Intracellular signaling through the *comRS* system in *Streptococcus mutans* genetic competence

Simon A.M. Underhill[1], Robert C. Shields[2], Justin R. Kaspar[2], Momin Haider[1], Robert A. Burne[2] and Stephen J. Hagen[*1]

[1]Department of Physics, University of Florida, 2001 Museum Road, Gainesville, Florida, USA,

[2] Department of Oral Biology, University of Florida, 1395 Center Drive, Gainesville, Florida, USA

*Corresponding author

Tel. 352 392 4716

Email: sjhagen@ufl.edu.

Short Title: Intracellular signaling in ComRS

Abstract

Entry into genetic competence in streptococci is controlled by ComX, an alternative sigma factor for genes that enable the import of exogenous DNA. In *Streptococcus mutans*, the immediate activator of *comX* is the ComRS signaling system, which consists of the cytosolic receptor ComR and the 7-residue signal peptide XIP, which is derived from ComS. Extracellular XIP is imported by an oligopeptide permease and interacts with ComR to form a transcriptional activator for both *comX* and *comS*, triggering *S. mutans* competence. It is not presently known how *S. mutans* produces XIP from ComS and externalizes it. The observation that *comX* is expressed bimodally under some environmental conditions suggests that ComR may also interact with endogenously produced XIP or ComS, creating an intracellular positive feedback loop in *comS* transcription. Here we use single cell and microfluidic methods to compare the effects of the native *comS* gene and extracellular XIP on *comX* expression. We find that deletion of *comS* reduces the response of *comX* to extracellular XIP. We also find that *comS*-overexpressing cells activate their *comX* even when their growth medium is rapidly exchanged, although this autoactivation requires an intact copy of *comS* under control of its own promoter. Further, *comS*-overexpressing cells did not activate *comS*-deficient mutants growing in coculture. These data show that individual cells can activate *comX* without exporting or importing the XIP or ComS signal, and that endogenously and exogenously produced ComS/XIP have inequivalent effects on *comX* behavior. These data are fully consistent with a model in which intracellular positive feedback in *comS* transcription plays a role in ComRS signaling, and is responsible for the bimodal expression of *comX*.

Author Summary

Heterogeneous gene expression in genetically identical populations plays an important role in bacterial persistence and survival under changing environmental conditions. In the oral pathogen *Streptococcus mutans,* the physiological state of genetic competence can exhibit bimodality, with only some cells becoming competent. *S. mutans* controls its entry into competence by using the ComRS signaling system to activate *comX,* a gene encoding the master regulator of competence. The ComRS system is understood as a quorum sensing system, in which the extracellular accumulation of the small signal peptide XIP, derived from ComS, induces *comX* expression. We coupled observation of bacteria that fluoresce when *comX* is active with mathematical analysis and in vitro binding assays to show that activation of *comX* does not necessarily require extracellular XIP or ComS, and that cells expressing *comX* do not necessarily export XIP. Our experiments and mathematical modeling indicate that a positive feedback loop in *comS* transcription allows a cell to activate *comX* in response to its own ComS or XIP in the absence of extracellular XIP, or to amplify its *comX* response if sufficient extracellular XIP is present. Such positive feedback loops often give rise to bimodal gene expression.

Introduction

Streptococcus mutans inhabits human oral biofilms and is a primary etiological agent of dental caries [1]. Many of the behaviors of *S. mutans* that facilitate its growth, competition, stress tolerance, and virulence are linked to the regulation of genetic competence, a transient physiological state during which the organism can internalize DNA from its environment. Biofilm formation, bacteriocin biosynthesis, tolerance of oxidative and acid stresses, carbohydrate utilization and many other behaviors of *S. mutans* interact with the pathway that controls entry into competence [2-7]. This competence pathway is complex, as it receives input from extracellular peptide signals and environmental cues such as pH [8, 9], carbohydrate source [10] and other growth conditions [11]. Competence regulation in *S. mutans* also involves regulatory feedback that can drive bimodality and other complex behaviors [11-14]. Consequently, although many elements of the *S. mutans* competence pathway have been described in detail, several key elements of the mechanisms and dynamics of regulation are not well understood.

