Effects of individual base-pairs on *in vivo* target search and destruction kinetics of small RNA

Anustup Poddar[1], Muhammad S. Azam[2,6], Tunc Kayikcioglu[1], Maksym Bobrovskyy[2,7], Jichuan Zhang[1], Xiangqian Ma[2], Piyush Labhsetwar[3,8], Jingyi Fei[4], Digvijay Singh[1,9], Zaida Luthey-Schulten[3], Carin K. Vanderpool[2,*] and Taekjip Ha[1,5,*]

[1]Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

[2]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

[3]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

[4]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

[5]Howard Hughes Medical Institute, Baltimore, Maryland 21205

[6]Present Address: Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

[7]Present Address: Department of Microbiology, University of Chicago, Chicago, Illinois 60637

[8]Present Address: The Land Institute, Salina, Kansas 67401

[9]Present Address: Division of Biological Sciences, University of California, San Diego, San Diego, California 92093

* Corresponding authors

Taekjip Ha, Ph.D. (tjha@jhu.edu)

Carin K. Vanderpool, Ph.D. (cvanderp@life.illinois.edu)

Abstract

Base-pairing interactions mediate intermolecular target recognition in many biological systems and applications, including DNA repair, CRISPR, microRNA, small RNA (sRNA) and antisense oligo therapies. Even a single base-pair mismatch can cause a substantial difference in biological activity but presently we do not yet know how the target search kinetics *in vivo* are influenced by single nucleotide level changes. Here, we used high-throughput sequencing to identify functionally relevant single point mutants of the bacterial sRNA, SgrS, and quantitative super-resolution microscopy to probe the mutational impact on the regulation of its primary target, *ptsG* mRNA. Our super-resolution imaging and analysis platform allowed us to further dissect mutational effects on SgrS lifetimes, and even subtle changes in the *in vivo* rates of target association, k_{on}, and dissociation, k_{off}. Mutations that disrupt Hfq binding and are distal to the mRNA annealing region still decreased k_{on} and increased k_{off}, providing an *in vivo* demonstration that Hfq directly facilitates sRNA-mRNA annealing. Single base-pair mismatches in the annealing region reduced k_{on} by 24-31% and increased k_{off} by 14-25%, extending the time it takes to find and destroy the target mRNA by about a third, depending on whether an AU or GC base-pair is disrupted. The effects of disrupting contiguous base-pairing are much more modest than that expected from thermodynamics, suggesting that Hfq also buffers base-pair disruptions.

Introduction

Myriad biological systems use base-pairing interactions for target recognition where proteins mediate base-pairing interactions between two physically separated strands. Such base-pairing-mediated targeting is found in a wide range of processes including DNA repair[1], noncoding RNA-based gene regulation[2,3], bacterial immunity using CRISPR[4], and therapies using anti-sense oligonucleotides[5]. They all rely on base-pairing interactions above a threshold for specificity. How do they achieve both accuracy and speed to sample through thousands of potential targets and rapidly reject non-targets? Recent advances in single-molecule imaging technologies made it possible to explore the kinetic parameters of target recognition and non-target rejection *in vitro*[6–12], and in a limited number of cases, inside living cells[13]. However, we do not yet know the impact of single nucleotide changes in *in vivo* target search kinetics, even though such minute changes can have large functional consequences. Our goal here is to quantify the mutational impact on base-pairing-mediated target search kinetics *in vivo* using bacterial gene regulation by small RNA (sRNA) as a model system.

Among the many examples of non-coding RNA-based gene regulation are microRNAs and long non-coding RNAs in eukaryotes and sRNAs in bacteria and archaea[14,3,15]. Often, bacterial sRNAs regulate gene expression at a post-transcriptional level during stress, for example, in iron limitation stress[16], osmotic and acid stress[17], and oxidative stress[18]. Our work, here, studied the sRNA SgrS, which is produced in response to glucose-phosphate stress.[19]

A disparity between sugar uptake and its metabolism gives rise to stress; a faster uptake rate leads to an accumulation of glucose-6-phosphate and activation of SgrR, a transcription factor. This stimulates the *sgrS* gene to transcribe SgrS, which reduces sugar transport, promotes efflux and reroutes cellular metabolism[20–22]. Sugar stress conditions are provoked in most studies by

65 subjecting cells to α-methylglucoside (αMG), a sugar analogue that gets phosphorylated during
66 import to form αMG-6-phosphate, which cannot be further processed metabolically. *E. coli*
67 SgrS, a 227-nt sRNA, binds reversibly and dynamically to its primary target, *ptsG* mRNA[23],
68 which codes for the EIICB domain of the glucose phosphotransferase system. Binding between
69 the RNAs, aided by a hexameric RNA chaperone protein Hfq, blocks the *ptsG* ribosome binding
70 site, thereby inhibiting translation of new glucose transporters (Fig. 1a). This sRNA-mRNA
71 complex also gets degraded by endoribonuclease RNase E, thus reducing the cellular
72 concentration of *ptsG* mRNA. Hfq is important for the stability of sRNAs in general and *in vitro*
73 studies have shown that Hfq increases the rate of annealing between sRNA and its target mRNA
74 sequences[24–26]. Whether Hfq also directly facilitates annealing between sRNA and mRNA *in*
75 *vivo* is unknown for any sRNA because it has not been possible to separate the effects of Hfq on
76 sRNA stability and sRNA-mRNA annealing.

77 SgrS contains a 3´ Hfq-binding region predicted to contain two stem-loops, the small stem-loop
78 and the terminator stem-loop that is larger, followed by a U-rich tail (Fig. 1b)[27,28]. An optimal
79 length of U-rich tail, with seven nucleotides or more[27,29], is required for the formation of
80 functional sRNAs and for efficient Hfq binding, and Hfq binding to the two stem-loops is critical
81 for target regulation[27,28,30–32].

82 Nucleotides 168-187 of SgrS are partially complementary to the *ptsG* 5´-UTR (Fig. 1c).[20]
83 Nucleotides 168-181, if presented as a 14 nt long oligonucleotide alone, are sufficient for full
84 repression of *ptsG* translation *in vitro* and *in vivo*[33]. Among these, G176 and G178 have been
85 shown to be most important for the annealing between SgrS and *ptsG* mRNA[24].

86 Previously, we developed a two-color 3D super-resolution imaging and modeling platform to
87 determine *in vivo* target search kinetics of wild-type SgrS regulation on *ptsG*[34]. The bimolecular
88 association rate constant k_{on} between the RNAs was $2x10^5$ $M^{-1}s^{-1}$, which is within the wide range
89 of reported Hfq-mediated sRNA and target mRNA association rates *in vitro* despite the crowded
90 cellular environment and large excess of non–target RNA molecules. The dissociation rate
91 constant k_{off} was 0.2 s^{-1}; 10 to 100-fold larger than *in vitro* estimates of other sRNA-mRNA
92 pairs[32,35,36]. Its non-zero value showed that even for the correct target, binding is reversible. The
93 large dissociation constant K_D (=k_{off}/k_{on}) of ~1 μM explained why more than a hundred SgrS
94 molecules are needed for rapid *ptsG* mRNA regulation. The rate constant for co-degradation, k_{cat},
95 was surprisingly high, 0.4 s^{-1}, suggesting that RNA degradation machineries accompany the
96 target search complex formed between SgrS and Hfq so that as soon as RNAs bind each other,
97 RNAs can be degraded without waiting for the arrival of downstream degradation machineries.
98 Here, by expanding the scale of this quantitative imaging-based investigation by an order of
99 magnitude to include 10 SgrS mutants, we aimed to determine how k_{on}, k_{off} and k_{cat} are affected
100 by single nucleotide changes.

101 We formulated a pipeline of experiments to identify and examine the key regions in SgrS
102 responsible for the annealing and regulation of *ptsG.* We used Sort-Seq, a high-throughput
103 method that can estimate the impact of different mutations on the overall activity of the
104 fluorescence reporter system chosen[37–39]. From the Sort-Seq results, we identified the regions in
105 the SgrS sequence important for the overall regulation and chose nine single nucleotide
106 substitution mutants. *E. coli* strains containing these mutations or one double substitution
107 mutation in their endogenous chromosomal copy were constructed and studied using single-
108 molecule fluorescence *in situ* hybridization (smFISH) followed by 2-color 3D super-resolution

109 imaging and modeling to determine k_{on}, k_{off} and k_{cat}. Our results show that the two stem-loops at
110 the 3´ end of SgrS play important roles in the activity of the sRNA. We also provide *in vivo*
111 evidence that Hfq directly facilitates SgrS-*ptsG* mRNA base-pairing. Importantly, we were able
112 to unambiguously ascribe relative contributions of single base-pairs to sRNA lifetimes and target
113 search kinetics, allowing us to quantify by how much the rates of mRNA binding and rejection
114 are influenced by eliminating a single base-pair between them.

115

116 **Results**

117 **Sort-Seq reveals SgrS nucleotides important for target regulation**

118 We employed a high-throughput Sort-Seq approach to identify SgrS regions important for *ptsG*
119 regulation. We created a low copy number reporter plasmid containing a partial *ptsG* sequence
120 (105 nt 5´- UTR along with the first 30 nt coding sequence of *ptsG* mRNA) and superfolder
121 GFP-coding sequence (*ptsG-sf*GFP)[40] (Supplementary Fig. 1) and transformed it into *E. coli*
122 strain MB1 (Δ*ptsG*, Δ*sgrS*, *lacI*^{q}, *tet*^{R}). The *sgrS* mutation library was constructed by random
123 mutagenesis PCR of a plasmid[40] containing the *sgrS* sequence (Supplementary Fig. 1) and was
124 then transformed into the MB1 strain containing the reporter plasmid (Fig. 2a). The expression of
125 *ptsG-sf*GFP and *sgrS* were under the control of $P_{Llac-O1}$ and $P_{Ltet-O1}$, respectively, and were
126 induced by Isopropyl β-D-1-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc). Upon
127 induction by IPTG, cells containing the target reporter (*ptsG-sf*GFP) alone showed bright
128 fluorescence, while those co-transformed with the plasmid containing wild-type *sgrS* showed
129 weak GFP fluorescence in the presence of both IPTG and aTc in single-cell imaging
130 (Supplementary Fig. 2) and flow cytometry analysis (Fig. 2b), indicating an effective repression

131 of the reporter. Cells co-transformed with the *sgrS* mutant library showed a broad distribution of
132 GFP fluorescence indicating highly variable levels of regulation by mutants (Fig. 2b). Based on
133 the flow cytometry results, the cells were collected in five intensity bins, and the plasmids were
134 extracted from the cells. For each bin, the mutated *sgrS* sequence from position 149 to 227 was
135 amplified by PCR and sequenced. Sequencing was limited to this region because the 5´ region,
136 up to nucleotide 153 and coding for the 43 amino acid peptide SgrT, is not involved in base-
137 pairing-dependent mRNA regulation[41,42].

138 Using the relative abundance of sequences in each bin and the GFP fluorescence levels from the
139 flow cytometry analysis, we calculated, for each single point mutation, the average fluorescence
140 intensity of cells sharing the same mutation as a measure of the regulation defect (Fig. 2d).[39]
141 High average single cell fluorescence would correspond to SgrS mutants that are highly
142 defective in regulation of *ptsG* reporter expression and vice versa. The degree of perturbation to
143 the regulatory capacity is color-coded in the heatmap grid, ranging from the least (blue) through
144 intermediate (white) to the most (red). Nucleotides 149 to 174 showed little to no perturbation of
145 SgrS regulation as shown by the blue squares (Fig. 2d). In contrast, the region where SgrS can
146 base-pair with *ptsG* mRNA (U175 to G186) displayed perturbations across a wide range as
147 shown by the white and red squares in the grid. Specifically, previous studies showed that
148 G176C or G178C eliminates the SgrS's ability to downregulate *ptsG* while C174G and G170C
149 only weakly perturbs SgrS function *in vivo* and *in vitro*[24,31]. The corresponding squares in our
150 heatmap grid (Fig. 2d) show red or dark red for G176C and G178C, and white or light blue for
151 C174G and G170C, validating our Sort-Seq results. We also see that SgrS regulation is
152 hampered if there are mutations in the small stem-loop region (nts 183-196 (Fig 1b)), the
153 terminator stem-loop region (nts 199-219) and the poly-U tail (nts 220-227). The largest effect is

154 seen in the stem region of the terminator stem-loop, C199 to G205 and C213 to G219, where we
155 see the darkest red, highlighting the importance of this stem-loop. These stem-loop regions and
156 the poly-U tail play a role in Hfq binding[27,28], and our Sort-Seq analysis therefore confirms that
157 Hfq interaction is important for SgrS function in the cell.

158 Based on Sort-Seq results, we picked nine single point mutations for further investigation. These
159 include mutations in the target annealing region (A177U, G178A, G178U, U181A), U-rich
160 region upstream of the small stem-loop (U182A), the small stem-loop (G184A), the terminator
161 stem-loop (G215A) and the poly-U tail (U224G, U224A).

162

163 **SgrS mutation effects on regulation of *ptsG* reporter**

164 To examine the effect of the selected SgrS mutations on *ptsG* regulation, we monitored the effect
165 of wild-type and seven of the SgrS mutants (plasmid-encoded and expressed from an inducible
166 promoter) on the activity of a chromosomal *ptsG'-'lacZ* translational fusion (Fig. 3a). The wild-
167 type SgrS almost completely eliminated β-galactosidase activity whereas the mutants showed
168 regulation defects of various degrees consistent with the Sort-Seq data. SgrS G215A, which
169 disrupts the terminator stem-loop structure, showed the largest defect. To test if the regulatory
170 defects can be explained by a reduction of SgrS levels, for example, due to shorter cellular
171 lifetimes associated with impaired Hfq binding, we performed Northern blot analysis. We found
172 that SgrS abundance is not affected for four of the mutants (A177U, G178U, G178A and
173 G184A) and is reduced by 40-50% for mutations in the terminator stem-loop or poly-U tail
174 (G215A, U224G and U224A) (Fig. 3c). Interestingly, the latter three mutants showed large
175 increases in readthrough transcription, suggesting that transcription termination is defective (Fig.

176 3d). These observations are consistent with a previous study which showed that SgrS molecules
177 with an extended 3′ region do not interact with Hfq[43]. Overall, single point mutations outside the
178 large terminator stem-loop and poly-U tail have minimal impact on SgrS abundance and their
179 regulatory defects cannot be explained by SgrS abundance changes.

