ILV]-labeled protein in the

489   respective titration step to the normalized signal intensity of free [ILV]-labeled protein,

490   was calculated and plotted against the sequence of NusG or the NusG variant,

491   respectively.

492

493   **qRT-PCR.** Total RNA was extracted from cells grown in M9 medium supplemented

494   with casamino acids (0.2%) at 37^{0}C to mid-log phase (O.D. 600=0.3). Fold-increase of

495   the PCR product was determined using qRT-PCR. cDNA was synthesized using RNA

496   was extracted from logarithmically growing cultures (O.D.$_{600}$=0.2-0.3) Where indicated,

497   cells were treated with BCM (100μg/ml) 1 min before induction with 1mM IPTG for

498   *lacZ.* Samples were removed (0.5ml) at the indicated times and total RNA extracted

499   RNA extracted using Qiagen RNeasy and Qiagen RNAprotect Bacteria Reagent (Qiagen,

500   Germantown, MD). cDNA was synthesized from the samples using High Capacity RNA

501   to cDNA kit (ThermoFisher, Waltham, MA). qRT-PCR reactions were performed using

502   Taqman Gene Expression Master Mix (Thermofisher, Waltham, MA) and Biorad DNA

503   Engine Opticon2 Real-Time Cycler (Bio-Rad Laboratories, Hercules, CA) and

504   PrimeTime qPCR probes (Integrated DNA Technologies, Coralville, IA). The *lacA*

505   transcript was probed with the following: probe:5'-/56-

506   FAM/CCACATGAC/ZEN/TTCCGATCCAGACGTT/3IABkFQ/-3'; primer1:5'-

24

507   ATACTACCCGCGCCAATAAC; primer2:5'-CCCTGTACACCATGAATTGAGA).

508   The reference gene was *ompA*(probe:5'-/56-

509   FAM/CAACAACAT/ZEN/CGGTGACGCACACAC /3IABkFQ/-3'; primer1:5'-

510   TGACCGAAACGGTAGGAAAC; primer2:5'-ACGCGATCACTCCTGAAATC). The

511   PCR was performed using the following conditions: 50 $^{\circ}$C for 10 min., 95 $^{\circ}$C for 2 min,

512   followed by 40 cycles each of 95 °C for 15s, and 60 °C 1min. All reactions were

513   performed in duplicate. Fold increases were calculated from measured C_t values using

514   the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001). Values are the average of three or more

515   independent experiments.

516

517   ***β–galactosidase assays.*** Cultures were grown in LB to early log phase (O.D. 600 =0.3)

518   at 37 ^{0}C. Where indicated bicyclomycin (BCM) (100μg/ml) was added to inhibit Rho-

519   dependent transcription termination prior to induction of *lacZ* with 1mm IPTG.   Where

520   indicated λN was expressed by incubation at 42 $^{\circ}$C. Reactions were terminated 15 min.

521   after induction. β-galactosidase was measured using a modification of the method of

522   Miller (Zhang and Bremer, 1995). Readthrough was calculated from the ratio of β-

523   galactosidase activity +/- BCM.

524

525   **Author Contributions:** R.W., Y.H., M.S., W.L., B.S. P.K.Z. and S.H. performed the

526   experiments. J.F., Y.H., M.S., R.W. P.K.Z. S.H.K. and M.G. designed the experiments

527   and interpreted the results. M.S. M.G. M.S. P.K.Z. S.H.K. and J.F. wrote the paper.

528

534

535   **Competing interests.** We declare no competing interests.

References

Adhya S, Gottesman M. 1978. Control of transcription termination. *Annu Rev Biochem* **47**:967–996. doi:10.1146/annurev.bi.47.070178.004535

Artsimovitch I, Knauer SH. 2019. Ancient Transcription Factors in the News. *MBio* **10**. doi:10.1128/mBio.01547-18

Belogurov GA, Vassylyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassylyev DG, Artsimovitch I. 2007. Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol Cell* **26**:117–129. doi:10.1016/j.molcel.2007.02.021

Burmann BM, Knauer SH, Sevostyanova A, Schweimer K, Mooney RA, Landick R, Artsimovitch I, Rösch P. 2012. An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* **150**:291–303. doi:10.1016/j.cell.2012.05.042

Burmann BM, Scheckenhofer U, Schweimer K, Rösch P. 2011. Domain interactions of the transcription-translation coupling factor *Escherichia coli* NusG are intermolecular and transient. *Biochem J* **435**:783–789. doi:10.1042/BJ20101679

Burmann BM, Schweimer K, Luo X, Wahl MC, Stitt BL, Gottesman ME, Rösch P. 2010. A NusE:NusG complex links transcription and translation. *Science* **328**:501–504. doi:10.1126/science.1184953

Burns CM, Richardson JP. 1995. NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. *Proc Natl Acad Sci USA* **92**:4738–4742. doi:10.1073/pnas.92.11.4738

Burova E, Hung SC, Sagitov V, Stitt BL, Gottesman ME. 1995. *Escherichia coli* NusG

559   protein stimulates transcription elongation rates in vivo and in vitro. *J Bacteriol*

560   **177**:1388–1392. doi:10.1128/jb.177.5.1388-1392.1995

561   Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, Nudler E. 2008.

562   Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*.

563   *Science* **320**:935–938. doi:10.1126/science.1152763

564   Conn AB, Diggs S, Tam TK, Blaha GM. 2019. Two Old Dogs, One New Trick: A

565   Review of RNA Polymerase and Ribosome Interactions during Transcription-Translation

566   Coupling. *Int J Mol Sci* **20**. doi:10.3390/ijms20102595

567   Das HK, Goldstein A, Lowney LI. 1967. Attachment of ribosomes to nascent messenger

568   RNA in *Escherichia coli*. *J Mol Biol* **24**:231–245. doi:10.1016/0022-2836(67)90329-4

569   Demo G, Rasouly A, Vasilyev N, Svetlov V, Loveland AB, Diaz-Avalos R, Grigorieff N,

570   Nudler E, Korostelev AA. 2017. Structure of RNA polymerase bound to ribosomal 30S

571   subunit. *Elife* **6**. doi:10.7554/eLife.28560

572   Drögemüller J, Strauß M, Schweimer K, Jurk M, Rösch P, Knauer SH. 2015.

573   Determination of RNA polymerase binding surfaces of transcription factors by NMR

574   spectroscopy. *Sci Rep* **5**:16428. doi:10.1038/srep16428

575   Dudenhoeffer BR, Schneider H, Schweimer K, Knauer SH. 2019. SuhB is an integral part

576   of the ribosomal antitermination complex and interacts with NusA. *Nucleic Acids Res*

577   **47**:6504–6518. doi:10.1093/nar/gkz442

578   Dutta D, Shatalin K, Epshtein V, Gottesman ME, Nudler E. 2011. Linking RNA

579   polymerase backtracking to genome instability in *E. coli*. *Cell* **146**:533–543.