S. mutans initiates entry into the competent state by increasing the transcription of the *comX* gene (sometimes referred to as *sigX*), which encodes an alternative sigma factor that is required for the expression of approximately 30 late competence genes [15]. Many of these genes encode products that are required for DNA uptake, processing of single-stranded DNA and homologous recombination [15, 16]. Expression of *comX* is controlled by the peptides CSP (competence stimulating peptide) and XIP (*sig*X-inducing peptide), and the efficacy of these peptides is strongly influenced by environmental conditions. CSP is derived by cleavage of a 21-residue peptide from

ComC and export through an ATP-binding cassette transporter. It is further processed to the active 18-residue peptide by the SepM protease [17]. Extracellular CSP is detected by the two-component signal transduction system ComDE, with the phosphorylated response regulator ComE activating genes for bacteriocin synthesis and immunity. ComE does not directly activate *comX*. Rather it affects *comX* indirectly via the intracellular bacteriocin CipB, in a way that is not understood [18].

In *S. mutans* and streptococci of the salivarius, bovis and pyogenic groups, the immediate regulator of *comX* is the ComRS system. ComR is an Rgg-like cytosolic transcriptional regulator and the type II ComS of *S. mutans* is a 17-residue peptide. The C-terminus of ComS contains the 7-residue small hydrophobic peptide XIP. Extracellular XIP is imported by the Opp permease and interacts with ComR to form a transcriptional activator for both *comX* and *comS* [19, 20]. Notably, the *S. mutans* competence pathway contains at least two elements of positive feedback, as XIP/ComR activates *comS* and ComX activates *comE* expression [19, 20].

An intriguing property of *S. mutans* competence is that although exogenous CSP and XIP can both activate *comX* and induce transformability, they do so under different environmental conditions and they elicit qualitatively different behaviors in the expression of *comX* [11]. Exogenous CSP elicits a bimodal response in which less than half of the population activates *comX;* exogenous XIP elicits a unimodal response in which all cells in the population activate *comX*. Further, CSP activates *comX* only in complex media containing small peptides, and only in cells carrying an intact *comS*. Activation by CSP does not require the permease gene *opp* [11]. By contrast, exogenous XIP activates *comX* only in defined media lacking small peptides and only in

cells with an intact *opp* gene [11, 21]. Activation by XIP does not require *comS* [20]. Therefore, although exogenous XIP can compensate for a *comS* deletion and activate *comX*, the bimodal *comX* response still requires an intact *comS* gene.

The observation that competence in several streptococcal species is directly stimulated by an extracellular ComS-derived peptide suggests that ComRS constitutes a novel type of Gram positive quorum signaling system, in which the ComS-derived signal XIP is processed and secreted, accumulates in the extracellular medium, and is then reimported. This interpretation in *S. mutans* is supported by several experimental observations. First, cells that carry *opp* take up exogenous XIP (in defined medium) and activate *comX* with high efficiency [20, 21]. Second, exogenous synthetic XIP is dramatically more effective in stimulating transformability than is exogenous full-length ComS [20]. Third, filtrates of *S. mutans* cultures grown to OD_{550} = 0.4 in defined medium were able to stimulate a P*comX* reporter strain [21]. Similarly, LC-MS/MS analysis found micromolar concentrations of XIP in supernates of *S. mutans* cultures grown to high density [22, 23]. Fourth, a transposon mutagenesis screen in *S. pyogenes* identified the widely conserved *pptAB* ABC transporter as a possible exporter of short hydrophobic peptides of the ComS type [24], raising the possibility that *S. mutans* may also possess a dedicated export mechanism for ComS or XIP.

However, although dedicated mechanisms that process and export the CSP signal in *S. mutans* are well characterized, corresponding mechanisms for ComS/XIP have not been identified in *S. mutans*. Although there is evidence that the Eep membrane protease facilitates the processing of *S. thermophilus* ComS [25], Eep did not affect processing of *S. mutans* ComS [22]. Further, although *S. mutans* encodes a

gene product with a fairly high degree of homology to *S. pyogenes* PptAB, deletion of the apparent *pptAB* genes in *S. mutans* had only a weak effect on competence induction in mid-exponential phase cultures [24]. Consequently it remains unknown how ComS is processed to XIP and exported, although cell lysis appears to play a larger role than active transport in the export of XIP. Coculture studies [26] showed that deletion of the gene *atlA* for a major autolysin impeded the ability of *comS*-overexpressing cells to induce *comX* in cocultured cells that lacked *comS*.

The import of XIP presents an additional puzzle for ComRS quorum signaling, as the permease OppA is not required for exogenous CSP to activate *comX*, but is required for XIP to activate *comX*. It has been suggested that bacteriocin production induced by CSP may create another entry route for extracellular XIP by increasing membrane permeability [12]. However, such a model implies, contrary to data [11], that CSP should also induce *comX* in defined growth media. In addition this permeability model does not explain the characteristic bimodal response of *comX* to extracellular CSP, as the CSP/ComDE pathway induces bacteriocin genes such as *cipB* unimodally (population-wide) [20].