180

181 **Super-resolution imaging of specific chromosomal SgrS mutants**

182 A set of 9 single point mutants of SgrS were chosen for further analysis using quantitative
183 imaging (A177U, G178A, G178U, G184A, U181A, U182A, G215A, U224A, U224G). To avoid
184 potential complications arising from SgrS overexpression, we created these mutations in the
185 endogenous chromosomal copy of SgrS. A177, G178 and U181 are in the seed (target base-
186 pairing) region, G184 is in the small stem-loop region. U181 and U182 are in the U-rich region
187 upstream of the small stem-loop, previously shown to bind Hfq[27]. We also constructed a double-
188 mutant G184A-C195U which restores the small stem-loop structure. G215 is in the terminator
189 stem-loop region and U224 is in the poly-U tail, both of which provide major binding sites for
190 Hfq. These mutant alleles in the background of strains with wild-type RNase E or a C-terminal
191 truncated RNase E were grown, and glucose-phosphate stress was induced using αMG for a
192 varied amount of time before cell fixation and permeabilization. We performed two-color 3D
193 super-resolution imaging of the SgrS sRNAs labeled with up to 9 FISH probes conjugated to
194 Alexa Fluor 647 and the *ptsG* mRNAs labeled with up to 28 FISH probes conjugated to CF568.
195 Δ*sgrS* and Δ*ptsG* strains were also examined to correct for the background arising from
196 nonspecific binding of FISH probes. The wild-type strain showed an increase in SgrS copy
197 number over time after sugar stress induction (Fig. 4, Supplementary Fig. 3). At the same time,
198 the copy number of *ptsG* mRNA showed a decrease (Fig. 4, Supplementary Fig. 3). We used a

199 density-based clustering algorithm[34] to determine the copy numbers of RNAs along with the
200 copy number of SgrS-*ptsG* mRNA complexes. Super-resolution imaging was especially
201 important for quantifying sRNA-mRNA complexes because at conventional microscopy
202 resolution there was too much false co-localization between sRNA and mRNA.

203 The accumulation of the mutant SgrS sRNAs was lower than for wild-type SgrS with an
204 accompanying impairment in *ptsG* mRNA degradation for all single point mutants examined,
205 showing that their regulatory functions are perturbed. The single cell distribution of RNA copy
206 numbers also showed a decreased accumulation of SgrS, with the histogram peaking at lower
207 copy numbers 20 min after αMG induction, and the histograms for *ptsG* mRNA peaked at higher
208 copy numbers per cell compared to the wild-type (Fig. 5b, d, h). The lowest accumulation of
209 SgrS was seen for G184A and G215A, and they also showed the most impaired mRNA
210 degradation (Fig. 4, 5a, c, g, Supplementary Fig. 9, 11). These two mutations occur in two
211 separate stem-loop regions, both of which participate in Hfq binding.[27,28] The double mutant,
212 G184A-C195U, which restores base-pairing in the small stem-loop via a compensatory mutation,
213 eliminated the negative impact of G184A as seen by recovery of SgrS accumulation and
214 regulation of *ptsG* mRNA (Fig. 4, 5e, f, Supplementary Fig. 10). This suggests that the
215 disruption of the stem-loop structure, not of G184 basepairing with *ptsG* mRNA, is primarily
216 responsible for regulatory defects of G184A.

217 These imaging data by themselves cannot tell us whether regulatory defects are due to changes in
218 target binding kinetics or due to changes in the SgrS stability. Therefore, we next determined the
219 lifetimes of wild-type and mutant SgrS molecules.

220

221 **Intrinsic lifetimes of SgrS mutants**

222 In order to calculate the target-independent lifetime of SgrS, we induced SgrS expression using
223 αMG and then added rifampicin to stop transcription globally. RT-qPCR was performed vs time
224 after rifampicin treatment to quantify the SgrS level. The wild-type SgrS showed minimal
225 intrinsic degradation over a period of 2 hours after the addition of rifampicin but it showed rapid
226 degradation in the presence of ongoing transcription (10.4 ± 0.7 min), suggesting that SgrS
227 degradation is normally dominated by co-degradation with its various target mRNAs (Fig.
228 Supplementary Fig. 25). The intrinsic degradation was also minimal for SgrS A177U mutant
229 (Supplementary Fig. 25), suggesting that in the absence of co-degradation, a mutation in the
230 target-annealing region does not destabilize SgrS. In contrast, intrinsic degradation of G184A
231 was rapid (lifetime of 6.3 min) and so was the intrinsic degradation of wild-type SgrS in Δ*hfq*
232 strain (lifetime of 5.1 min) (Supplementary Fig. 25), indicating that Hfq is required for the target-
233 independent stability of SgrS and the small stem-loop is important for Hfq binding.

234

235 **Lifetime of SgrS mutants**

236 In order to determine the effective lifetime of SgrS mutants, which includes the contributions
237 from intrinsic degradation and co-degradation with target mRNA, the strains carrying
238 chromosomal mutations were treated with αMG for 10 minutes before rinsing it away. SgrS
239 decay over time was then monitored through imaging of fixed cells. The wild-type SgrS showed
240 a degradation rate of 0.0016 ± 0.0001 s^{-1} (lifetime of 10.4 ± 0.7 min) and all of the mutants
241 showed higher rates, the highest being for G184A with 0.0046 ± 0.0003 s^{-1} (lifetime of 3.6 ± 0.2
242 min), followed by G215A with 0.00345 ± 0.0003 s^{-1} (lifetime of 4.8 ± 0.4 min) (Fig. 6a,

243 Supplementary Fig. 26-32, 34-36). Because G184A and G215A disrupt the small and terminator
244 stem-loop regions, respectively, our data suggest that both stem-loop regions are important for
245 SgrS stability *in vivo*. G184A-C195U recovered the stability of SgrS to the wild-type level with
246 an identical degradation rate within error (Fig. 6a, Supplementary Fig. 33). Rifampicin-chase
247 experiments did not show any difference in the lifetime of *ptsG* mRNA between all mutant
248 strains (Supplementary Fig. 26-37), showing that the mutations in SgrS have no effect on *ptsG*
249 mRNA stability when SgrS is not induced.

250 The degradation rate of SgrS in Δ*hfq* strains was much higher, about $0.022 \pm 0.004\ s^{-1}$ (lifetime
251 of 0.76 ± 0.14 min), for wild-type and all SgrS mutants (Fig. 6a, Supplementary Fig. 38-40) [23].
252 This 14-fold increase in degradation rate for sRNAs in Δ*hfq* strains confirms that Hfq is
253 indispensable for the stability of SgrS[27,44]. Because none of the SgrS mutants in the hfq^+ cells
254 showed degradation rates as high as in Δ*hfq* strains, these SgrS mutations are only partially
255 deleterious to the interactions with Hfq.

256 Mutations in the base-pairing regions (A177U, G178U and G178A) reduced the lifetime of SgrS
257 in the imaging-based experiment even though they are not expected to alter Hfq binding.
258 Because our Northern blot analysis of overexpressed SgrS showed that these mutations do not
259 change SgrS abundance, the intrinsic degradation is unlikely to be affected by the mutations.
260 Instead, we attribute the discrepancy to mutation-induced alterations in co-degradation of SgrS
261 with other SgrS target mRNAs.

262

263 **Target search and destruction kinetics of SgrS mutants**

264 Once we obtain the average copy numbers of SgrS, *ptsG* mRNA and the SgrS-*ptsG* complex per
265 cell as a function of time after SgrS induction, we can use a previously developed deterministic
266 kinetic model to describe the SgrS-*ptsG* regulation kinetics (Fig. 1a)[34]. We used the
267 experimentally-determined degradation rate for *ptsG* mRNA, β_p, to calculate the *ptsG*
268 transcription rate α_p using $\alpha_p = \beta_p \times [p]_0$, where $[p]_0$ is the steady state copy number of *ptsG*
269 mRNA at t=0. By globally fitting the six time courses of the three RNA species with or without
270 RNase E mutation that inhibits co-degradation, we obtained k_{on}, k_{off} and k_{cat} for the wild-type and
271 mutant SgrS.

272 k_{on} for the wild-type strain was (1.9 ± 0.2) x 10^5 M^{-1}-s^{-1} and k_{off} was 0.22 ± 0.02 s^{-1} giving a
273 K_D of 1.16 ± 0.14 μM, comparable to the previously published results (Fig. 6b-d)[34]. k_{on} was
274 lower for all single point mutants compared to the wild-type and the reduction ranged from 24%
275 for A177U to 53% for G184A. k_{off} was higher for all the mutants and the increase ranged from
276 14% for A177U to 33% for G184A giving a dissociation constant, K_D of 1.67 ± 0.13 μM and
277 3.08 ± 0.33 μM, respectively (Fig. 6b-d). k_{cat} was not affected by the mutations within error
278 (Supplementary Fig. 43). To test the possibility that the apparent changes in k_{on} and k_{off} are due
279 to fitting errors and that the regulatory deficiencies can be explained solely by reduction in SgrS
280 lifetimes, we repeated the global fitting procedure while keeping the k_{on} and k_{off} values fixed at
281 the wild-type values. The fits were considerably worse, and were especially poor for copy
282 number curves of *ptsG* mRNA and SgrS/mRNA complex (Supplementary Fig. 44-51).
283 Therefore, our procedure of obtaining the mutation effects on k_{on} and k_{off} is robust.

284 When we attempted to restore the base-pairing in the small stem-loop by adding a compensatory
285 mutation to G184A that showed the largest changes to k_{on} and k_{off} (G184A-C195U), k_{on} and k_{off}

286 returned to the wild-type values within error (Fig. 6c-d). Nucleotides 168-187 in SgrS were
287 originally proposed to participate in base-pairing with the *ptsG* mRNA[20] but a subsequent study
288 showed that only nucleotides 168-181 are required for basepairing[33]. Here we found that binding
289 kinetics is similar between the wild-type and G184C-C195U, strongly suggesting that a mutation
290 at G184 primarily acts through disruption of the small stem-loop structure, thereby affecting Hfq
291 binding, instead of through direct disruption of base-pairing of G184 with the target strand. Even
292 though our data suggest that G184 is not involved with SgrS-*ptsG* mRNA base-pairing, its
293 mutation negatively affected annealing kinetics, decreasing k_{on} and increasing k_{off}. Therefore, our
294 results support the dual roles of Hfq: first to increase sRNA stability (Fig. 6a) and second to
295 directly facilitate SgrS-*ptsG* binding.

296 The A to U mutation at position 177 removes an AU base-pair, breaking 8 base-pairs, the longest
297 stretch of contiguous base-pairing between SgrS and *ptsG* mRNA into segments of 4 and 3 base-
298 pairs. This disruption gives a reduction in association rate of 24% and an increase in dissociation
299 rate by 14%. The two mutations at position 178 eliminate a GC base-pair and breaking the same
300 8 base-pairs into segments of 3 and 4 base-pairs. G178A and G178U mutants gave a reduction of
301 association rate by 31-32% and an increase of dissociation rate by 23-25%. The larger effects of
302 G178 mutations compared to A177U are likely due to the loss of GC over AU base-pair.
303 Consistent with this suggestion, a mutation at U181, losing an AU base-pair, decreased the
304 association rate by 26% and increased the dissociation rate by 18%, very similar to A177U
305 values.

306 The G215A mutation in the terminator stem-loop and the mutations U224A and U224G in the
307 poly-U tail showed k_{on} decreases and k_{off} increases even though they should not change
308 complementarity between SgrS and *ptsG* mRNA. The substantial effects on binding kinetics

309 must therefore be due to defects in Hfq's ability to facilitate the annealing reaction, further
310 providing *in vivo* evidence of direct facilitation of base-pairing between sRNA and mRNA by
311 Hfq.

312

313 **Difference in regulation outcome between imaging and Sort-Seq experiments**

314 To examine if the regulation outcomes for SgrS mutants is consistent between our quantitiave
315 imaging experiments and Sort-Seq analysis, we used the fractional decrease of *ptsG* mRNA over
316 the first 20 min after sugar stress induction as a measure of the SgrS regulation of *ptsG* mRNA
317 target in imaging-based analysis. (Fig. 6e-g). Plotting these values vs the inferred GFP signals
318 obtained from the Sort-Seq experiments, we observed a relatively weak correlation (Pearson's R
319 = 0.71), suggesting that the translation reporter-based Sort-Seq method is not able to faithfully
320 capture the regulation defects of SgrS mutations. For example, G178A which had a large
321 deficiency in regulation in imaging experiments showed almost the wild-type level regulation in
322 Sort-Seq. A large defect in regulation was shown in previous studies where nucleotide 178 was
323 mutated and our imaging-based approach is in accordance to this finding[24]. SgrS overexpression
324 in Sort-Seq may have overcome the negative effect of mutations through mass action when the
325 defect is primarily in binding kinetics. For the G215A, U224A and U224G mutant strains,
326 however, Sort-Seq showed large regulatory deficiencies, suggesting that their defects cannot be
327 overcome by overexpression. There are several possible explanations. First, because Sort-Seq
328 relies on the translational output, mutations that disrupt translational inhibition but not RNA co-
329 degradation may not be scored well in imaging-based experiments. However, *in vitro* translation
330 experiments showed that mutation of G178 to C eliminates translation inhibition by SgrS[33],
331 making it unlikely that defects in RNA-RNA annealing do not affect translational repression.

332 Second, even when these mutants can bind Hfq, the complex may be defective in mediating
333 RNA annealing. Third, these mutations also interfere with proper termination as shown by
334 readthrough transcripts (Fig. 3). It was shown previously that the readthrough products of SgrS
335 transcription do not bind Hfq *in vivo* and *in vitro*[43]. Weakening of the terminator stem-loop or
336 reduction of the slippery Us must be causing transcription readthroughs that produce regulation-
337 defective products, and much of the effect of G215A, U224A or U224G may be due to improper
338 termination. It has also been shown that readthrough transcription of SgrS is suppressed under
339 stress conditions, providing an additional layer of regulation[43]. Finally, the incorporation of the
340 GFP in the reporter system may have affected the stability of the mRNA.

341

342 **Discussion**

343 Previous studies have measured the effect of mutations in the regulation of mRNA targets of
344 SgrS[45–47] and have shown the importance of Hfq, sequence complementarity between SgrS and
345 its targets, and RNA secondary structures[27,33,48,49]. Hfq has also been shown to promote structural
346 changes to the RNAs, which in turn helps in the annealing and, consequently, regulation[50–52].
347 Our study provides a quantitative description of the process of target search and off-target
348 rejection by determining the kinetic parameters as a function of single nucleotide changes in
349 functionally important regions in SgrS. The k_{on} and k_{off} values determined in this study depict the
350 apparent rate constants because we did not explicitly include Hfq binding in our model. k_{on} in
351 particular should have contributions from Hfq binding to SgrS, target search by SgrS/Hfq and
352 subsequent annealing.

353 We used IntaRNA[53–56] to predict the energy of interaction between SgrS and *ptsG* mRNA and
354 found that it changes by ~6.4 kcal/mol for the G178 point mutations whereas the change is only

355 around ~4.3 kcal/mol for A177U. Our study agreed with the ranking because we saw lower rates
356 of association and higher rates of dissociation for G178 than A177U. However, the magnitude of
357 the effect is much more modest compared to a simple prediction based on the energetic penalty.
358 Both mutations introduce a mismatch within eight contiguous base-pairs, incurring large
359 energetic penalties. For example, 6.4 kcal/mol would correspond to a change in the equilibrium
360 binding constant by a factor ~60,000 instead of ~2 we observed. Therefore, Hfq must be
361 buffering the effect of breaking internal base-pairs in short helices. How this is achieved is
362 presently unknown.