580   doi:10.1016/j.cell.2011.07.034

581   Fan H, Conn AB, Williams PB, Diggs S, Hahm J, Gamper HB, Hou Y-M, O'Leary SE,

582 Wang Y, Blaha GM. 2017. Transcription-translation coupling: direct interactions of RNA

583 polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res* **45**:11043–11055.

584 doi:10.1093/nar/gkx719

585 Fu J, Hashem Y, Wower I, Lei J, Liao HY, Zwieb C, Wower J, Frank J. 2010.

586 Visualizing the transfer-messenger RNA as the ribosome resumes translation. *EMBO J*

587 **29**:3819–3825. doi:10.1038/emboj.2010.255

588 Grassucci RA, Taylor DJ, Frank J. 2007. Preparation of macromolecular complexes for

589 cryo-electron microscopy. *Nat Protoc* **2**:3239–3246. doi:10.1038/nprot.2007.452

590 Kang JY, Mooney RA, Nedialkov Y, Saba J, Mishanina TV, Artsimovitch I, Landick R,

591 Darst SA. 2018. Structural Basis for Transcript Elongation Control by NusG Family

592 Universal Regulators. *Cell* **173**:1650-1662.e14. doi:10.1016/j.cell.2018.05.017

593 Knowlton JR, Bubunenko M, Andrykovitch M, Guo W, Routzahn KM, Waugh DS,

594 Court DL, Ji X. 2003. A spring-loaded state of NusG in its functional cycle is suggested

595 by X-ray crystallography and supported by site-directed mutants. *Biochemistry* **42**:2275–

596 2281. doi:10.1021/bi0272508

597 Kohler R, Mooney RA, Mills DJ, Landick R, Cramer P. 2017. Architecture of a

598 transcribing-translating expressome. *Science* **356**:194–197. doi:10.1126/science.aal3059

599 Krupp F, Said N, Huang Y-H, Loll B, Bürger J, Mielke T, Spahn CMT, Wahl MC. 2019.

600 Structural Basis for the Action of an All-Purpose Transcription Anti-termination Factor.

601 *Mol Cell*. doi:10.1016/j.molcel.2019.01.016

602 Kyrpides NC, Woese CR, Ouzounis CA. 1996. KOW: a novel motif linking a bacterial

603 transcription factor with ribosomal proteins. *Trends Biochem Sci* **21**:425–426.

604 Langlois R, Pallesen J, Ash JT, Nam Ho D, Rubinstein JL, Frank J. 2014. Automated

605   particle picking for low-contrast macromolecules in cryo-electron microscopy. *J Struct*

606   *Biol* **186**:1–7. doi:10.1016/j.jsb.2014.03.001

607   Larson MH, Mooney RA, Peters JM, Windgassen T, Nayak D, Gross CA, Block SM,

608   Greenleaf WJ, Landick R, Weissman JS. 2014. A pause sequence enriched at translation

609   start sites drives transcription dynamics *in vivo*. *Science* **344**:1042–1047.

610   doi:10.1126/science.1251871

611   Lawson MR, Ma W, Bellecourt MJ, Artsimovitch I, Martin A, Landick R, Schulten K,

612   Berger JM. 2018. Mechanism for the regulated control of transcription by a universal

613   adapter protein. *Mol Cell* **71**:1–12.

614   Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time

615   quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**:402–408.

616   doi:10.1006/meth.2001.1262

617   Luo X, Hsiao H-H, Bubunenko M, Weber G, Court DL, Gottesman ME, Urlaub H, Wahl

618   MC. 2008. Structural and functional analysis of the *E. coli* NusB-S10 transcription

619   antitermination complex. *Mol Cell* **32**:791–802. doi:10.1016/j.molcel.2008.10.028

620   Mehdi Q, Yudkin MD. 1967. Coupling of transcription to translation in the induced

621   synthesis of beta-galactosidase. *Biochim Biophys Acta* **149**:288–290. doi:10.1016/0005-

622   2787(67)90710-1

623   Meyer O, Schlegel HG. 1983. Biology of aerobic carbon monoxide-oxidizing bacteria.

624   *Annu Rev Microbiol* **37**:277–310. doi:10.1146/annurev.mi.37.100183.001425

625   Miller OL, Hamkalo BA, Thomas CA. 1970. Visualization of bacterial genes in action.

626   *Science* **169**:392–395. doi:10.1126/science.169.3943.392

627   Mitra P, Ghosh G, Hafeezunnisa M, Sen R. 2017. Rho Protein: Roles and Mechanisms.

30

628   *Annu Rev Microbiol* **71**:687–709. doi:10.1146/annurev-micro-030117-020432

629   Mooney Rachel A., Davis SE, Peters JM, Rowland JL, Ansari AZ, Landick R. 2009.

630   Regulator trafficking on bacterial transcription units *in vivo*. *Mol Cell* **33**:97–108.

631   doi:10.1016/j.molcel.2008.12.021

632   Mooney Rachel Anne, Schweimer K, Rösch P, Gottesman M, Landick R. 2009. Two

633   structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct

634   interactions with RNA polymerase and regulators. *J Mol Biol* **391**:341–358.

635   doi:10.1016/j.jmb.2009.05.078

636   Newton WA, Beckwith JR, Zipser D, Brenner S. 1965. Nonsense mutants and polarity in

637   the *lac* operon of *Escherichia coli*. *J Mol Biol* **14**:290–296. doi:10.1016/s0022-

638   2836(65)80250-9

639   Pani B, Nudler E. 2017. Mechanistic insights into transcription coupled DNA repair.

640   *DNA Repair (Amst)* **56**:42–50. doi:10.1016/j.dnarep.2017.06.006

641   Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE.

642   2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J*

643   *Comput Chem* **25**:1605–1612. doi:10.1002/jcc.20084

644   Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD,

645   Kalé L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem*

646   **26**:1781–1802. doi:10.1002/jcc.20289

647   Proshkin S, Rahmouni AR, Mironov A, Nudler E. 2010. Cooperation between translating

648   ribosomes and RNA polymerase in transcription elongation. *Science* **328**:504–508.