The lack of an established mechanism for processing and export of ComS, and the experimental link between *comX* bimodality and the endogenous production (via *comS)* – but not the import (via Opp) – of XIP raise the question of whether the ComRS system can activate *comX* through purely intracellular signaling, at least under some environmental conditions. Intracellular transcriptional feedback is often a cause of bimodality in bacterial gene expression [27]. An ability to activate *comX* through endogenously produced, intracellular ComS (or XIP) would potentially allow individual

S. mutans within a population to exhibit different competence behavior, without requiring accumulation or import of extracellular XIP [14]. Here we used a combination of microfluidic and continuous flow experiments, including coculture studies using *comS* deletion and *comS* overexpressing strains, to test whether activation of *comX* is necessarily accompanied by import or export of XIP and to assess the contribution of endogenous ComS production to *comX* activation in individual cells.

Results

An intact copy of *comS* alters the *comX* response to exogenous XIP

Fig. 1A compares the *comX* response to exogenous XIP in *S. mutans* UA159 (wild-type) and a *comS* deletion (Δ*comS*) genetic backgrounds, as measured by the fluorescence of a P*comX-gfp* plasmid-borne reporter. Cells were imaged while adhered within a microfluidic chamber and supplied with a constant flow of defined medium (FMC) containing synthetic XIP. Although both strains respond to exogenous XIP, the Δ*comS* strain consistently showed roughly 1.5-fold lower P*comX* activity than the wild type, at all XIP concentrations. Even saturating concentrations of exogenous XIP did not induce *comX* to equivalent levels in the wild-type and Δ*comS* strains. This result is similar to the roughly two-fold difference in XIP-induced transformability observed for the wild-type and *ΔcomS* strains in a previous study [20]. Fig. 1B shows that the threshold for *comX* response occurred at a roughly 2-fold lower XIP concentration in UA159 than in the Δ*comS* strain. Therefore the deletion of *comS* elevated the threshold for a response to extracellular XIP and reduced the overall response at saturation.

The deletion of *comS* also affected cell-to-cell variability (noise) in *comX* expression. Figs. 1C and 1D show that the histograms of reporter fluorescence differ in wild-type and Δ*comS* cells. Wild type showed a generally broader (noisier) *comX* response than did Δ*comS*. We quantified this difference by fitting the histograms to a gamma distribution $\Gamma(n \mid a,b)$, a two-parameter probability distribution that can be used to model cell-to-cell variability in n, the copy number for a bacterial protein [28]. In a simple physical model, the parameter a of the gamma distribution is related to the number of mRNAs produced during the cell division time, while b is related to the number of protein copies produced per mRNA transcript [29]. As shown in Fig. 1E and 1F, the UA159 background has a roughly 2-fold higher value for parameter a (transcription rate), while parameter b (translation) is similar for the two strains. As this difference persists even at XIP concentrations exceeding 1 μM, these data emphasize the significant effect of *comS* deletion on *comX* expression, even when excess extracellular XIP is provided at concentrations that saturate the *comX* response.

Fluid replacement did not alter induction of *comX* in a *comS* overexpressing strain

To test whether *comX* activation in complex growth media requires the accumulation of extracellular XIP, we tested the effect of fluid flow rate on *comS* overexpressing cells that were adhered in a microfluidic flow chamber. Cells carrying the 184*comS* overexpression plasmid had previously been observed to activate *comX* in defined medium lacking exogenous CSP or XIP [26]. We anticipated that in an experiment where the cells were immobilized and supplied with a continuous flow of

fresh medium (lacking XIP or CSP), high flow rates would remove extracellular XIP (or ComS), leading to diminished *comX* activity. We loaded 184*comS* P*comX-rfp* cells, which carry both the *comS* overexpression plasmid and a plasmid-borne P*comX-rfp* reporter, into four different microfluidic flow chambers. Chambers were supplied with fresh complex medium flowing at rates between 0.02 ml h^{-1} and 1 ml h^{-1}. These flow rates were sufficient to completely replace the growth medium within each chamber on time intervals ranging from 6 seconds to 10 minutes. We also studied (1) a P*comX-rfp* reporter in a UA159 background (negative control), and (2) a 184*comS* overproducing strain that was lacking a start codon (ATG point-mutated to AAG) on its chromosomal *comS* (184*comS* P*comX-rfp* Δ*comS*).