363 We found that the rate of co-degradation remains high, ~0.3 s^{-1}, even with SgrS mutations. The
364 co-degradation of the SgrS-*ptsG* complex is brought about by the degradosome, in which RNase
365 E is a key component. Hfq copurifies with RNase E and SgrS[57], and at least one sRNA (MicC)
366 has been shown to mediate the interaction between Hfq and the C-terminal part of RNase E *in*
367 *vitro*[58] and *in vivo*[59]. It has been hypothesized that the sRNA-Hfq-RNase E complex forms first
368 and subsequently the complementary mRNA binds to this complex, aided by Hfq, followed by a
369 coupled or sequential degradation of the RNA pair[58]. The changes in k_{on} and k_{off} in our study
370 account for the disruption of the SgrS-*ptsG* mRNA annealing, but once a stable complex with all
371 four components forms, the co-degradation occurs at the same rate irrespective of the SgrS
372 mutations.

373 Because k_{cat} did not change with SgrS mutations, the probability that a single binding event will
374 cause co-degradation of sRNA and mRNA decreases with a mutation-induced increase in k_{off}. On
375 average, the wild-type SgrS would take 1.73 (= $(k_{off}+k_{cat})/k_{cat}$) binding events before co-
376 degradation. This number increases to 1.81 when AU basepairing is disrupted by a mutation and
377 increases further to 1.9 when GC basepairing is disrupted. In addition, the binding rate at the full

378 SgrS accumulation condition, 100-200 copies per cell, would correspond to 0.48 μM (assuming
379 0.7 μm^3 per cell), and the wild-type SgrS would take about 11 s to bind *ptsG* mRNA. If a mutant
380 SgrS were present at the same concentration, it would take ~13.6 s and ~14.4 s to bind for
381 disrupting a single AU and GC base-pair, respectively. The overall time it takes to degrade the
382 target would increase from 19 s to 24.6 s and 27.4 s for disrupting an AU and GC base-pair,
383 respectively. Although we examined only the effect of SgrS mutation in this study, if we assume
384 that a mutation in the target mRNA breaking a single base-pair has a similar impact, we can
385 conclude that target search and destruction will take 29-44 % longer for nearly cognate off-target
386 RNA containing a single base-pair mismatch.

387 The poly-U sequence at the 3' end of sRNAs is an important Hfq binding module[29,43] and binds
388 to the proximal face of the ring-shaped Hfq hexamer[29,50] . Because Hfq forms a stable 1:1
389 complex with SgrS[27], a single Hfq hexamer must bind the poly-U tail, both of the stem-loops,
390 and the UA-rich region upstream of the small stem-loop simultaneously. How this is achieved
391 requires further structural studies, but the UA-rich region has been found in other sRNAs and has
392 been proposed to bind the rim of the Hfq hexamer[27,51,60–62]. The binding of Hfq with target
393 mRNAs, however, has been studied extensively and it is known that Hfq brings about a
394 distortion in the mRNA structure, promoting the base-pairing between the RNAs[63,64]. In this
395 study we showed that the U224 mutations in the poly-U tail caused the rate of association of the
396 RNAs to decrease and the rate of dissociation to increase. Because U224 is distant from the
397 mRNA annealing region of SgrS, our data showed that Hfq directly facilitates RNA-RNA
398 annealing *in vivo*. The same effect was observed from other mutants that disrupt Hfq binding
399 without changing the mRNA annealing region of SgrS, and collectively our work presents the
400 first *in vivo* evidence that Hfq directly facilitates target binding. It should be emphasized that a

401 careful accounting of SgrS mutations' effects on SgrS lifetimes was necessary to reach this
402 conclusion. Microscopic mechanisms for Hfq's role in sRNA-mRNA annealing are still a subject
403 of active research[60,65,66], and may be investigated in the future using our analysis platform.

404

405 **Materials and Methods**

406 **Construction of plasmids for Sort-Seq studies**

407 A ptsG-*sf*GFP reporter system was constructed, containing 105 nt 5´ UTR and 30 nt coding
408 sequence of *ptsG* mRNA, which coded for the first 10 amino acids of PtsG protein, and this was
409 fused by a 42 nt linker sequence and the superfolder GFP coding sequence. The reporter system
410 was subcloned from the pZEMB8 plasmid. A plasmid, pAS06 was constructed by inserting this
411 reporter sequence into the low copy plasmid pAS05 between the XhoI and XbaI restriction sites
412 and the expression of the reporter system was under the control of $P_{Llac\text{-}O1}$.

413 The SgrS sRNA sequence was inserted in between the NdeI and BamHI restriction sites of the
414 medium copy plasmid pZAMB1 and its expression was under the control of $P_{Ltet\text{-}O1}$. The sgrS
415 mutation library was prepared by using the plasmid pZAMB1 as a template for mutagenesis PCR
416 and also as a vector to insert the sgrS mutation sequence.

417 **Cell culture and induction for Sort-Seq studies**

418 The *E. coli* MB1 strain (Δ*ptsG*, Δ*sgrS*, $lacI^q$, tet^R) was transformed with plasmids (pAS06 for
419 ptsG-*sf*GFP and pZAMB1 for *sgrS* or the SgrS mutation library) and grown at 37 ºC in LB Broth
420 Miller (EMD) overnight with the respective antibiotics (100 µg/ml ampicillin (Gold
421 Biotechnology, Inc.) for pAS06 plasmid and 30 µg/ml chloramphenicol (Sigma-Aldrich) for

422 pZAMB1 plasmid and the *sgrS* mutation library). The following day, the cell culture was diluted
423 200-fold into fresh LB Broth with respective antibiotics and were grown until OD_{600} reached 0.1-
424 0.2 as measured using an Educational Spectrophotometer (Fisher Scientific Education). The
425 culture was diluted again to an OD_{600} of 0.001 and supplemented with 1 mM IPTG (Sigma-
426 Aldrich) to induce the expression of PtsG-*sf*GFP and 50 ng/ml aTc to induce the expression of
427 SgrS or the SgrS mutation library. The *E. coli* cells were collected and treated further for the next
428 set of experiments.

429 **SgrS sRNA mutagenesis experiment**

430 Agilent Genemorph II Random Mutagenesis Kit (Agilent Technologies) was used to perform
431 mutagenesis PCR on SgrS using the protocol adapted from previously published work from
432 Levine's lab[39]. 1 ng of pZAMB1 plasmid, with the sgrS sequence, was used to conduct
433 mutagenesis PCR for 15 cycles. The yields of the individual mutants were increased by
434 amplifying the product using Phusion® High-Fidelity DNA Polymerase (New England Biolabs)
435 and purified using QIAquick Spin Columns (Qiagen) . The PCR products were then digested
436 with NdeI (New England Biolabs) and BamHI (New England Biolabs) and purified by QIAquick
437 Spin Columns (Qiagen). The pZAMB1 vector was also prepared by digestion with NdeI and
438 BamHI followed by purification using QIAquick Spin Columns (Qiagen). The vector and the
439 PCR insert were used to prepare 4 ligation reactions by mixing with T4 Ligase (New England
440 Biolabs).

441 The products from all the reactions were combined and purified using QIAquick Spin Columns
442 (Qiagen) into water. 5 μl of the purified ligation product was then transformed into MB1 strain,
443 which was pre-transformed with pAS06 plasmid expressing ptsG-*sf*GFP. These transformed cells
444 were then recovered and diluted into LB Broth supplemented with 100 μg/ml ampicillin and 30

445 μg/ml chloramphenicol and grown overnight at 37 ºC. The following day, the culture was
446 centrifuged, aliquoted as frozen stocks and used for imaging and flow cytometry experiments.

447 **Epifluorescence Imaging**

448 1 ml culture of the *E. coli* strain to be imaged was grown from an overnight culture till OD_{600} =
449 0.1-0.2. It was then chilled on ice followed by centrifugation at 6000 g, 4 ºC for 1 minute to form
450 a cell pellet. Then they were washed with ice-cold 1X PBS twice and resuspended in 100 μl 1X
451 PBS.

452 1.5% (w/v) agarose gel was prepared by dissolving agarose in 1X PBS. A few μl cell suspension
453 was sandwiched between a No. 1.5 glass coverslip (VWR) and a thin slab of the agarose gel. The
454 sample was then imaged.

455 The epifluorescence images were acquired by a Nikon Ti Eclipse microscope (Nikon
456 Instruments, Inc.) using an oil immersion objective (1.46 NA, 100X) which spans an area of
457 around 133x133 μm^2 for DIC (no filter, autofluorescence) and fluorescence imaging (Ex 480-
458 500 nm, Em 509-547 nm, exposure time 200 ms). The images were acquired using an EMCCD
459 camera (Andor). They were processed using the NIS-Element AR software (Nikon Instruments,
460 Inc.).

461 **Fluorescence-Activated Cell Sorting**

462 The *E. coli* strain to be sorted was cultured overnight in LB Broth with appropriate antibiotics.
463 The following day, the liquid culture was diluted 200-fold and cultured with antibiotics until
464 OD_{600} = 0.1-0.2. The cells were then diluted to OD_{600} = 0.001 in LB Broth with antibiotics and 1
465 mM IPTG and/or 50 ng/ml aTc were added corresponding to the strain of *E. coli* and the

466 plasmids it is carrying. They were grown till OD_{600} = 0.1-0.2, washed with ice-cold 1X PBS
467 twice and kept on ice before flow cytometry analysis or fluorescence-activated cell sorting. The
468 sorting and analysis were done in a MoFlo XDP Cell Sorter (Beckman Coulter) using a 488nm
469 200mW laser.

470 **Preparation of the sample for sequencing**

471 The cells sorted into the batches were grown in LB Broth supplemented with 30 μg/ml
472 chloramphenicol to saturation. We extracted the plasmids with E.Z.N.A. Plasmid Mini kit
473 (Omega, D6942-02). To generate sequencing amplicons, we followed Illumina 16S sequencing
474 protocol. We used 5 ng of each plasmid elute as PCR template and amplified out the portion of
475 interest using 0.5 μM of primers annealing to the region of the *sgrS* sequence under
476 consideration with Phusion 2X Mastermix (NEB, M0531L). We employed 20 cycles of 10 s at
477 98 ºC denaturation, 20 s at 63 ºC primer annealing, 10 s at 72 ºC elongation phases preceded by
478 additional initial denaturation at 98 ºC for 30 s and followed by 72 ºC final extension for 2 min.
479 To clean up the product, we incubated the PCR product with 20 μl Ampure XP beads (Beckman
480 Coulter, A63880) for 5 min. We retained the bead-bound material after keeping for 2 minutes on
481 a magnetic rack (GE, 1201Q46). We washed the beads twice with 80% ethanol, air-dried for 10
482 min and eluted the material in 53 μl 10 mM Tris pH 8.5 by incubation for 2 min. We collected
483 45-50 μl bead-free liquid 2 min after placing the material on a magnetic rack.

484 **Illumina Next-Gen Sequencing**

485 We performed 8 additional cycles of PCR with Nextera 24-Index kit for indexing before sample
486 pooling (Illumina, FC-121-1011), for which we used 7.5 μl of the above elute as template, 7.5 μl
487 each of the suitable i5 and i7 primers with 38 μl Phusion 2X Mastermix. We followed

488 manufacturer's recommended thermal cycling protocol (95 ºC 3min, 98 ºC 30s, 55 ºC 30s, 72 ºC
489 30s, 72 ºC 5min). We also bead-purified 55 µl of this final product with 56 µl Ampure-XP beads
490 and eluted with 28 µl 10 mM Tris pH 8.5 buffer. We pooled the final products based on their
491 Nanodrop reading at equal molar stoichiometry and diluted the sample down to 4 nM in 10 mM
492 Tris buffer. We alkaline-denatured by mixing 5 µl DNA sample with 0.2 M NaOH and
493 incubating for 5 min at room temperature. We diluted this product down to 20 pM in Hbf buffer.
494 We loaded a final mixture of 465 µl Hbf buffer, 120 µl pooled 20 pM library and 15 µl denatured
495 20 pM PhiX control library (Illumina, FC-110-3001) after a 2 minute heat treatment at 96 ºC
496 followed by a 5 min incubation on ice. We used a 150 cycles MiSeq v3 reagent kit (Illumina,
497 MS-102-3001) to perform a single-end sequencing for 150 cycles. We used the manufacturer's
498 default algorithm for base calling and de-multiplexing of the constituent samples.

499

500 **Intensity Moment Calculation**

501 We parsed the raw .fastq output files via a simple home-made C++ script compiled with GCC v.
502 7.5 and plotted with GNU Octave v. 4.2. The analysis scripts can be accessed via the Gitlab page
503 https://gitlab.com/tuncK/sortseq/-/tree/master and the raw data can be obtained from TH upon
504 request.

505 We only imported the base calls of each read, thus including all sequences regardless of their
506 quality factors. We directly extracted from each read the subsequence excluding the PCR
507 adaptors, i.e. bases 23 to 128. Out of this list of subsequences, we detected ones that are exact
508 duplicates of each other by building a red-black binary search tree. Among all such groups, we
509 only considered SgrS sequence variants that are represented by at least 10 distinct reads in the

510 data. We compared the observed sequence of each group with the wild-type SgrS sequence i.e.
511 that of the plasmid used as error-prone PCR template.

512 We normalized the raw number of reads of each group by both the total number of reads and the
513 fraction of cells falling under each gate. As such, we defined a weighted average intensity to
514 each individual mutant along this 106 base long SgrS segment that we probed. Referred to as the
515 "intensity moment" from now on, we calculated the following quantity:

516

517 $$K_{ij} = \frac{\sum_{k=-2}^{2} \langle I_k \rangle c^k n_{ij}^k / N^k}{\sum_{k=-2}^{2} c^k n_{ij}^k / N^k} \qquad (1)$$

518

519 where, K_{ij} is the intensity moment of the mutant carrying a single substitution mutation at the i'th
520 base position to nucleotide type j rather than the wt base. c^k is the overall fraction of cells that
521 are sorted into the k'th bin based on the GFP intensity histogram that FACS acquisition
522 software reports. n^k is the number of reads carrying a single substitution mutation at base
523 position i to base type j and detected in the k'th FACS bin. N^k is the total number of acceptable
524 reads in the dataset. $N^k \geq \sum_{ij} n_{ij}^k$ due to experimental errors as well as reads carrying multiple
525 substitutions due to the stochastic nature of error-prone PCR. $\langle I_k \rangle$ is the median intensity of
526 the cells falling into the k'th bin as reported by FACS. For the representative intensity of
527 each bin, we used the median intensity reported by the FACS device.