649   doi:10.1126/science.1184939

650   Rae CD, Gordiyenko Y, Ramakrishnan V. 2019. How a circularized tmRNA moves

651   through the ribosome. *Science* **363**:740–744. doi:10.1126/science.aav9370

652   Roche ED, Sauer RT. 1999. SsrA-mediated peptide tagging caused by rare codons and

653   tRNA scarcity. *EMBO J* **18**:4579–4589. doi:10.1093/emboj/18.16.4579

654   Said N, Krupp F, Anedchenko E, Santos KF, Dybkov O, Huang Y-H, Lee C-T, Loll B,

655   Behrmann E, Bürger J, Mielke T, Loerke J, Urlaub H, Spahn CMT, Weber G, Wahl MC.

656   2017. Structural basis for λN-dependent processive transcription antitermination. *Nat*

657   *Microbiol* **2**:17062. doi:10.1038/nmicrobiol.2017.62

658   Sambrook J, Russel DW. 2001. Molecular Cloning: A Laboratory Manual, Third Edition.

659   ed. New York: Cold Spring Harbor Press.

660   Saxena S, Myka KK, Washburn R, Costantino N, Court DL, Gottesman ME. 2018.

661   *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol Microbiol*

662   **108**:495–504. doi:10.1111/mmi.13953

663   Scheres SHW. 2012. RELION: implementation of a Bayesian approach to cryo-EM

664   structure determination. *J Struct Biol* **180**:519–530. doi:10.1016/j.jsb.2012.09.006

665   Sevostyanova A, Belogurov GA, Mooney RA, Landick R, Artsimovitch I. 2011. The β

666   subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Mol*

667   *Cell* **43**:253–262. doi:10.1016/j.molcel.2011.05.026

668   Sharan SK, Thomason LC, Kuznetsov SG, Court DL. 2009. Recombineering: a

669   homologous recombination-based method of genetic engineering. *Nat Protoc* **4**:206–223.

670   doi:10.1038/nprot.2008.227

671   Silhavy TJ, Berman ML, Enquist LW, editors. 1984. Experiments with Gene Fusions.

672   Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

673   Squires CL, Greenblatt J, Li J, Condon C, Squires CL. 1993. Ribosomal RNA

674   antitermination in vitro: requirement for Nus factors and one or more unidentified cellular

675   components. *Proc Natl Acad Sci USA* **90**:970–974.

676   Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, Stagg S, Potter CS,

677   Carragher B. 2005. Automated molecular microscopy: the new Leginon system. *J Struct*

678   *Biol* **151**:41–60. doi:10.1016/j.jsb.2005.03.010

679   Thommen M, Holtkamp W, Rodnina MV. 2017. Co-translational protein folding:

680   progress and methods. *Curr Opin Struct Biol* **42**:83–89. doi:10.1016/j.sbi.2016.11.020

681   Trabuco LG, Villa E, Mitra K, Frank J, Schulten K. 2008. Flexible fitting of atomic

682   structures into electron microscopy maps using molecular dynamics. *Structure* **16**:673–

683   683. doi:10.1016/j.str.2008.03.005

684   Vogel U, Jensen KF. 1995. Effects of the antiterminator *BoxA* on transcription elongation

685   kinetics and ppGpp inhibition of transcription elongation in *Escherichia col*i. *J Biol Chem*

686   **270**:18335–18340. doi:10.1074/jbc.270.31.18335

687   Vogel U, Jensen KF. 1994. Effects of guanosine 3',5'-bisdiphosphate (ppGpp) on rate of

688   transcription elongation in isoleucine-starved *Escherichia coli*. *J Biol Chem* **269**:16236–

689   16241.

690   Vvedenskaya IO, Vahedian-Movahed H, Bird JG, Knoblauch JG, Goldman SR, Zhang Y,

691   Ebright RH, Nickels BE. 2014. Interactions between RNA polymerase and the "core

692   recognition element" counteract pausing. *Science* **344**:1285–1289.

693   doi:10.1126/science.1253458

694   Washburn RS, Gottesman ME. 2011. Transcription termination maintains chromosome

695   integrity. *Proc Natl Acad Sci USA* **108**:792–797. doi:10.1073/pnas.1009564108

696   Weis F, Bron P, Giudice E, Rolland J-P, Thomas D, Felden B, Gillet R. 2010. tmRNA-

697   SmpB: a journey to the centre of the bacterial ribosome. *EMBO J* **29**:3810–3818.

698   doi:10.1038/emboj.2010.252

699   Werner F. 2012. A nexus for gene expression-molecular mechanisms of Spt5 and NusG

700   in the three domains of life. *J Mol Biol* **417**:13–27. doi:10.1016/j.jmb.2012.01.031

701   Zhang X, Bremer H. 1995. Control of the Escherichia coli rrnB P1 promoter strength by

702   ppGpp. *J Biol Chem* **270**:11181–11189. doi:10.1074/jbc.270.19.11181

703   Zuber PK, Artsimovitch I, NandyMazumdar M, Liu Z, Nedialkov Y, Schweimer K,

704   Rösch P, Knauer SH. 2018. The universally-conserved transcription factor RfaH is

705   recruited to a hairpin structure of the non-template DNA strand. *Elife* **7**.

706   doi:10.7554/eLife.36349

707   Zuber PK, Schweimer K, Rösch P, Artsimovitch I, Knauer SH. 2019. Reversible fold-

708   switching controls the functional cycle of the antitermination factor RfaH. *Nat Commun*

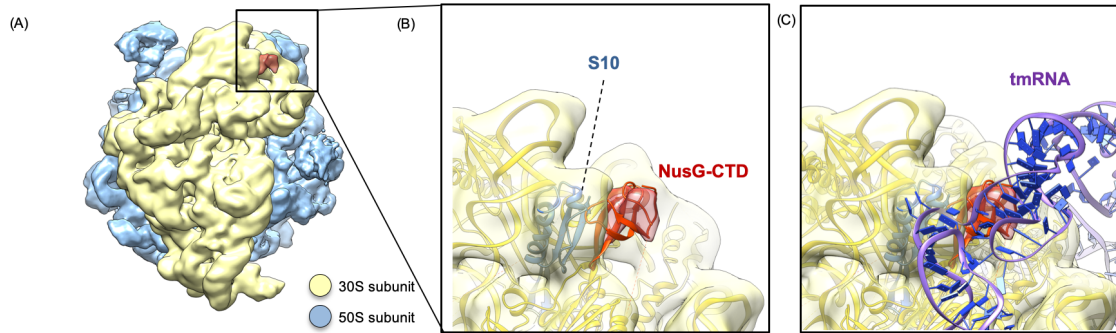709   **10**:702. doi:10.1038/s41467-019-08567-6

710

711

712

34

713   **Figures**

714

715

716

717   **Figure 1. Structure of NusG-CTD bound to 70S ribosome. (A)** Cryo-EM density of

718   the 70S ribosome:NusG complex. The density of the 50S subunit is shown in light blue,

719   the density of the 30S subunit in yellow, the density corresponding to NusG-CTD in red.