Figs. 2A-2B show that P*comX* was not activated in the negative control (wild-type background) under any condition. However, the strain that overexpressed *comS* and harbored an intact chromosomal copy of *comS* showed a highly heterogeneous response, indicating that a subpopulation of these cells strongly activated *comX* in the flowing complex medium. Most significantly, the rate of fluid flow had little effect on either the median or the variance in *comX* expression. Fig. 2B shows the fluorescence of the particular cells in which the signal exceeded the maximum P*comX* activity (roughly 14 fluorescence units) seen in the UA159 negative control. The median *comX* activity in the *comS*-overexpressing cells did not decline at the highest flow rates; instead it showed a weak increase, which was smaller than the cell-to-cell variability. These data show that overexpression of *comS* can allow *S. mutans* to activate *comX*, even in complex medium containing small peptides; in such media *comX* transcription is not generally inducible by exogenously supplied XIP. The finding that this activation is

unaffected by very rapid replacement of the medium implies that the *comX* response in the *comS* overexpressing strain did not require extracellular XIP (or ComS) to accumulate in the medium. However this *comX* response did require a chromosomal copy of *comS*: Fig. 2B shows that very few Δ*comS* cells activated *comX*, even though they harbored the *comS* overexpression plasmid. Together with Fig. 1, these data show that the chromosomal *comS* plays a role in *comX* activation that is not fully complemented either by saturating concentrations of exogenous XIP or by endogenous overproduction of ComS.

The presence of *comS* in the chromosome increases *comS* mRNA levels

Figure 1 suggests that transcription from the native *comS* plays an important role in the response of the wild type to exogenous XIP, while Figure 2 suggests that the native *comS* is essential to the self-activation of the ComS-overexpression strain. Therefore we used RT-qPCR to measure *comS* and *comX* transcript copy numbers in mid-exponential phase cultures of these strains. Strains were grown in complex medium (BHI) or defined medium (FMC) ± synthetic XIP, with 16S rRNA serving as a control for normalization. The amount of *comS* mRNA was significantly greater in cells of wild type treated with exogenous XIP, in comparison with controls lacking XIP or grown in complex medium (Fig. 3A). In addition, cells that overexpressed *comS* in a wild type background (184*comS*) and were grown in complex medium had *comS* transcript levels that were significantly higher than in a strain that overexpressed *comS* but lacked a chromosomal copy of *comS* (184*comS* Δ*comS*). As expected, the Δ*comS* strain in defined medium showed only a baseline signal for transcription, whether or not XIP was

provided. The *comX* transcript levels were low compared to *comS*, and spanned a smaller range overall, with a statistically significant enhancement only in the wild type grown with XIP (Fig. 3B). Overall these data verify that addition of exogenous XIP (in defined medium) enhances *comS* transcription, but they also show that the presence of the chromosomal copy of *comS* significantly boosts *comS* mRNA levels when *comS* is overexpressed from a plasmid.

Population density of *comS* overexpressing cells does not determine the *comX* response

To test whether *comS*-overexpressing cells release extracellular XIP that can activate *comX* in Δ*comS* cells, we measured *comX* activation in cocultures of 184*comS* P*comX-rfp* (senders) and P*comX-gfp* Δ*comS* (receivers). We prepared cocultures by mixing sender (*comS* overexpressing) and receiver (*comS* deficient) cultures in different ratios. We loaded the cocultures into microfluidic chambers containing static, defined medium (FMC) without exogenous XIP. We anticipated that, if senders released XIP or ComS into the extracellular medium, both senders (RFP reporter) and receivers (GFP reporter) would respond by activating *comX*, and that the average activation would increase with the ratio of senders to receivers.

We analyzed the green and red fluorescence of the cocultures to generate histograms of individual cell fluorescence that reveal both the sender (red) and receiver (green) *comX* response, shown in Figs. 4A and 4B. Representative microscopy images are shown in Figs. 4C-4H. As expected, a control chamber containing only receiver cells in 50 nM XIP showed enhanced GFP fluorescence but only baseline RFP fluorescence

(second column in Fig. 4A, 4B, and Fig. 4D). Similarly, a control chamber containing only sender cells showed enhanced RFP fluorescence but only baseline GFP fluorescence (rightmost column in Figs. 4A, 4B, and Fig. 4H). When exogenous XIP was not provided, the GFP fluorescence of cocultures was unchanged as the sender/receiver ratio increased. The GFP fluorescence histograms remained at the baseline level of the negative control (Δ*comS* strain alone) in the first column of Fig. 4A. Further, the GFP response of the cocultures did not change appreciably over a period of four hours (Supporting Fig. S1). Even when present in abundance, senders did not activate *comX* in the Δ*comS* receivers.