528 In the figures, we reported the standard score of each entry given by

529 $$z = \frac{K_{ij} - \langle K_{ij} \rangle}{\sigma_{ij}(K_{ij})} \qquad (2)$$

530 **Construction of bacterial strains**

531 The oligonucleotides, plasmids, and strains used in this study are listed in Supplementary Tables
532 1 and 2. *E. coli* K12 MG1655 derivatives were used for all experiments. P1 transduction[67] or λ-
533 red recombination[68] were used to move alleles between strains. DNA fragments were PCR
534 amplified using Q5® Hot Start High-Fidelity 2X Master Mix (NEB) and oligonucleotides
535 described in Supplementary Table 2. A set of plasmids (Supplementary Table 1) were used as
536 templates to PCR amplify the wild-type and *sgrS* mutants A177T, G178T, G178A, and G184A
537 using single-stranded oligos (Supplementary Table 2) containing 5' and 3´ homology to the
538 flanking regions of *cat-sacB* cassette (MB205). DNA fragments containing the *sgrS* mutants
539 G215A, T224G, T224A, T181A, and T182A were PCR amplified from MG1655 genomic DNA
540 using oligonucleotides listed in Supplementary Table 2.

541 **Measurement of intrinsic degradation rates of SgrS** Strain DB166, MB206, MB209 and
542 XM199 were cultured overnight at 37 °C and diluted 1:100 to a fresh LB medium and the
543 cultures were grown at 37 °C to OD_{600} ~ 0.3. To induce SgrS expression, αMG was added to
544 final concentration of 0.5% and the cells were grown for additional 30 minutes. Rifampicin was
545 added to final concentration of 250 μg/ml and the cells were grown for another 5 minutes. At this
546 point, cells were harvested (t=0 timepoint) for RNA extraction. Three biological replicates were
547 harvested for each time-point.

548 Cells from 1.0 ml of culture were mixed with 2 ml of RNA protect reagent (Qiagen). The
549 mixture was pelleted at 4000 rpm for 10 minutes and then discard the supernatant. Total RNA
550 was isolated using Direct-Zol RNA miniPrep (Zymo) kit following the manufacturer's
551 instruction. Genomic DNA was removed by DNaseI provided by the Kit. Finally total RNA

552 was eluted in 40 µl of nuclease free water. First-strand cDNA was synthesized from 1 µg of total
553 RNA using Superscript™ IV First-Strand cDNA Synthesis SuperMix kit according to the
554 manufacturer's protocol (Invitrogen, USA).

555 The primers used to amplify SgrS are: OSA499 (GATGAAGCAAGGGGGTGCCC) and
556 OSA500 (CAATACTCAGTCACACATGATGCAGGC)

557 The primers used to amplify housekeeping gene *rrsA* are: OXM187
558 (ATTCCGATTAACGCTTGCAC) and OXM188 (AGGCCTTCGGGTTGTAAAGT)

559 Real-time PCR was performed using SYBR Green master mix (Fisher) and Eppendorf Realplex
560 in a 96-well plate. Each reaction is comprised of 1x SYBR Green master mix, 100 nM of each
561 primer, 2 µl of 1:50 diluted cDNA in a total of 10 µl reaction volumes. Each plate contains "no
562 template" controls for individual transcripts as well as housekeeping transcripts such as *rrsA* for
563 every sample as an internal control.

564 Delta delta Ct method was used to analyze the qPCR data. The transcripts turnover rates were
565 calculated based on the non-linear fit with one phase exponential decay curves using GraphPad
566 software.

567 **Cell culture, fixation and permeabilization for smFISH and super-resolution imaging**

568 The wild-type *E. coli* strain (DJ480) was grown overnight at 37 ºC, 250 rpm in LB Broth Miller
569 (EMD), the RNase E mutant was grown in 25 µg/ml kanamycin (Kan) (Fisher Scientific), the
570 SgrS A177U, G178U, G178A, U181A, U182A, G184A, G184A-C195U, G215A, U224A,
571 U224G mutants were grown in LB Broth with 50 µg/ml spectinomycin (Spec) (Sigma-Aldrich)
572 and the RNase E mutants of the respective SgrS mutations were grown in LB Broth with 25
573 µg/ml kanamycin and 50 µg/ml spectinomycin. The following day, the overnight cultures were

574 diluted 100-fold into MOPS EZ rich defined medium (Teknova) with 0.2% glucose and the
575 respective antibiotics, and allowed to grow at 37 ºC and 250 rpm until the OD_{600} reached 0.15-
576 0.25. α-methyl D-glucopyranoside (αMG) (Sigma-Aldrich) was used to introduce sugar-
577 phosphate stress and subsequently induce SgrS sRNA expression. A specific volume of liquid
578 was taken out of the culture after 0, 2, 4, 6, 8, 10, 15, 20 minutes of incubation and mixed with
579 formaldehyde (Fisher Scientific) to a final concentration of 4% for the fixation of the cells.

580 Δ*sgrS* and Δ*ptsG* strains were grown overnight in LB Broth Miller (EMD) at 37 ºC and 250 rpm
581 using 25 µg/ml kanamycin and 10 µg/ml tetracycline (Tet) (Sigma-Aldrich) respectively. The
582 next day the cultures were diluted 100-fold into MOPS EZ rich defined medium (Teknova) with
583 0.2% glucose (Sigma-Aldrich) and the respective antibiotics and left to grow at 37 ºC and 250
584 rpm again till the OD_{600} reached 0.2. The cells were then mixed with formaldehyde (Fisher
585 Scientific) to a final concentration of 4% to fix the cells.

586 Following the formaldehyde fixation, the cells were incubated at room temperature for 30
587 minutes and subsequently centrifuged at 3214 x g for 10 minutes at room temperature. The
588 pellets were resuspended in 200 µl 1X PBS and then washed 3 times, each time performing
589 centrifugation at 600 x g for 4 minutes and resuspending in 200 µl 1X PBS. The cells were then
590 permeabilized with 70% ethanol, shaken at room temperature for 1 hour and stored at 4 ºC prior
591 to fluorescence *in situ* hybridization.

592 **Single-molecule fluorescence *in situ* hybridization (smFISH)**

593 Stellaris Probe Designer was used to design the smFISH probes and they were ordered from
594 Biosearch Technologies (https://www.biosearchtech.com/). The probe labeling was performed
595 by using equal volumes of each probe. The final volume of sodium bicarbonate was adjusted to

596 0.1 M by adding 1/9 reaction volume of 1 M sodium bicarbonate (pH = 8.5). 0.05-0.25 mg of
597 Alexa Fluor 647 succinimidyl ester (Life Technologies) or CF 568 succinimidyl ester (Biotium)
598 dissolved in 5 µl DMSO was mixed with the probe solution. The dyes were kept at a molar
599 excess of 20-25 fold relative to the probes. The reaction mixture was incubated in the dark at 37
600 ºC with gentle vortexing overnight. The following day the reaction was quenched by using 1/9
601 reaction volume of 3 M sodium acetate (pH = 5). Ethanol precipitation followed by P-6 Micro
602 Bio-Spin Columns (Bio-Rad) were employed to remove unconjugated dyes.

603 60 µl of permeabilized cells were centrifuged at 600 x g for 4 minutes and the pellets were
604 washed with FISH wash solution (10% formamide in 2X Saline Sodium Citrate (SSC) buffer).
605 They were then resuspended along with the probes in 15 µl of FISH hybridization buffer (10%
606 dextran sulfate (Sigma-Aldrich), 1 mg/ml *E. coli* tRNA (Sigma-Aldrich), 0.2 mg/ml Bovine
607 Serum Albumin (BSA) (NEB), 2 mM vanadyl ribonucleoside complexes (Sigma-Aldrich), 10%
608 formamide (Fisher Scientific) in 2X SSC). The number of probes used for sRNA SgrS was 9,
609 they were labeled with Alexa Fluor 647 and the concentration of the labeled probes was 50 nM.
610 The number of probes used for *ptsG* mRNA was 28, they were labeled with CF 568 and the
611 labeled probe concentration was 15 nM. The reaction mixtures were incubated in the dark at 30
612 ºC overnight. The following day, the cells were suspended in 20X volume FISH wash solution
613 and centrifuged. They were resuspended in FISH wash solution, incubated at 30 ºC for 30
614 minutes and centrifuged, and this was repeated 3 times. After the final washing step, the cells
615 were pelleted and resuspended in 20 µl 4X SSC and stored at 4 ºC prior to imaging.

616 **Single-molecule localization-based super-resolution imaging**

617 The labeled cells were immobilized on 1.0 borosilicate chambered coverglass (Thermo Scientific
618 Nunc Lab-Tek) treated with poly-L-lysine (Sigma-Aldrich) and imaged with imaging buffer (50
619 mM Tris-HCl (pH = 8.0), 10% glucose (Sigma-Aldrich), 1% β-mercaptoethanol (Sigma-
620 Aldrich), 0.5 mg/ml glucose oxidase (Sigma-Aldrich) and 0.2% catalase (Sigma-Aldrich) in 2X
621 SSC).

622 3D super-resolution imaging was performed using an Olympus IX-71 inverted microscope with a
623 100X NA 1.4 SaPo oil immersion objective. Sapphire 568-100 CW CDRH (568 nm) (Coherent)
624 and DL-640-100-AL-O (647 nm) (Crystalaser) were used for two-color imaging and DL405-025
625 (405 nm) (Crystalaser) was used for the reactivation of the dyes. The laser excitation was
626 controlled by mechanical shutters (LS6T2, Uniblitz). The laser lines were reflected to the
627 objective using a dichroic mirror (Di01-R405/488/561/635, Semrock) The emission signal was
628 collected by the objective and then they passed through an emission filter (FF01-594/730-25,
629 Semrock for Alexa Fluor 647 or HQ585/70M 63061, Chroma for CF 568) and the excitation
630 laser was cleaned using notch filters (ZET647NF, Chroma; NF01-568/647-25x5.0, Semrock and
631 NF01-568U-25, Semrock). The images were captured on a 512x512 Andor EMCCD camera
632 (DV887ECS-BV, Andor Tech). 3D imaging was achieved by introducing astigmatism using a
633 cylindrical lens with focal length 2 m (SCX-50.8-1000.0-UV-SLMF-520-820, CVI Melles Griot)
634 in the emission path between two relay lenses of focal lengths 100 mm and 150 mm. Each pixel
635 corresponded to 100 nm in this setup. The z-drift of the setup was controlled by the CRISP
636 (Continuous Reflective Interface Sample Placement) system (ASI) and the region of interest for
637 imaging was selected using an xy-sample stage (BioPrecision2, Ludl Electronic Products). The
638 storm-control software written in Python by Zhuang's group and available at GitHub
639 (https://github.com/ZhuangLab/storm-control) was used for image acquisition.

640 After acquiring a DIC image of the sample area, two-color super-resolution imaging was
641 performed. 568 nm laser excitation was used for CF 568 after completing the image acquisition
642 for Alexa Fluor 647 using 647 nm laser excitation. Fluorophore bleaching was compensated and
643 moderate signal density was maintained by increasing the 405 nm laser power slowly. Imaging
644 was completed when most of the fluorophores had photobleached and the highest reactivation
645 laser power was reached.

646 Fluorescent nanodiamonds (140 nm diameter, Sigma Aldrich) were utilized for mapping of the
647 two channels. These nanodiamonds nonspecifically attached to the surface of the imaging
648 chambers and were excited by both 647 nm and 568 nm lasers. They generated localization spots
649 in the final reconstructed images that was used for mapping.

650 **Image Analysis**

651 The raw data was acquired using the Python-based acquisition software and it was analyzed
652 using a data analysis algorithm which was based on work published previously by Zhuang's
653 group.[69,70] The peak identification and fitting were performed using the method described
654 before.[34] The z-stabilization was done by the CRISP system and the horizontal drift was
655 calculated using Fast Fourier Transformation (FFT) on the reconstructed images of subsets of the
656 super-resolution image, comparing the center of the transformed images and corrected using
657 linear interpolation.

658 **Clustering Analysis and copy number calculation**

659 A density-based clustering analysis algorithm (DBSCAN) was employed to calculate the RNA
660 copy numbers. The algorithm used was the same as previously published,[34] but the Nps and Eps
661 values were updated for the SgrS and *ptsG* images, since, we used CF 568 instead of Alexa Fluor

662 568 and we also used a different 405 nm laser to reactivate the dyes. The SgrS (9 probes labeled
663 with Alexa Fluor 647) images were clustered using Nps = 3 and Eps = 15 and the *ptsG* (28
664 probes labeled with CF 568) images were clustered using Nps = 10 and Eps = 25 and these
665 numbers were empirically chosen. A MATLAB code was used as before for the cluster analysis.

666 Δ*sgrS* and Δ*ptsG* strains were grown, prepared, imaged and analyzed in the same manner as
667 before and they were used for the measurement of the background signal due to the non-specific
668 binding of Alexa Fluor 647 and CF 568.

669 The SgrS image with no αMG induction for the wild-type *E. coli* cells (DJ480) was considered to
670 be the low SgrS copy number sample, where it was assumed that one cluster was equivalent to
671 one RNA and the *ptsG* image with 20 minute αMG induction for the wild-type *E. coli* cells was
672 considered to be the low *ptsG* copy number sample. The copy numbers of the RNAs were
673 calculated in the same manner using MATLAB codes as described previously.[34]

674 **Colocalization analysis**

675 To calculate the copy number of SgrS-*ptsG* complexes, colocalization analysis was performed in
676 order to calculate the percentage of *ptsG* colocalized with SgrS. The average radius of a *ptsG*
677 mRNA cluster was calculated to be around 40 nm. That value was used as the radius to consider
678 a 3D spherical volume from the center of the *ptsG* cluster. The SgrS spots corresponding to
679 clusters found in this volume were taken to be colocalized with the *ptsG* cluster. The base-
680 pairing mutant strain was considered a negative control (Supplementary Fig. 41a) and percentage
681 of colocalization was plotted against SgrS copy number and fit with a line (y = a*x) to act as a
682 calibration for colocalization by chance (Supplementary Fig. 41b). The coefficient, a, was used a

683 correction factor for colocalization calculation as, final colocalization = calculated colocalization
684 – a*SgrS copy number.

685 **SgrS and *ptsG* mRNA half-life measurements**

686 The *ptsG* mRNA degradation rates were calculated using a rifampicin-chase experiment. The
687 wild-type (DJ480) *E. coli* cells, the SgrS A177U, G178A, G178U, U181A, U182A, G184A,
688 G184A-C195U, G215A, U224A, and U224G were grown in LB Broth with the respective
689 antibiotics at 37 ºC, 250 rpm overnight. The following day, the overnight cultures were diluted
690 100-fold in MOPS EZ rich defined medium supplemented with 0.2% glucose and they were
691 grown at 37 ºC, 250 rpm. When the OD_{600} reached 0.15-0.25 rifampicin (Sigma-Aldrich) was
692 added to a final concentration of 500 µg/ml. This was taken as the 0-minute time point for the
693 experiment and aliquots were taken at 2, 4, 6, 8, 10, 15, 20 minutes after the addition of
694 rifampicin and fixed in the same manner described before. The cells were labeled by FISH
695 probes, imaged and analyzed by the same process mentioned. The natural logs of the copy
696 numbers were plotted against time and the slope of the linear fitting was used to calculate the
697 lifetime of the RNA. The reciprocal of the lifetimes gave the degradation rates.