720   **(B)** Close-up view of the region boxed in **(A)**. 70S (yellow), S10 (blue), and NusG-CTD

721   (red) are in ribbon representation, cryo-EM density is shown as transparencies.   **(C)**

722   Superposition of the 70S:NusG complex with the 70S:tmRNA complex (tmRNA is in

723   ribbon representation, purple and dark blue; EMD 5234, PDB: 3IZ4). 30S and NusG-

724   CTD are displayed as in **(B)**.

725

726

35
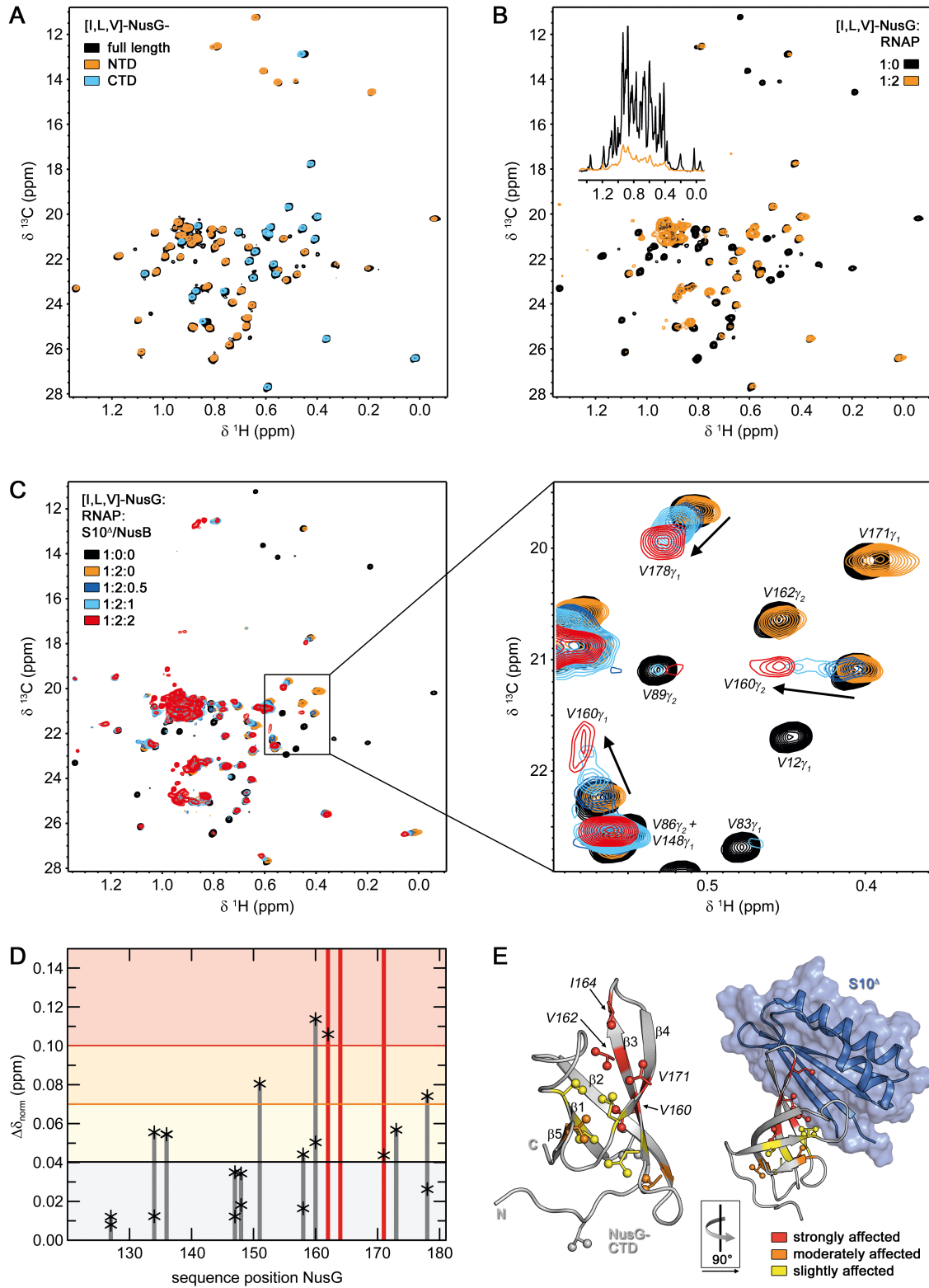
Figure 2. RNAP-bound NusG interacts with S10. (A) Superposition of 2D [^{1}H, ^{13}C]-methyl-TROSY spectra of [ILV]-NusG (black, 20 µM), [ILV]-NusG-NTD (orange, 100

727

728

729

36

730   μM), and [ILV]-NusG-CTD (cyan, 30 μM). **(B)** 2D [^{1}H, ^{13}C]-methyl-TROSY spectra of

731   [ILV]-NusG in the absence (black, 20 μM) and presence (orange, 18 μM) of two

732   equivalents of RNAP. Inset: Normalized 1D [^{1}H,^{13}C]-methyl TROSY spectra, colored as

733   2D spectra. **(C)** 2D [^{1}H, ^{13}C]-methyl-TROSY spectra of [ILV]-NusG alone (20 μM), in

734   the presence of a twofold molar excess of RNAP (18 μM [ILV]-NusG), and upon

735   titration of [ILV]-NusG:RNAP with 218 μM S10^Δ:NusB. The molar ratio of [ILV]-

736   NusG:RNAP:S10^Δ:NusB is indicated in color. The panel on the right shows an

737   enlargement of the boxed region. Selected signals are labeled and arrows indicate

738   chemical shift changes upon S10^Δ:NusB addition. **(D)** [^{1}H,^{13}C]-methyl-TROSY-derived

739   normalized chemical shift changes of [ILV]-NusG-CTD methyl group signals of RNAP-