698 The SgrS degradation rates were calculated for the above strains and the wild-type Δ*hfq*, A177U
699 Δ*hfq*, G184A Δ*hfq* mutants by stopping the transcription of SgrS by removing αMG from the
700 media. The wild-type *E. coli* cells, the mutants and the RNase E mutants were grown overnight
701 as described before in LB Broth with the respective antibiotics. The cells were diluted the
702 following day and grown in MOPS EZ rich defined medium with the respective antibiotics till
703 OD_{600} 0.15-0.25. SgrS transcription was induced in the cells using αMG and growing them for
704 10 minutes. The cells were then washed twice with centrifugation and resuspension with cold,
705 fresh media devoid of αMG and finally resuspended in pre-warmed media at 37 ºC. Aliquots

706 were taken at 0, 2, 4, 6, 8, 10, 15, 20 minutes (0, 2, 4, 6, 8 minutes for the Δ*hfq* strains) and fixed
707 as described before. The cells were then treated, imaged and analyzed to calculate the
708 degradation rates as mentioned before.

709 **Modeling of SgrS-induced *ptsG* mRNA degradation**

710 Kinetic model and experimental measurements of the parameters

711 The mass-action equations used for the wild-type *E. coli* cells and the chromosomal mutations
712 are shown below:

713 $$\frac{d[p]}{dt} = \alpha_p - \beta_p[p] - k_{on}[S][p] + k_{off}[Sp] \qquad (3)$$

714 $$\frac{d[S]}{dt} = \alpha_S - \beta_{S,p}[S] - k_{on}[S][p] + k_{off}[Sp] \qquad (4)$$

715 $$\frac{d[Sp]}{dt} = k_{on}[S][p] - k_{off}[Sp] - k_{cat}[Sp] \qquad (5)$$

716 In the above equations, the changes in the concentration of *ptsG*, SgrS and the SgrS-*ptsG*
717 complex over time are shown. α_p, α_S are the transcription rates of the *ptsG* mRNA and SgrS
718 respectively; $\beta_p, \beta_{S,p}$ are respectively the endogenous degradation rate of ptsG mRNA and the
719 degradation rate of SgrS excluding the co-degradation with *ptsG* mRNA; k_{on}, k_{off} are the rates
720 of association and dissociation of SgrS and *ptsG* mRNA and k_{cat} is the RNase E-mediated co-
721 degradation of SgrS-*ptsG* complex.

722 We calculated the endogenous degradation rate of *ptsG* mRNA (β_p) of the wild-type *E. coli*,
723 chromosomal mutations and the RNase E mutants from the super-resolution imaging and
724 analysis. The degradation rate of SgrS for the cells were calculated by stopping the transcription
725 of SgrS, but this method takes into account target-dependent and target-independent degradation

726 ($\beta_{S,total}$). We also calculated the degradation rate for the respective RNase E mutant strains and
727 this measurement gave us target-independent degradation and other RNase E–independent
728 degradation (β_{S0}). These two values provided a higher and lower bound for the endogenous
729 degradation rate of SgrS ($\beta_{S,p}$).

730 The transcription rate of *ptsG* mRNA was calculated using $\alpha_p = \beta_p \times [p]_0$ and in this equation
731 $[p]_0$ is the concentration of *ptsG* mRNA before the induction of sugar stress in all of the cases.
732 This was done because it was observed previously[34] that the *ptsG* mRNA reached an equilibrium
733 in the cells without SgrS-induced degradation. We calculated this for all the cases, viz., wild-type
734 *E. coli*, SgrS mutants and the RNase E mutants and the transcription rate of *ptsG* mRNA did not
735 show any significant change.

736 RNase E mutant cells are not able to degrade SgrS-*ptsG* complex efficiently, but it is a
737 possibility that the complex can degrade endogenously or via other minor degradation pathways.
738 We kept k_{cat} as a fitting parameter and used the measured parameters, α_p, β_p, $\beta_{S,total}$and β_{S0}
739 and the above equations to fit the time courses for all the strains to estimate the 5 parameters; α_S,
740 $\beta_{S,p}$, k_{on}, k_{off} and k_{cat}.

741 <u>Parameter search</u>

742 Poisson weighting (total sum of the squares, $SS_{tot} = \Sigma_i(y_i - \bar{y})^2$ and residual sum of the
743 squares, $SS_{res} = \Sigma_i(y_i - f_i)^2$, where y_i is the experimental data and f_i is the fitted data) was
744 used in the fitting of global R^2 according to the equation:

745 $$R^2 \equiv 1 - \frac{SS_{res}}{SS_{tot}} \qquad (6)$$

746 so that no bias was introduced for a particular species. The parameters were selected to maximize
747 the global R^2 for the time course curves of each of the species. The concentrations of the SgrS-
748 *ptsG* complex in all the strains were very close to the background and as a result the total
749 variance became small. R^2 was not helpful to estimate the quality of the fit in these cases.
750 Instead, χ^2's were calculated as

751 $$\chi^2 \equiv \sum_i \frac{(y_i - f_i)^2}{f_i} \qquad (7)$$

752 for all those cases and the significance levels (α) were reported.

753

754 **References**

755 1. Malkova, A. & Ira, G. Break-induced replication: Functions and molecular mechanism.
756 *Current Opinion in Genetics and Development* (2013). doi:10.1016/j.gde.2013.05.007

757 2. Geisler, S. & Coller, J. RNA in unexpected places: Long non-coding RNA functions in
758 diverse cellular contexts. *Nature Reviews Molecular Cell Biology* (2013).
759 doi:10.1038/nrm3679

760 3. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by Small RNAs in Bacteria:
761 Expanding Frontiers. *Molecular Cell* **43**, 880–891 (2011).

762 4. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems
763 in bacteria and archaea. *Nature* (2012). doi:10.1038/nature10886

764 5. Askari, F. K. & McDonnell, W. M. Antisense-Oligonucleotide Therapy. *N. Engl. J. Med.*
765 **334**, 316–318 (1996).

766 6. Singh, D. *et al.* Real-time observation of DNA target interrogation and product release by
767 the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proc. Natl. Acad. Sci. U. S. A.*
768 **115**, 5444–5449 (2018).

769 7. Singh, D., Sternberg, S. H., Fei, J., Doudna, J. A. & Ha, T. Real-time observation of DNA
770 recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* **7**, 1–8
771 (2016).

772 8. Globyte, V., Lee, S. H., Bae, T., Kim, J. & Joo, C. CRISPR /Cas9 searches for a
773 protospacer adjacent motif by lateral diffusion . *EMBO J.* **38**, (2019).

774 9. Globyte, V., Kim, S. H. & Joo, C. Single-Molecule View of Small RNA–Guided Target
775 Search and Recognition. *Annu. Rev. Biophys.* **47**, 569–593 (2018).

776 10. Ragunathan, K., Liu, C. & Ha, T. RecA filament sliding on DNA facilitates homology
777 search. *Elife* **2012**, (2012).

778 11. Lee, J. Y. *et al.* Base triplet stepping by the Rad51/RecA family of recombinases. *Science*
779 *(80-.).* **349**, 977–981 (2015).

780 12. Qi, Z. *et al.* DNA sequence alignment by microhomology sampling during homologous
781 recombination. *Cell* **160**, 856–869 (2015).

782 13. Jones, D. L. *et al.* Kinetics of dCas9 target search in Escherichia coli. *Science (80-.).* **357**,
783 1420–1424 (2017).

784 14. Richards, G. R. & Vanderpool, C. K. Molecular call and response: The physiology of
785 bacterial small RNAs. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*
786 **1809**, 525–531 (2011).

787 15. Desnoyers, G., Bouchard, M. P. & Massé, E. New insights into small RNA-dependent
788 translational regulation in prokaryotes. *Trends in Genetics* **29**, 92–98 (2013).

789 16. Massé, E., Vanderpool, C. K. & Gottesman, S. Effect of RyhB Small RNA on Global Iron
790 Use in Escherichia coli. *J. Bacteriol.* **187**, 6962–6971 (2005).

791 17. Lease, R. A., Smith, D., McDonough, K. & Belfort, M. The small noncoding DsrA RNA
792 is an acid resistance regulator in Escherichia coli. *J. Bacteriol.* **186**, 6179–6185 (2004).

793 18. Domenech, P., Honoré, N., Heym, B. & Cole, S. T. Role of OxyS of Mycobacterium
794 tuberculosis in oxidative stress: Overexpression confers increased sensitivity to organic
795 hydroperoxides. *Microbes Infect.* **3**, 713–721 (2001).

796 19. Zhang, A. *et al.* Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.*
797 **50**, 1111–1124 (2003).

798 20. Bobrovskyy, M. & Vanderpool, C. K. The small RNA SgrS: Roles in metabolism and
799 pathogenesis of enteric bacteria. *Frontiers in Cellular and Infection Microbiology* **4**,
800 (2014).

801 21. Richards, G. R., Patel, M. V., Lloyd, C. R. & Vanderpool, C. K. Depletion of glycolytic
802 intermediates plays a key role in glucose-phosphate stress in escherichia coli. *J. Bacteriol.*
803 **195**, 4816–4825 (2013).

804 22. Bobrovskyy, M. *et al.* Determinants of target prioritization and regulatory hierarchy for
805 the bacterial small RNA SgrS. *Mol. Microbiol.* (2019). doi:10.1111/mmi.14355

806 23. Rice, J. B. & Vanderpool, C. K. The small RNA SgrS controls sugar-phosphate
807 accumulation by regulating multiple PTS genes. *Nucleic Acids Res.* **39**, 3806–3819

808 (2011).

809 24. Kawamoto, H., Koide, Y., Morita, T. & Aiba, H. Base-pairing requirement for RNA
810 silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol.*
811 *Microbiol.* **61**, 1013–1022 (2006).

812 25. Panja, S. & Woodson, S. A. Hfq proximity and orientation controls RNA annealing.
813 *Nucleic Acids Res.* (2012). doi:10.1093/nar/gks618

814 26. Zuzanna, W. & Mikolaj, O. Hfq assists small RNAs in binding to the coding sequence of
815 ompD mRNA and in rearranging its structure. *RNA* **22**, 979–994 (2016).

816 27. Ishikawa, H., Otaka, H., Maki, K., Morita, T. & Aiba, H. The functional Hfq-binding
817 module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich
818 sequence and followed by a 3′ poly(U) tail. *RNA* **18**, 1062–1074 (2012).

819 28. Morita, T., Nishino, R. & Aiba, H. Role of the terminator hairpin in the biogenesis of
820 functional Hfq-binding sRNAs. *RNA* **23**, 1419–1431 (2017).

821 29. Otaka, H., Ishikawa, H., Morita, T. & Aiba, H. PolyU tail of rho-independent terminator
822 of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U. S. A.* **108**,
823 13059–13064 (2011).

824 30. Horler, R. S. P. & Vanderpool, C. K. Homologs of the small RNA SGRS are broadly
825 distributed in enteric bacteria but have diverged in size and sequence. *Nucleic Acids Res.*
826 **37**, 5465–5476 (2009).

827 31. Maki, K., Uno, K., Morita, T. & Aiba, H. RNA, but not protein partners, is directly
828 responsible for translational silencing by a bacterial Hfq-binding small RNA. *Proc. Natl.*

829 *Acad. Sci. U. S. A.* **105**, 10332–10337 (2008).

830 32. Soper, T. J. & Woodson, S. A. The rpoS mRNA leader recruits Hfq to facilitate annealing
831 with DsrA sRNA. *RNA* (2008). doi:10.1261/rna.1110608

832 33. Maki, K., Morita, T., Otaka, H. & Aiba, H. A minimal base-pairing region of a bacterial
833 small RNA SgrS required for translational repression of ptsG mRNA. *Mol. Microbiol.* **76**,
834 782–792 (2010).

835 34. Fei, J. *et al.* Determination of in vivo target search kinetics of regulatory noncoding RNA.
836 *Science (80-.).* **347**, 1371–1374 (2015).

837 35. Arluison, V. *et al.* Spectroscopic observation of RNA chaperone activities of Hfq in post-
838 transcriptional regulation by a small non-coding RNA. *Nucleic Acids Res.* (2007).
839 doi:10.1093/nar/gkl1124

840 36. Hwang, W., Arluison, V. & Hohng, S. Dynamic competition of DsrA and rpoS fragments
841 for the proximal binding site of Hfq as a means for efficient annealing. *Nucleic Acids Res.*
842 (2011). doi:10.1093/nar/gkr075

843 37. Rutherford, S. T., Valastyan, J. S., Taillefumier, T., Wingreen, N. S. & Bassler, B. L.
844 Comprehensive analysis reveals how single nucleotides contribute to noncoding RNA
845 function in bacterial quorum sensing. *Proc. Natl. Acad. Sci. U. S. A.* **112**, (2015).

846 38. Kinney, J. B., Murugan, A., Callan, C. G. & Cox, E. C. Using deep sequencing to
847 characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc.*
848 *Natl. Acad. Sci. U. S. A.* **107**, 9158–9163 (2010).

849 39. Peterman, N., Lavi-Itzkovitz, A. & Levine, E. Large-scale mapping of sequence-function

850 relations in small regulatory RNAs reveals plasticity and modularity. *Nucleic Acids Res.*
851 **42**, 12177–12188 (2014).

852 40. Bobrovskyy, M. & Vanderpool, C. K. Diverse mechanisms of post-transcriptional
853 repression by the small RNA regulator of glucose-phosphate stress. *Mol. Microbiol.* **99**,
854 254–273 (2016).

855 41. Wadler, C. S. & Vanderpool, C. K. A dual function for a bacterial small RNA: SgrS
856 performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc.*
857 *Natl. Acad. Sci.* **104**, 20454–20459 (2007).

858 42. Balasubramanian, D. & Vanderpool, C. K. Deciphering the interplay between two
859 independent functions of the small RNA regulator SgrS in Salmonella. *J. Bacteriol.*
860 (2013). doi:10.1128/JB.00586-13

861 43. Morita, T., Ueda, M., Kubo, K. & Aiba, H. Insights into transcription termination of Hfq-
862 binding sRNAs of Escherichia coli and characterization of readthrough products. *RNA* **21**,
863 1490–1501 (2015).

864 44. De Lay, N., Schu, D. J. & Gottesman, S. Bacterial small RNA-based negative regulation:
865 Hfq and its accomplices. *Journal of Biological Chemistry* **288**, 7996–8003 (2013).

866 45. Sun, Y. & Vanderpool, C. K. Physiological consequences of multiple-target regulation by
867 the small RNA SgrS in escherichia coli. *J. Bacteriol.* **195**, 4804–4815 (2013).

868 46. Rice, J. B., Balasubramanian, D. & Vanderpool, C. K. Small RNA binding-site
869 multiplicity involved in translational regulation of a polycistronic mRNA. *Proc. Natl.*
870 *Acad. Sci. U. S. A.* **109**, (2012).

871 47. Wadler, C. S. & Vanderpool, C. K. Characterization of homologs of the small rna sgrs
872 reveals diversity in function. *Nucleic Acids Res.* **37**, 5477–5485 (2009).

873 48. Gottesman, S. & Storz, G. Bacterial small RNA regulators: Versatile roles and rapidly
874 evolving variations. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).

875 49. Aiba, H. Mechanism of RNA silencing by Hfq-binding small RNAs. *Current Opinion in*
876 *Microbiology* **10**, 134–139 (2007).

877 50. Sauer, E. & Weichenrieder, O. Structural basis for RNA 3′-end recognition by Hfq. *Proc.*
878 *Natl. Acad. Sci. U. S. A.* **108**, 13065–13070 (2011).

879 51. Panja, S., Schu, D. J. & Woodson, S. A. Conserved arginines on the rim of Hfq catalyze
880 base pair formation and exchange. *Nucleic Acids Res.* (2013). doi:10.1093/nar/gkt521

881 52. Panja, S., Paul, R., Greenberg, M. M. & Woodson, S. A. Light-Triggered RNA Annealing
882 by an RNA Chaperone. *Angew. Chemie Int. Ed.* **54**, 7281–7284 (2015).

883 53. Mann, M., Wright, P. R. & Backofen, R. IntaRNA 2.0: Enhanced and customizable
884 prediction of RNA-RNA interactions. *Nucleic Acids Res.* **45**, W435–W439 (2017).

885 54. Wright, P. R. *et al.* CopraRNA and IntaRNA: Predicting small RNA targets, networks and
886 interaction domains. *Nucleic Acids Res.* **42**, (2014).

887 55. Busch, A., Richter, A. S. & Backofen, R. IntaRNA: Efficient prediction of bacterial sRNA
888 targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**, 2849–
889 2856 (2008).

890 56. Raden, M. *et al.* Freiburg RNA tools: A central online resource for RNA-focused research
891 and teaching. *Nucleic Acids Res.* **46**, W25–W29 (2018).

892 57. Morita, T., Maki, K. & Aiba, H. RNase E-based ribonucleoprotein complexes: Mechanical
893 basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* **19**,
894 2176–2186 (2005).

895 58. Bruce, H. A. *et al.* Analysis of the natively unstructured RNA/protein-recognition core in
896 the Escherichia coli RNA degradosome and its interactions with regulatory RNA/Hfq
897 complexes. *Nucleic Acids Res.* **46**, 387–402 (2018).

898 59. Caillet, J., Baron, B., Boni, I. V., Caillet-Saguy, C. & Hajnsdorf, E. Identification of
899 protein-protein and ribonucleoprotein complexes containing Hfq. *Sci. Rep.* **9**, 1–12 (2019).

900 60. Updegrove, T. B., Zhang, A. & Storz, G. Hfq: The flexible RNA matchmaker. *Current*
901 *Opinion in Microbiology* **30**, 133–138 (2016).

902 61. Schu, D. J., Zhang, A., Gottesman, S. & Storz, G. Alternative Hfq- sRNA interaction
903 modes dictate alternative mRNA recognition . *EMBO J.* (2015).
904 doi:10.15252/embj.201591569

905 62. Dimastrogiovanni, D. *et al.* Recognition of the small regulatory RNA RydC by the
906 bacterial Hfq protein. *Elife* **3**, (2014).

907 63. Geissmann, T. A. & Touati, D. Hfq, a new chaperoning role: Binding to messenger RNA
908 determines access for small RNA regulator. *EMBO J.* **23**, 396–405 (2004).

909 64. Peng, Y., Curtis, J. E., Fang, X. & Woodson, S. A. Structural model of an mRNA in
910 complex with the bacterial chaperone Hfq. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17134–
911 17139 (2014).

912 65. Storz, G., Opdyke, J. A. & Zhang, A. Controlling mRNA stability and translation with

913 small, noncoding RNAs. *Current Opinion in Microbiology* **7**, 140–144 (2004).

914 66. Valentin-Hansen, P., Eriksen, M. & Udesen, C. The bacterial Sm-like protein Hfq: A key
915 player in RNA transactions. *Molecular Microbiology* **51**, 1525–1533 (2004).

916 67. Miller, J. H. *Experiments in molecular genetics*. ((Cold Spring Harbor, N.Y.: Cold Spring
917 Harbor Laboratory, 1972).

918 68. Yu, D. *et al.* An efficient recombination system for chromosome engineering in
919 Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5978–5983 (2000).

920 69. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical
921 reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–5 (2006).

922 70. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution
923 imaging by stochastic optical reconstruction microscopy. *Science (80-.).* **319**, 810–813
924 (2008).

925

926 **Acknowledgements**

927 We would like to thank Erel Levine, Divya Balasubramanian, D. Jin for plasmids and strains.
928 We thank Hao Zhang at the Cell Sorting Core Facility (Bloomberg School of Public Health) for
929 helping us with the flow cytometry and sorting experiments. We appreciate and thank Prof.
930 Sarah Woodson for going through the manuscript and providing insightful suggestions. This
931 work was supported by grants from National Institutes of Health R01 GM112659 (M.B., M.S.A.,
932 T.H., J.Z., and A.P.), R35 GM122569 (T.H., J.Z., and A.P.), National Science Foundation PHY
933 1430124 (T.H., J.Z., and A.P). T.H. is an investigator with the Howard Hughes Medical Institute.

934

935 **Author Contributions**

936 A.P., C.K.V., T.H. designed the experiments, with help from M.B., J.Z. A.P., T.K., J.Z.
937 performed the Sort-Seq experiments and the Sort-Seq data was analyzed by A.P., T.K., P.L.
938 M.B. performed the β-galactosidase and Northern blot experiments. M.S.A., X.M. made the
939 strains that were used in this work. M.S.A., X.M. performed the qPCR experiments to calculate
940 RNA lifetimes. A.P. performed all super-resolution imaging experiments, with some help from
941 J.Z. A.P. performed the analysis for the imaging experiments with the MATLAB package
942 written by D.S., J.F. A.P., Z.L.S, C.K.V., T.H. discussed the data. A.P., T.K., M.S.A., X.M., J.F.,
943 C.K.V., T.H. wrote the manuscript.

944

945 **Competing Interests**

946 The authors declare no competing interests.

947

948 **Figure Legends**

949 **Figure 1. Target search kinetics of SgrS. (a)** Kinetic scheme of *ptsG* mRNA degradation
950 induced by wild-type SgrS sRNA and the different point mutants of SgrS. The figure shows one
951 of the SgrS mutant strains, A177U. The steps are described in detail in the main text and the inset
952 shows the mutants used in this study. [*p*], [*S*] and [*Sp*] are the concentrations of *ptsG* mRNA,
953 SgrS and the SgrS-*ptsG* complex, respectively, in their mass-action equations. **(b)** Secondary
954 structure of SgrS sRNA from nucleotide 168 to the poly-U tail. **(c)** Base-pairing interaction

955 between SgrS and *ptsG* mRNA showing the complementary region, start site and ribosome
956 binding site.

957 **Figure 2. Mapping efficacy of SgrS regulation of ptsG mRNA with respect to its sequence.**
958 **(a)** Preparation of SgrS mutation library. Mutations are introduced into the SgrS plasmid using
959 mutagenesis PCR. This library is then transformed into an *E. coli* strain already transformed with
960 the *ptsG* mRNA plasmid fused with a GFP reporter. **(b)** Sorting of the cells and sequencing. The
961 cells with two-plasmid co-expression were sorted using flow cytometry. The SgrS library (blue)
962 shows GFP fluorescence that spans the region from the wild-type SgrS (black) to target (*ptsG*)-
963 only (green). Cells were sorted into 5 evenly spaced (log scale) fluorescence bins and the
964 occupancy percentages were 18.74%, 33.76%, 30.91%, 13.83% and 2.76% respectively. The
965 cells from each bin were amplified, DNA was purified, barcoded and sequenced using Illumina
966 sequencing platform. **(c)** Histogram of the Sort-Seq measurements from the SgrS library from
967 two replicates combined. The mean fluorescence for the wild-type SgrS is shown in red. **(d)** Heat
968 map showing the effect of mutations in SgrS regulation of *ptsG* mRNA starting from nucleotide
969 149 to 227. The color in the boxes are scaled from blue (low) to red (high), according to the level
970 of perturbation of SgrS regulation. Black squares represent the wild-type base at each position
971 and the black boxes with white crosses show the position of the mutants missing in the
972 experiment. Text shows the wild type sequence of SgrS. Insets show the four regions of SgrS,
973 viz. base-pairing region, small stem-loop, terminator stem-loop and the poly-U tail.

974 **Figure 3. Regulation of *ptsG'-'lacZ* translational fusion by SgrS point-mutants. (a)**
975 Regulation of chromosomal *ptsG'-'lacZ* translational fusion by wild-type SgrS and A177U,
976 G178U, G178A, G184A, G215A, U224A, U224G mutant variants (plasmid-encoded) was
977 assessed using β-galactosidase activity assay. Standard error was calculated based on data from

978 four biological replicates. **(b)** RNA was extracted simultaneously with β-galactosidase activity
979 assay and Northern blot was performed using probes specific for SgrS sRNA and 5S rRNA
980 (control). Full-length (227 nt), properly terminated SgrS transcripts are labeled as "termination"
981 products, and longer transcripts that arose due to transcriptional readthrough are labeled as
982 "readthrough" products. **(c)** Band intensities of total SgrS transcripts (termination+readthrough)
983 were measured, and 5S-normalized values were plotted as "SgrS abundance" (steady state
984 transcript abundance of SgrS mutants). **(d)** Band intensities of SgrS termination and readthrough
985 products were measured and 5S-normalized ratios (readthrough/termination) were calculated and
986 plotted for each SgrS mutant as "SgrS readthrough ratio".

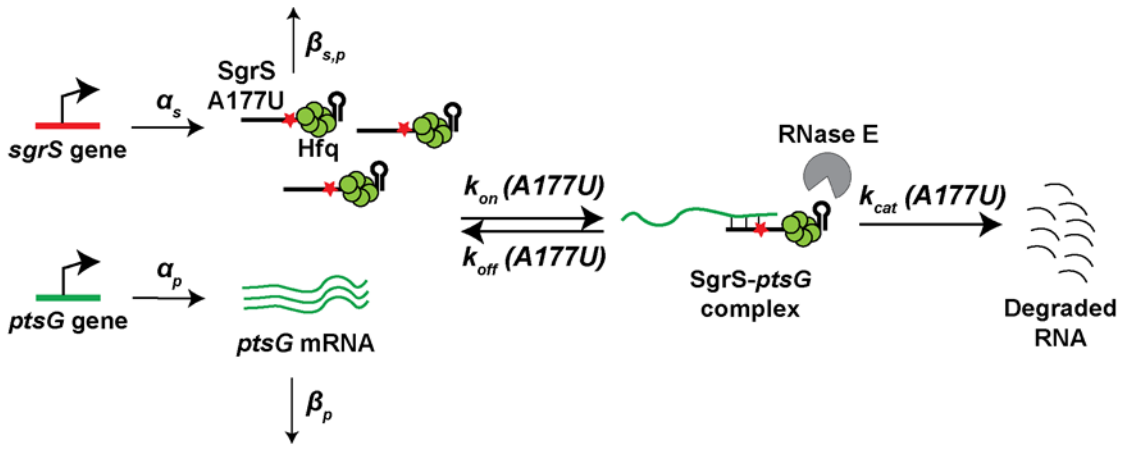
987 **Figure 4. 3D super-resolution images projected in 2D planes.** The panels show SgrS (red) and
988 *ptsG* mRNA (green) labeled by smFISH for the wild-type and the mutant strains, A177U,
989 G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G before and
990 after 20 minutes of αMG (non-metabolizable sugar analogue) induction. Cell boundaries are
991 denoted by white solid lines. Scale bar is 2 μm.

992 **Figure 5. Time dependent changes in the copy numbers of SgrS and *ptsG* mRNA and**
993 **estimation of kinetic parameters.** Time course changes and corresponding modeling curves for
994 the SgrS, *ptsG* mRNA and SgrS-*ptsG* complex in **(a)** wild-type, **(c)** G184A mutant strain, **(e)**
995 G184A-C195U mutant strain, **(g)** G215A mutant strain. Average copy numbers per cell are
996 plotted against time. Rate constants obtained for these mutants are shown in Figure 6 and in
997 Supplementary Table 3. Weighted R^2's for modeling are also reported in Supplementary Table 4.
998 Error bars in **(a)**, **(c)**, **(e)**, **(g)** are standard errors from 80 to 150 cells in each case. Histograms
999 showing the change in distribution of SgrS and *ptsG* mRNA copy numbers for **(b)** wild-type, **(d)**

1000 G184A mutant strain, **(f)** G184A-C195U mutant strain, **(h)** G215A mutant strain for 80-150 cells
1001 in each case.