740   bound [ILV]-NusG upon complex formation with S10^Δ:NusB. Asterisks mark the values

741   of individual methyl group signals, bars represent the highest values. Red bars indicate

742   vanishing signals. Horizontal lines are thresholds for affected methyl groups: slightly

743   affected (0.04 ppm ≤ $\Delta\delta_{norm}$ < 0.07 ppm; black), moderately affected (0.07 ppm ≤ $\Delta\delta_{norm}$

744   < 0.1 ppm; orange), and strongly affected ($\Delta\delta_{norm}$ ≥ 0.10 ppm; red). **(E)** Mapping of

745   affected methyl groups on the structure of isolated NusG-CTD (left; PDB ID: 2JVV) and

746   NusG-CTD in complex with S10^Δ (right; PDB ID 2KVQ). NusG-CTD is shown in ribbon

747   (gray), S10^Δ in ribbon and surface (blue) representation. Affected Ile, Leu, and Val

748   residues are colored according to **(D),** non-affected Ile, Leu, and Val residues are gray.

749   Side chains of Ile, Leu, and Val residues are depicted as sticks, their methyl groups as

750   spheres. Strongly affected Ile, Leu, and Val residues are labeled. The orientation of

751   NusG-CTD in the complex relative to the isolated state is indicated.

752

753
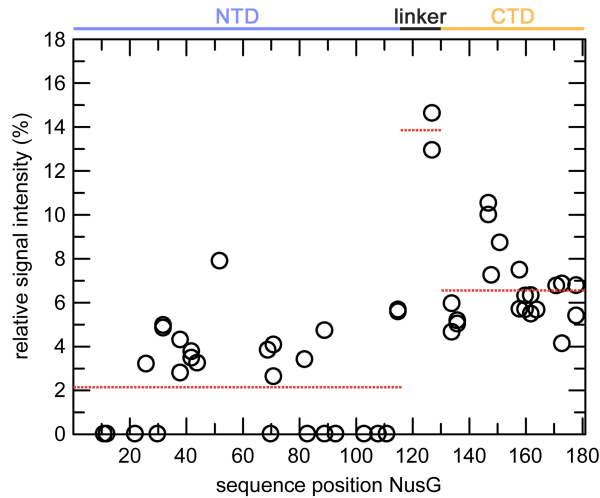
**Figure 2-figure supplement 1: Binding of [ILV]-NusG to RNAP.** [^{1}H,^{13}C]-methyl-

TROSY derived relative signal intensities of [ILV]-NusG methyl groups after addition of

two equivalents of RNAP (see Fig. 2B). Red, dashed horizontal lines indicate average

relative signal intensities of NusG-NTD, the linker, and NusG-CTD (domain organization

is indicated at the top). Related to Figure 2B.

759

760
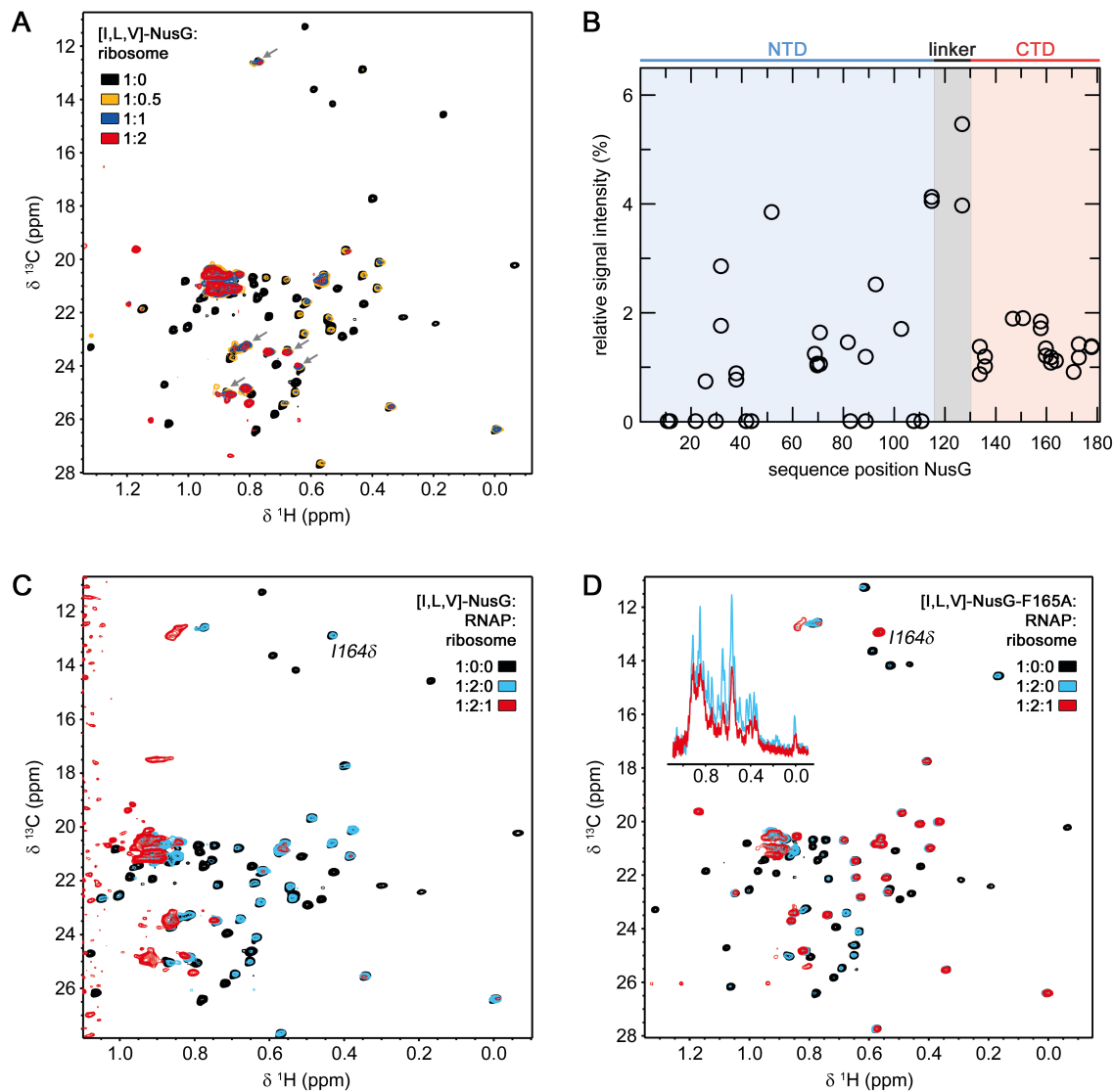
Figure 3: RNAP-bound NusG interacts with the 70 S ribosome. (A,B) NusG interacts with 70S ribosome via its CTD. **(A)** 2D [^{1}H, ^{13}C]-methyl-TROSY spectra of free [ILV]-NusG (11 µM, black) and [ILV]-NusG in the presence of 70S ribosome (molar ratio [ILV]-NusG:ribosome = 1:0.5 (6.6 µM [ILV]-NusG, orange); = 1:1 (7.5 µM [ILV]-NusG, blue); =1:2 (4 µM [ILV]-NusG, red). Arrows indicate [ILV]-NusG-NTD signals that are well visible in the [ILV]-NusG:ribosome complex. **(B)** Quantitative analysis of [ILV]-NusG methyl group signal intensities in the presence of 0.5 equivalents of 70S ribosome. Relative signal intensities are plotted versus the sequence position of NusG.