1002 **Figure 6. Calculation of various parameters and correlation with Sort-Seq. (a)** Degradation
1003 rates of SgrS for the wild-type and the strains A177U, G178A, G178U, U181A, U182A, G184A,
1004 G184A-C195U, G215A, U224A, U224G, Δ*hfq* wild type, Δ*hfq* A177U, Δ*hfq* G184A for full
1005 length RNase E and RNase E mutants. Error bars represent standard deviation from two
1006 experimental replicates. **(b-d)** k_{on}, k_{off}, K_D measured from the time dependent modeling curves of
1007 the SgrS, *ptsG* mRNA and SgrS-*ptsG* mRNA complexes for the wild-type and strains A177U,
1008 G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G. These were
1009 determined simultaneously in the wild-type and RNase E mutants. Error bars report standard
1010 deviation from the independent fitting on two replicates. **(e)** Time course changes in *ptsG* mRNA
1011 for the wild-type, G184A, G184A-C195U, G215A mutant strains. Error bars represent standard
1012 errors from 80-150 cells in each case. **(f)** Fractional change in *ptsG* mRNA copy numbers for the
1013 wild-type and the mutants A177U, G178A, G178U, U181A, U182A, G184A, G184A-C195U,
1014 G215A, U224A, U224G before and after 20 minutes αMG induction. Error bars represent
1015 standard errors from 80-150 cells in each case. **(g)** Comparison of the SgrS regulation efficacy
1016 calculated from Sort-Seq assay and the imaging-based analysis. Error bars in the x-axis are
1017 standard deviations calculated from two experimental replicates and those in the y-axis are as
1018 described in **(f)**. The fitting is shown in blue and the grey region shows the 95% confidence
1019 region.

a



$$\frac{d[p]}{dt} = \alpha_p - \beta_p[p] - k_{on}[S][p] + k_{off}[Sp]$$

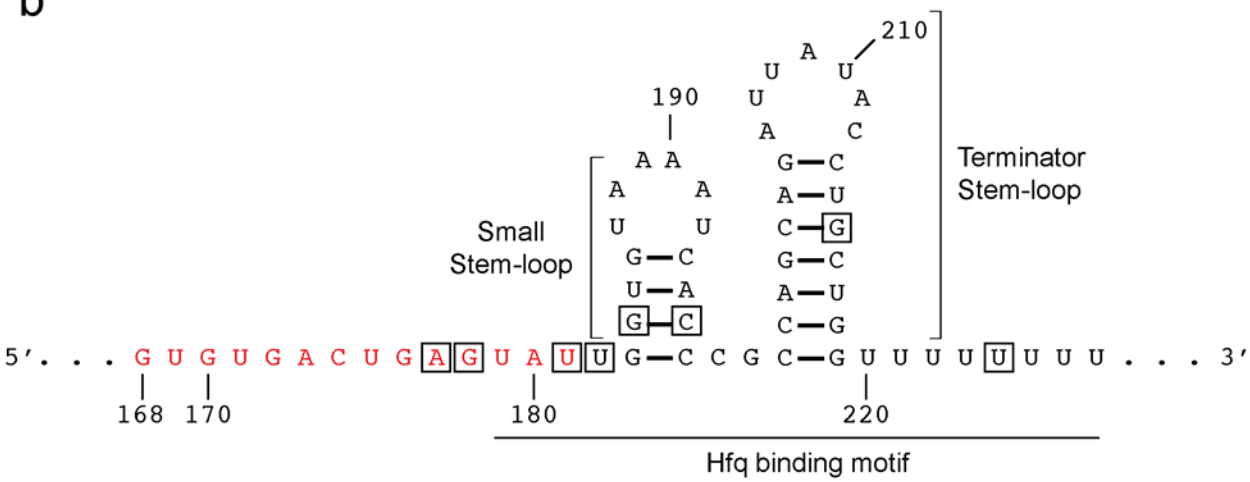
$$\frac{d[S]}{dt} = \alpha_S - \beta_{S,p}[S] - k_{on}[S][p] + k_{off}[Sp]$$

$$\frac{d[Sp]}{dt} = k_{on}[S][p] - k_{off}[Sp] - k_{cat}[Sp]$$

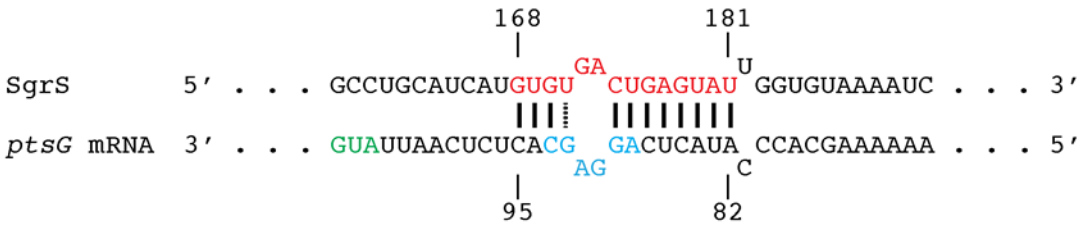
Mutants used in this study

A177U	G184A
G178A	G184A-C195U
G178U	G215A
U181A	U224A
U182A	U224G

b



c



Green nucleotides form the Start Site
Blue nucleotides form the Ribosome Binding Site

Figure 1. Target search kinetics of SgrS. (a) Kinetic scheme of *ptsG* mRNA degradation induced by wild-type SgrS sRNA and the different point mutants of SgrS. The figure shows one of the SgrS mutant strains, A177U. The steps are described in detail in the main text and the inset shows the mutants used in this study. [*p*], [*S*] and [*Sp*] are the concentrations of *ptsG* mRNA, SgrS and the SgrS-*ptsG* complex, respectively, in their mass-action equations. **(b)** Secondary structure of SgrS sRNA from nucleotide 168 to the poly-U tail. **(c)** Base-pairing interaction between SgrS and *ptsG* mRNA showing the complementary region, start site and ribosome binding site.

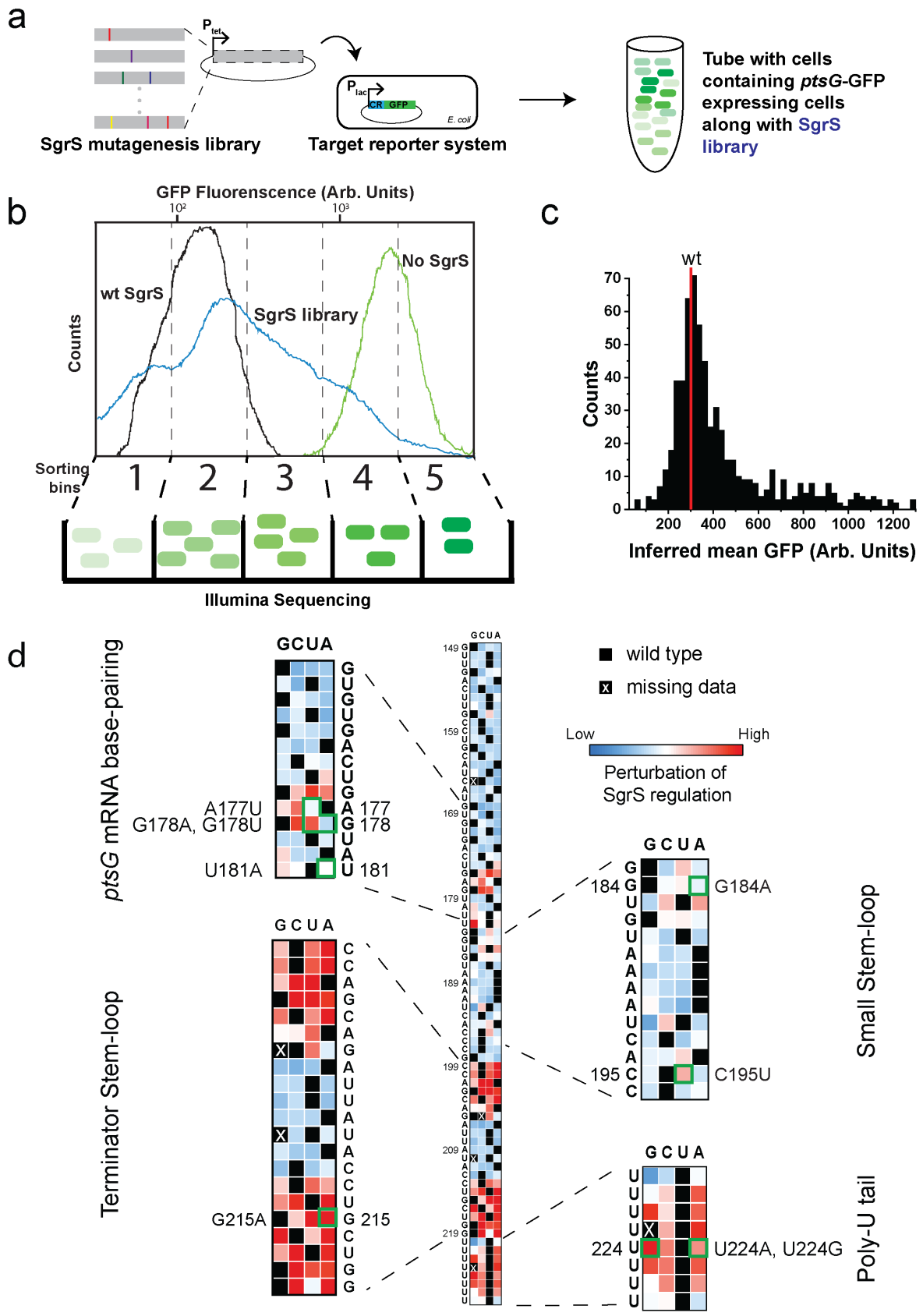

a
SgrS mutagenesis library
Target reporter system
Tube with cells containing ptsG-GFP expressing cells along with SgrS library
b
GFP Fluorescence (Arb. Units)
wt SgrS
SgrS library
No SgrS
Counts
Sorting bins
1
2
3
4
5
Illumina Sequencing
c
wt
Counts
Inferred mean GFP (Arb. Units)
d
ptsG mRNA base-pairing
Terminator Stem-loop
wild type
missing data
Low
High
Perturbation of SgrS regulation
A177U
G178A, G178U
U181A
G215A
G184A
C195U
U224A, U224G
Small Stem-loop
Poly-U tail

Figure 2. Mapping efficacy of SgrS regulation of *ptsG* mRNA with respect to its sequence. **(a)** Preparation of SgrS mutation library. Mutations are introduced into the SgrS plasmid using mutagenesis PCR. This library is then transformed into an *E. coli* strain already transformed with the *ptsG* mRNA plasmid fused with a GFP reporter. **(b)** Sorting of the cells and sequencing. The cells with two-plasmid co-expression were sorted using flow cytometry. The SgrS library (blue) shows GFP fluorescence that spans the region from the wild-type SgrS (black) to target (*ptsG*)-only (green). Cells were sorted into 5 evenly spaced (log scale) fluorescence bins and the occupancy percentages were 18.74%, 33.76%, 30.91%, 13.83% and 2.76% respectively. The cells from each bin were amplified, DNA was purified, barcoded and sequenced using Illumina sequencing platform. **(c)** Histogram of the Sort-Seq measurements from the SgrS library from two replicates combined. The mean fluorescence for the wild-type SgrS is shown in red. **(d)** Heat map showing the effect of mutations in SgrS regulation of *ptsG* mRNA starting from nucleotide 149 to 227. The color in the boxes are scaled from blue (low) to red (high), according to the level of perturbation of SgrS regulation. Black squares represent the wild-type base at each position and the black boxes with white crosses show the position of the mutants missing in the experiment. Text shows the wild type sequence of SgrS. Insets show the four regions of SgrS, viz. base-pairing region, small stem-loop, terminator stem-loop and the poly-U tail.

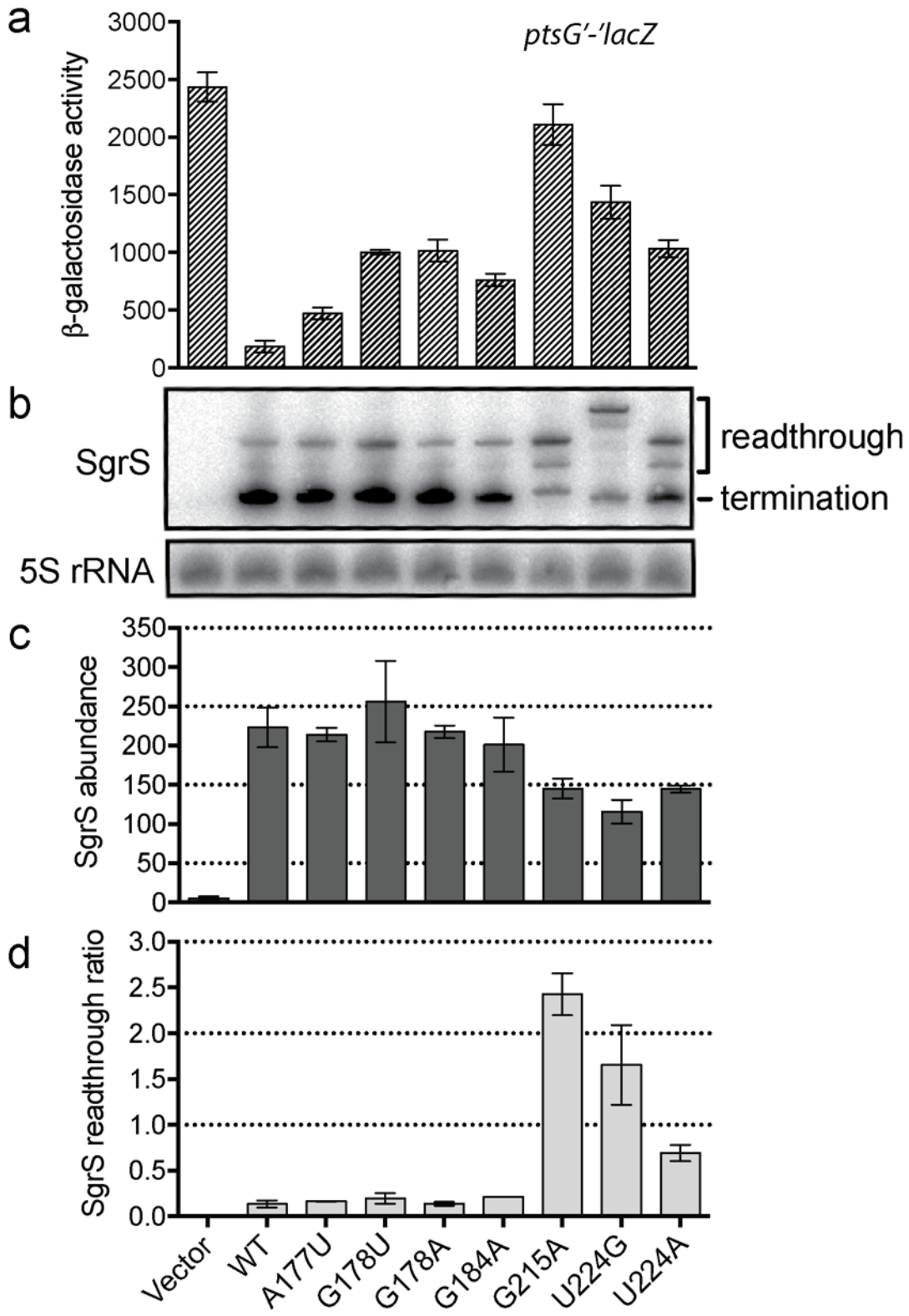

a
ptsG'-'lacZ
β-galactosidase activity
3000
2500
2000
1500
1000
500
0
b
SgrS
readthrough
termination
5S rRNA
c
SgrS abundance
350
300
250
200
150
100
50
0
d
SgrS readthrough ratio
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Vector
WT
A177U
G178U
G178A
G184A
G215A
U224G
U224A

Figure 3. Regulation of *ptsG'-'lacZ* translational fusion by SgrS point-mutants. (a) Regulation of chromosomal *ptsG'-'lacZ* translational fusion by wild-type SgrS and A177U, G178U, G178A, G184A, G215A, U224A, U224G mutant variants (plasmid-encoded) was assessed using β-galactosidase activity assay. Standard error was calculated based on data from four biological replicates. **(b)** RNA was extracted simultaneously with β-galactosidase activity assay and Northern blot was performed using probes specific for SgrS sRNA and 5S rRNA (control). Full-length (227 nt), properly terminated SgrS transcripts are labeled as "termination" products, and longer transcripts that arose due to transcriptional readthrough are labeled as "readthrough" products. **(c)** Band intensities of total SgrS transcripts (termination+readthrough) were measured, and 5S-normalized values were plotted as "SgrS abundance" (steady state transcript abundance of SgrS mutants). **(d)** Band intensities of SgrS termination and readthrough products were measured and 5S-normalized ratios (readthrough/termination) were calculated and plotted for each SgrS mutant as "SgrS readthrough ratio".

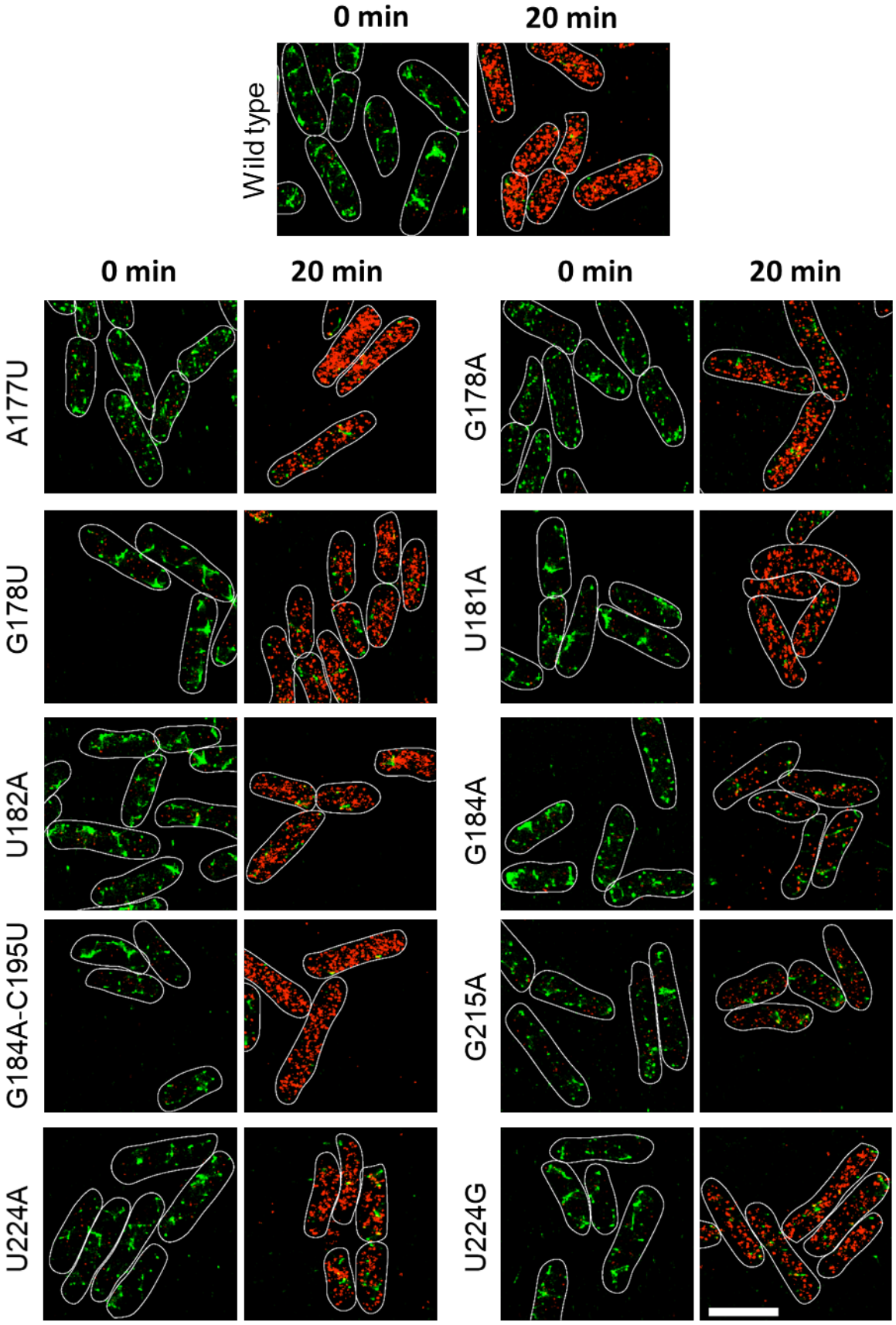

0 min
20 min
Wild type
0 min
20 min
A177U
G178U
U182A
G184A-C195U
U224A
0 min
20 min
G178A
U181A
G184A
G215A
U224G
SgrS
ptsG mRNA

Figure 4. 3D super-resolution images projected in 2D planes. The panels show SgrS (red) and *ptsG* mRNA (green) labeled by smFISH for the wild-type and the mutant strains, A177U, G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G before and after 20 minutes of αMG (non-metabolizable sugar analogue) induction. Cell boundaries are denoted by white solid lines. Scale bar is 2 μm.

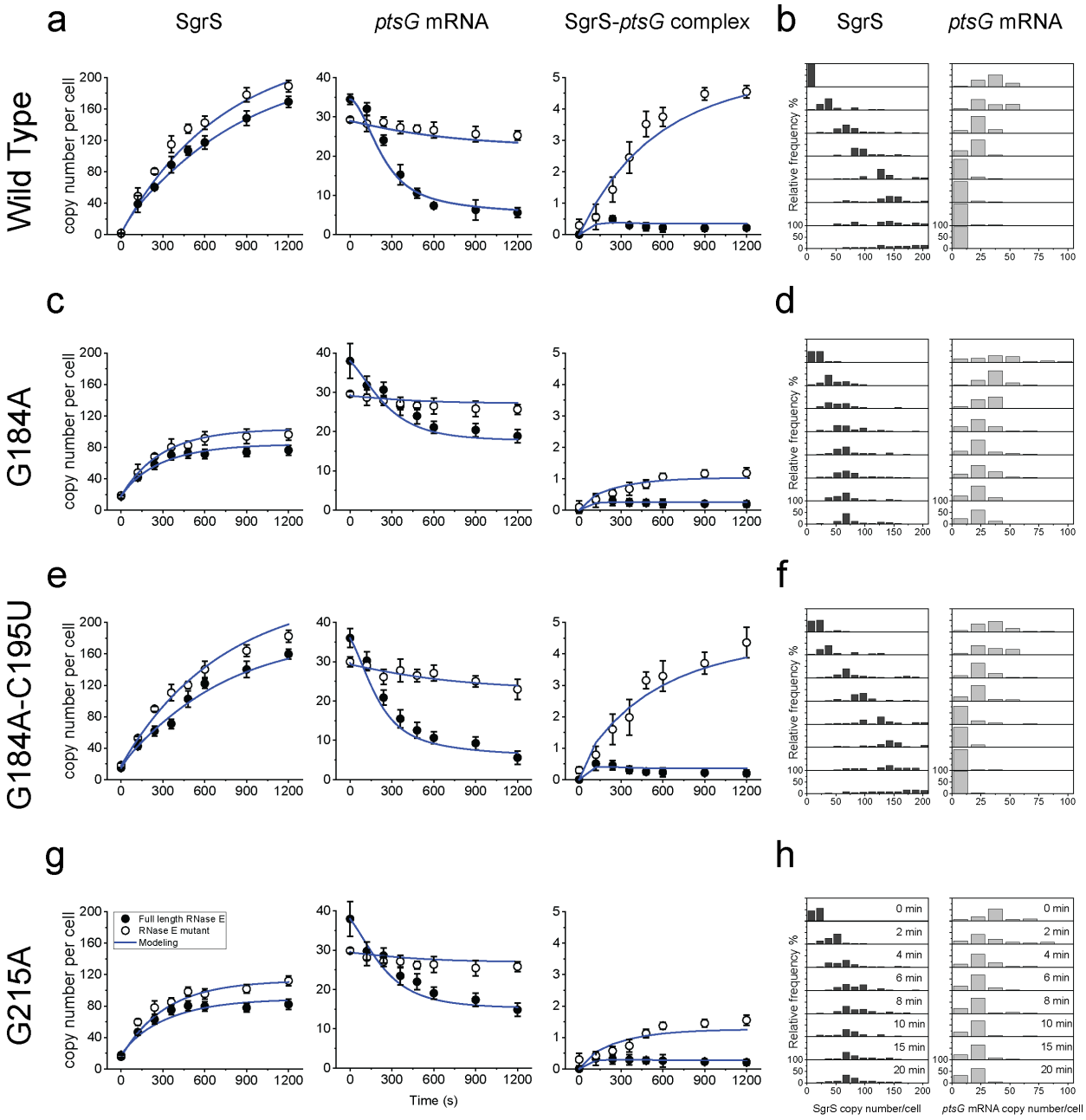


Figure 5. Time dependent changes in the copy numbers of SgrS and *ptsG* mRNA and estimation of kinetic parameters. Time course changes and corresponding modeling curves for the SgrS, *ptsG* mRNA and SgrS-*ptsG* complex in **(a)** wild-type, **(c)** G184A mutant strain, **(e)** G184A-C195U mutant strain, **(g)** G215A mutant strain. Average copy numbers per cell are plotted against time. Rate constants obtained for these mutants are shown in Figure 6 and in

Supplementary Table 3. Weighted R^2's for modeling are also reported in Supplementary Table 4. Error bars in **(a)**, **(c)**, **(e)**, **(g)** are standard errors from 80 to 150 cells in each case. Histograms showing the change in distribution of SgrS and *ptsG* mRNA copy numbers for **(b)** wild-type, **(d)** G184A mutant strain, **(f)** G184A-C195U mutant strain, **(h)** G215A mutant strain for 80-150 cells in each case.

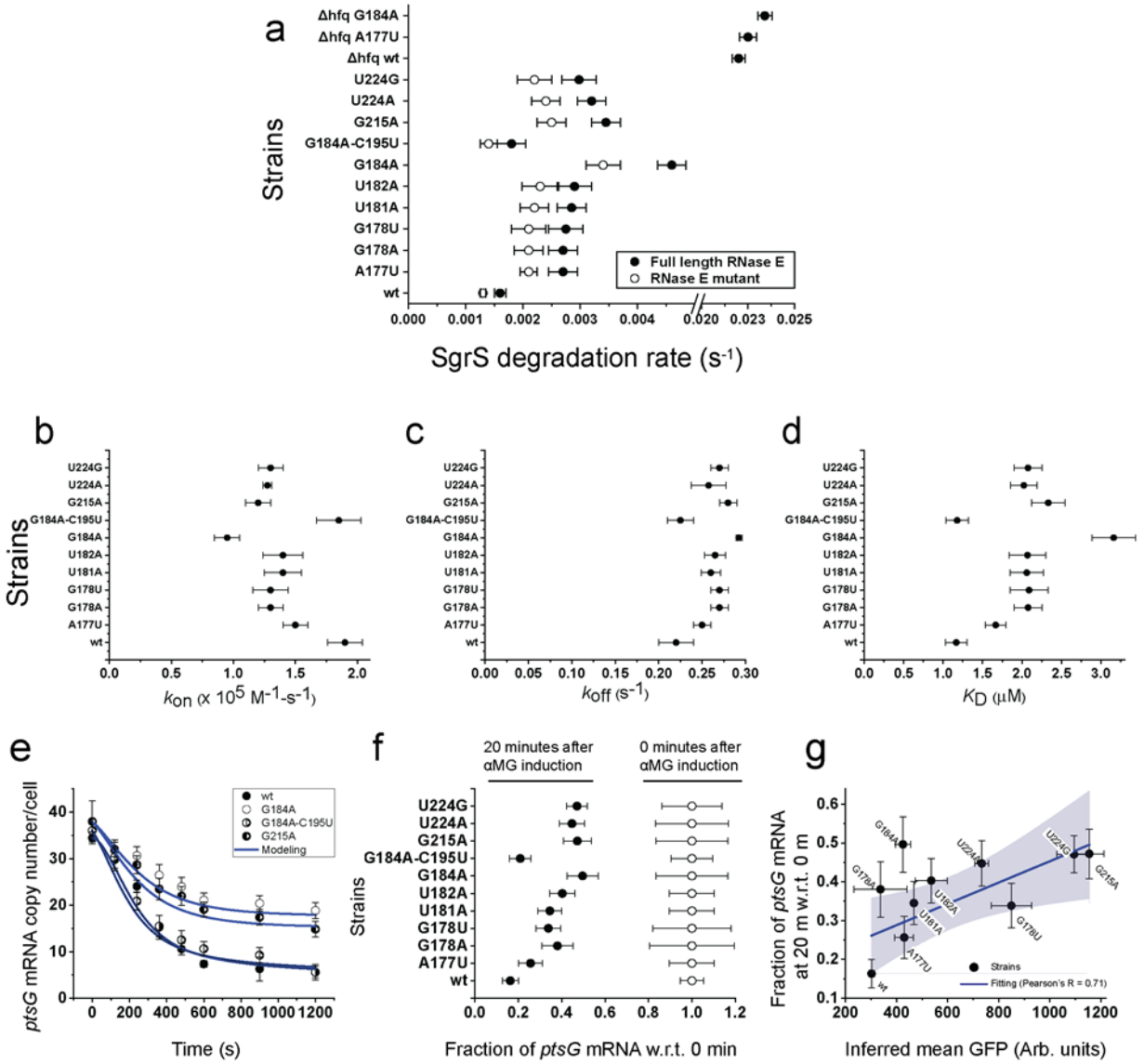


Figure 6. Calculation of various parameters and correlation with Sort-Seq. (a) Degradation rates of SgrS for the wild-type and the strains A177U, G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G, Δ*hfq* wild type, Δ*hfq* A177U, Δ*hfq* G184A for full length RNase E and RNase E mutants. Error bars represent standard deviation from two experimental replicates. **(b-d)** k_{on}, k_{off}, K_D measured from the time dependent modeling curves of the SgrS, *ptsG* mRNA and SgrS-*ptsG* mRNA complexes for the wild-type and strains A177U,

G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G. These were determined simultaneously in the wild-type and RNase E mutants. Error bars report standard deviation from the independent fitting on two replicates. **(e)** Time course changes in *ptsG* mRNA for the wild-type, G184A, G184A-C195U, G215A mutant strains. Error bars represent standard errors from 80-150 cells in each case. **(f)** Fractional change in *ptsG* mRNA copy numbers for the wild-type and the mutants A177U, G178A, G178U, U181A, U182A, G184A, G184A-C195U, G215A, U224A, U224G before and after 20 minutes αMG induction. Error bars represent standard errors from 80-150 cells in each case. **(g)** Comparison of the SgrS regulation efficacy calculated from Sort-Seq assay and the imaging-based analysis. Error bars in the x-axis are standard deviations calculated from two experimental replicates and those in the y-axis are as described in **(f)**. The fitting is shown in blue and the grey region shows the 95% confidence region.