39

769   The domain organization of NusG is indicated above the diagram. **(C)** 2D [^{1}H, ^{13}C]-

770   methyl-TROSY spectra of [ILV]-NusG (11 µM, black), [ILV]-NusG in the presence of

771   RNAP (molar ratio 1:2, 6 µM [ILV]-NusG, blue), and [ILV]-NusG in the presence of

772   RNAP and 70S ribosome (molar ratio 1:2:1, 6 µM [ILV]-NusG, red). **(D)** 2D [^{1}H, ^{13}C]-

773   methyl-TROSY spectra of [ILV]-NusG^{F165A} (20 µM, black), [ILV]-NusG^{F165A} in the

774   presence of RNAP (molar ratio 1:2, 6 µM [ILV]-NusG^{F165A}, blue), and [ILV]-NusG^{F165A}

775   in the presence of RNAP and 70S ribosome (molar ratio 1:2:1, 6 µM [ILV]-NusG^{F165A},

776   red). The inset shows the normalized 1D spectra of the corresponding titration step.

777
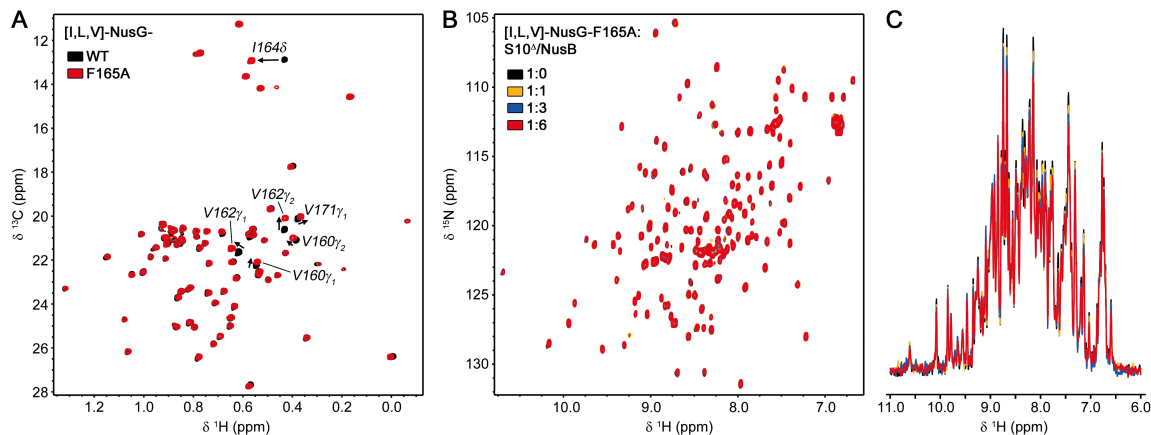
Figure 3-figure supplement 1: NusG-F165A does not interact with S10ᐞ/NusB. (A) 2D [^{1}H, ^{13}C]-methyl-TROSY spectra of [ILV]-NusG (11 µM, black) and [ILV]-NusG^{F165A} (20 µM, red). Arrows and labels indicate NusG-CTD methyl groups affected in their resonance frequencies by the F165A amino acid substitution. **(B,C)** 2D **(B)** and normalized 1D **(C)** [^{1}H, ^{15}N]-HSQC spectra of 20 µM [ILV]-NusG^{F165A} upon titration with 432 µM S10ᐞ/NusB (colors as indicated).
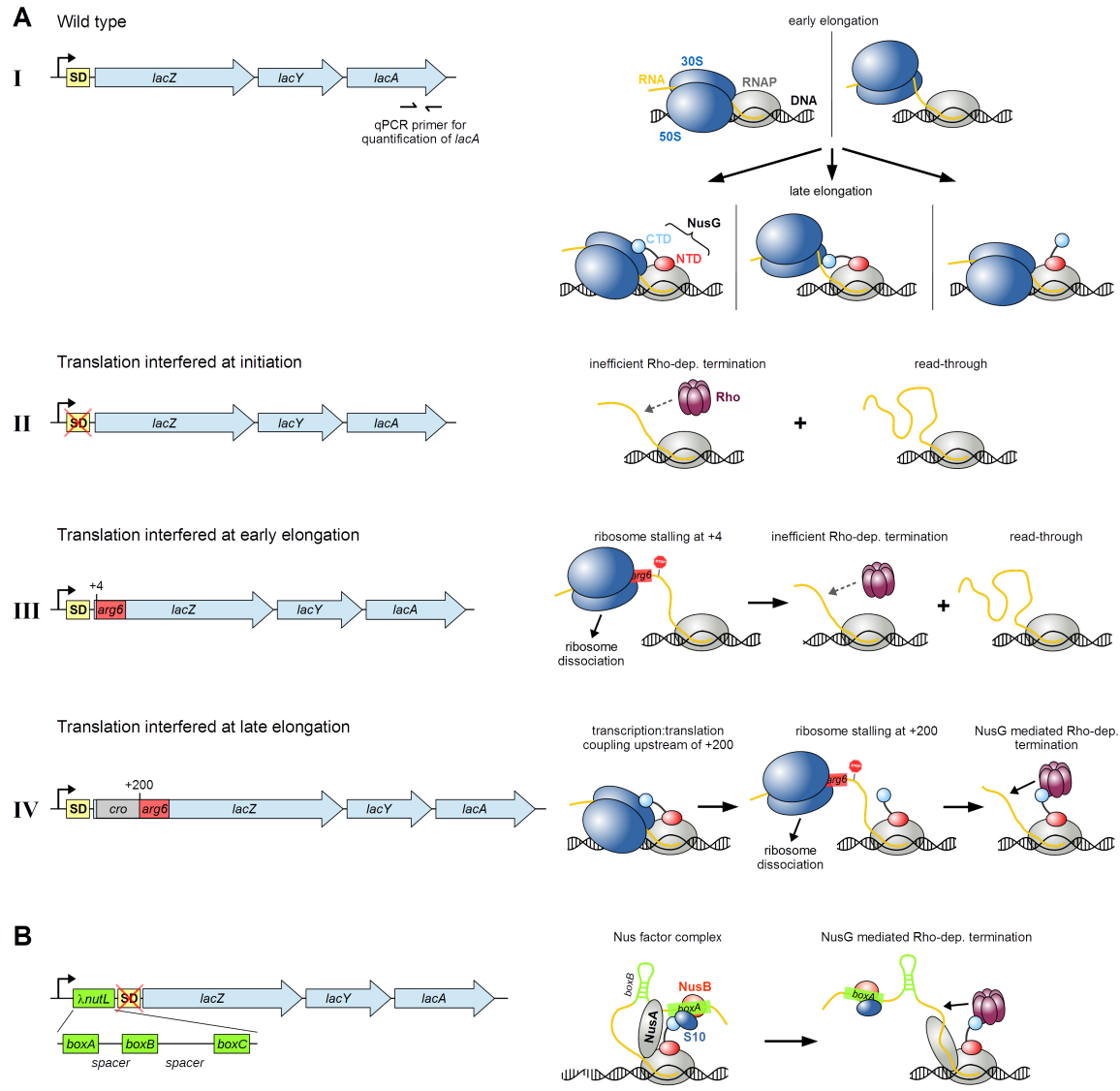
Figure 4: Translation is required for NusG recruitment to the TEC. (A, B) Left:

Organization of the *E. coli lac* operon in strains MDS42 (A-I; wild type *lacZ*), RSW1225

(A-II; mutant (inactive) *lacZ* SD sequence), RSW1245 (A-III; in-frame insertion of six

rare Arg codons (*arg6*) at position +4 of *lacZ*), RSW1276 (A-IV; in-frame insertion of

λ*cro* and six rare Arg codons at position +4 of *lacZ* (equivalent to *arg6* being at position

+200 of the gene)), and RSW1297 (B; λ*nutL* site upstream of mutant *lacZ* SD sequence).

SD sequences of *lacY* and *lacA* were omitted for clarity. qPCR primers specific to the 3'

end of *lacA* (position indicated in A-I) were used to measure mRNA levels and thereby

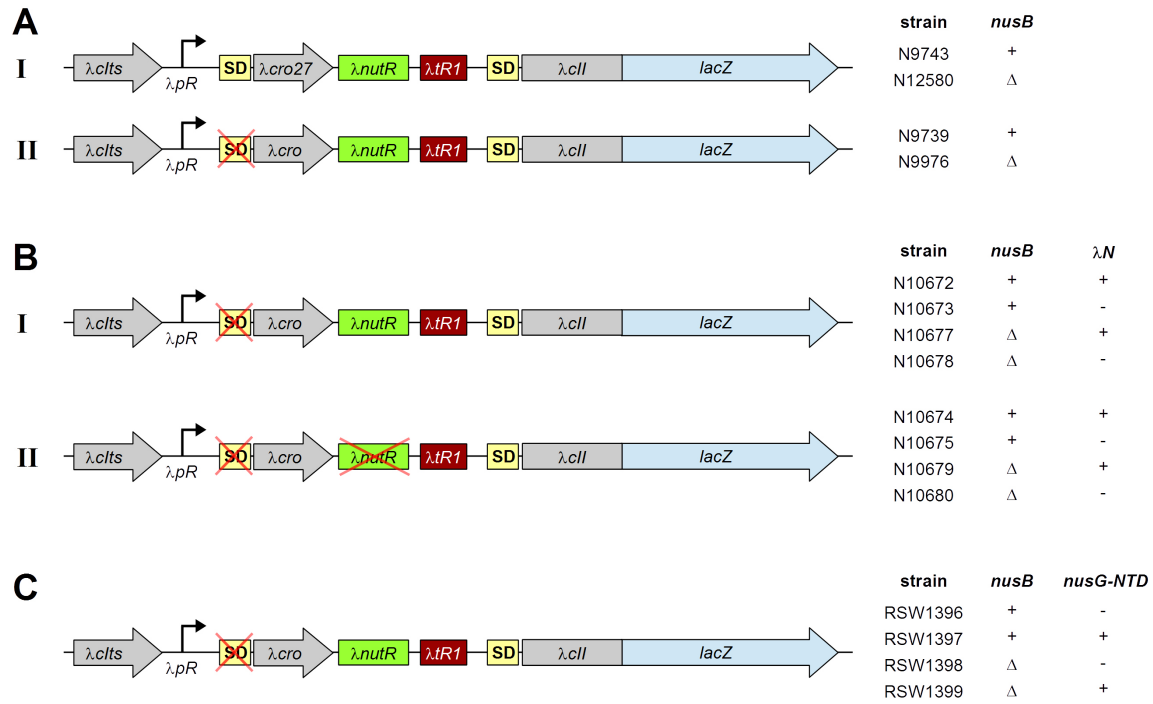797   read-through of *lacA* (see Table 1). Right: Schemes of possible effects on

798   transcription:translation coupling and Rho-dependent termination within *lacZ*. A-I, top:

799   Ribosomes are recruited in the early elongation phase, leading to a directly coupled

800   RNAP:ribosome complex (left) or uncoupled transcription and translation (right). A-I,

801   bottom: NusG is recruited in late elongation, resulting in a NusG-coupled complex with

802   (left) or without direct RNAP:ribosome contacts (middle), or modifying the pre-existing

803   RNAP:ribosome complex without establishing an CTD:S10 interaction (right). A-II:

804   Failure of NusG recruitment results in inefficient Rho-dependent termination and high

805   *lacA* read-through. A-III: *arg6* stops the translating ribosome at position +4, whereas

806   transcription elongation proceeds (left), resulting in ribosome dissociation and no NusG

807   recruitment. Transcription proceeds and is only inefficiently terminated by Rho (right).

808   A-IV: NusG couples transcription and translation (left) until *arg6* stops the ribosome at

809   position +200 (middle), allowing efficient, NusG-stimulated Rho-dependent termination

810   (right). B: λ*nutL* recruits NusA, NusG and the S10/NusB dimer, creating a Nus complex

811   (left). NusG can thus support Rho-dependent termination.

812

813

A

I — λcIts — λpR — SD — λcro27 — λnutR — λtR1 — SD — λcII — lacZ

strain	nusB
N9743	+
N12580	Δ

II — λcIts — λpR — SD — λcro — λnutR — λtR1 — SD — λcII — lacZ

strain	nusB
N9739	+
N9976	Δ

B

I — λcIts — λpR — SD — λcro — λnutR — λtR1 — SD — λcII — lacZ

strain	nusB	λN
N10672	+	+
N10673	+	-
N10677	Δ	+
N10678	Δ	-

II — λcIts — λpR — SD — λcro — λnutR — λtR1 — SD — λcII — lacZ

strain	nusB	λN
N10674	+	+
N10675	+	-
N10679	Δ	+
N10680	Δ	-

C

λcIts — λpR — SD — λcro — λnutR — λtR1 — SD — λcII — lacZ

strain	nusB	nusG-NTD
RSW1396	+	-
RSW1397	+	+
RSW1398	Δ	-
RSW1399	Δ	+

814

815 **Figure 5: NusG can be recruited *via* a Nus complex.** Genetic constructs used to

816 monitor NusG mediated Rho-dependent termination are shown with the corresponding

817 strains and their properties indicated on the right side. Transcription is started from the

818 λpR promoter, followed by WT-λcro or λcro carrying a missense mutation at codon 27

819 (λcro27), a WT or mutant λnutR site (B), the Rho-dependent terminator λtR1 and a

820 λcII::lacZ transcriptional fusion with corresponding SD site. All strains encode a

821 temperature sensitive λcI construct (λcIts) to allow temperature-controlled induction of

822 gene expression from the λpR promoter. λN$^+$ strains listed in (B) further encode the λN

823 protein; NusG-NTD for strains listed in (C) was supplied from plasmid pRM442. See

824 also tables 2A-C.

825

826   **Tables**

827

828   **Table 1. NusG couples late after transcription initiation.** β-galactosidase was induced

829   for 20 min from the *lac* operon with1mM IPTG. Where indicated, Rho-dependent

830   termination was inhibited by adding 100μg/ml BCM 1 minute prior to induction. Read-

831   through was calculated from the fold-increase of *lacA RNA* compared to *ompA* RNA in

832   the presence or absence of BCM. RNA levels were measured using qRT-PCR and the

833   fold- increase calculated using the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001).

834   RSW1225 carries two G to A mutations in the *lacZ* ribosome-binding site. RSW1245

835   carries an insertion of six rare arginine codons (atg-acc-atg-AGG-AGA-CGA-AGG-

836   AGA-CGA) at the amino terminus of *lacZ*. RSW1276 contains six rare arginine codons

837   200nt distal to the start of translation. RSW1297 carries an insertion of λ*nutL*

838   immediately 5' to the mutated ribosome binding site.

strain	*lacZ*	*nutL*	BCM$^-$	BCM$^+$	RT (%)
MDS42	wt	-	.25	.26	99
RSW1225	SD$^-$	-	.12	.56	21
RSW1245	arg(6) - early	-	.13	.49	29
RSW1276	arg(6) - late	-	<.01	.12	<1
RSW1297	SD$^-$	+	.01	.59	1

839

840   **Table 2. NusG recruitment depends on translation. (A**) NusG coupling at *nutR* requires

841   NusB. Expression of β-galactosidase was induced from a chromosomal *cII*::lacZ

842   transcriptional fusion (λ*cIts-pR-cro-nutR-tR1-cII::lacZ*) by incubating at 42^0 C for 30 min.

843   N9743 and N12580 carry a missense mutation at *cro* codon 27, N9739 and 9976 have a G

844   to C mutation in the *cro* Shine Dalgarno sequence (*SD-*), N12580 and N9976 are deleted

845   for *nusB*. Where indicated, bicyclomycin (BCM) was added to 100μg/ml prior to

846   induction of β-galactosidase. Read-through (RT) was calculated from the ratio of β-

847   galactosidase activity in the presence or absence of BCM. **(B)** BoxA mutations block

848   NusG coupling at *nutR*. Expression of β-galactosidase was induced from a chromosomal

849   *cII*::lacZ transcriptional fusion (λ*cIts-pR-cro* (*SD⁻*) *-nutR-tR1-cII::lacZ*) by incubating at

850   42^0 C for 30 min. Strains N10672, N10674, N10677 and N10679 express the λN

851   transcription termination inhibitor. *boxA69* and Δ*nusB* strain numbers are indicated in

852   Table 3. Read-through (RT) was calculated from the ratio of β-galactosidase activity in the

853   presence or absence of λN. **(C)** NusG-NTD uncouples. Expression of β-galactosidase was

854   induced from a chromosomal *cII*::lacZ transcriptional fusion (λ*cIts-pR-cro(*SD⁻*)-nutR-tR1-*

855   *cII::lacZ*) by incubating at 42^0 C for 30 min. The NusG-NTD was induced from the

856   plasmid pRM442 *tac* promoter with 1mM IPTG for 10 min prior to induction of

857   β−galactosidase in strains RSW1397 and RSW 1399. Strains RSW1396 and RSW1398

858   carried an empty vector (p*trc99A*) and were exposed to IPTG as above. Where indicated

859   bicyclomycin (BCM) was added to 100μg/ml prior to induction of β-galactosidase. Read-

860   through (RT) was calculated from the ratio of β-galactosidase activity in the presence or

861   absence of BCM.

862

A	strain	*cro*	*nusB*	BCM⁻	BCM⁺	RT (%)
	9743	*ms27*	+	530	680	78
	12580	*ms27*	D	890	1150	78
	9739	SD-	+	141	613	**23**
	9976	SD-	D	1191	1290	92

863

B	strain	*boxA*	*nusB*	λN⁻	λN⁺	RT (%)
	10673, 10672	+	+	126	946	**13**
	10675, 10674	*69*	+	1212	2211	55
	10678, 10677	+	D	2874	2616	100
	10680, 10679	*69*	D	1896	2416	78

864

C	strain	*nusG-NTD*	*nusB*	BCM⁻	BCM⁺	RT (%)
	RSW1396	-	+	247	862	**29**
	RSW1397	+	+	944	1013	93

	RSW1398	–	Δ	2013	2314	93
	RSW1399	+	Δ	2360	2760	86

865

866

867