By contrast, the median RFP fluorescence of the cocultures (Fig. 4B, 4F-4H) did increase at high sender/receiver ratios. The median RFP fluorescence of cells in the coculture increased approximately in proportion to the density of senders, as expected if each sender activated its own *comX*. RFP fluorescence was constant over a period of four hours (Supporting Fig. S1). These data show that although overexpression of *comS* stimulates *comX* within individual cells, this activation does not cause extracellular, diffusible XIP to accumulate to levels that activate nearby Δ*comS* cells.

Growth phase-dependent release of XIP

We previously showed that intercellular signaling by *S. mutans* ComRS is impeded by deletion of the *atlA* gene, encoding a major autolysin. Loss of AtlA inhibits cell lysis, which appears to occur primarily in stationary phase [26]. We therefore tested whether signaling from a sender (*comS* overexpressing) strain to a receiver (Δ*comS*) strain is enhanced in the latter phases of growth. We prepared sender/receiver

cocultures in different ratios in defined medium. Every 2 h the pH of the cultures was adjusted to 7.0 by addition of 2 N NaOH, the OD_{600} was recorded, and an aliquot of the culture was collected for fluorescence imaging of *comX* promoter activity. Low pH suppresses the *comX* response, so the pH adjustment ensured that cells were able to respond to *comX*-activating signals when present [9, 13]. The GFP fluorescence histograms of Fig. 5A show that *comX* expression in the Δ*comS* strain is slightly higher at 12 h than at 2 h or 8 h. This increase was more pronounced at higher ratios of sender to receiver cells, consistent with increased release of XIP from the senders later in growth phase. The histograms of Fig. 5B show a strong RFP response (due to senders) as in Fig. 4B, with a slightly stronger response at earlier times.

The median GFP and RFP signals in the above histograms do not shift dramatically with either time or coculture ratio. However the histograms in Fig. 5A suggest moderate, density-dependent increases in receiver (green) fluorescence at 12 h. Figs. 5C and 5D highlight these changes by showing the value of the 99th percentile of red and green fluorescence respectively in the cultures, versus optical density. In Fig. 5C the response of the most active receivers increases slightly at higher OD_{600} values, a change that is slightly more pronounced at higher sender:receiver ratios. By contrast Fig. 5D shows no strong trend in the fluorescence of the 99th percentile of senders, versus OD_{600}. None of the cocultures exhibited as strong a GFP response as the positive control (receiver + 50 nM exogenous XIP, Fig. 5C), indicating that even after 12 h extracellular XIP had not accumulated to the level used for the positive control. Overall these data are consistent with a robust autoactivation of the

comS overexpressing senders, accompanied by some release of XIP (or ComS) to the extracellular medium in late growth.

In vitro binding of ComR to the *comS* and *comX* promoters

The observation that the overexpressing (sender) strain activated *comX* without releasing XIP into the medium implies that activation of *comX* does not require export and reimport of XIP or ComS, if the chromosomal copy of *comS* is intact. To test whether ComS could serve as an intracellular signal to activate *comX*, we tested whether unprocessed ComS could enable ComR binding to the *comS* and *comX* promoter regions *in vitro*. We used a fluorescence polarization assay, with purified recombinant ComR, synthetic XIP or ComS, and a fluorescently labeled DNA oligomer corresponding to the *S. mutans comX* promoter region containing the ComR binding site. As histidine tagging of ComR was found to reduce its affinity for the *comX* promoter region (Supporting Fig. S2), untagged ComR was used for these assays. Fig. 6A shows the fluorescence polarization as a function of ComR concentration, both in the presence and absence of excess (10 μM) ComS or XIP. Because of the excess of peptide (10 μM) relative to fluorescent probe (1 nM), the probe polarization depends primarily on the concentration of ComR added. In the absence of XIP or ComS, ComR caused a weak rise in the fluorescence polarization of the DNA oligomer, indicating a weak binding affinity like that of ComR of other streptococci [19]. However, in the presence of ComS or XIP the binding isotherm saturated at low ComR concentrations, indicating formation of a complex with high affinity for the *comX* promoter.

Fig. 6B shows a competition assay in which unlabeled ('cold') P*comX* and P*comS* DNA oligomers having the same stem-loop structure as the labeled probe were titrated into samples containing 1.5 μM ComR, 10 μM ComS or XIP, and the labeled DNA (1 nM). The systematic decrease in polarization is consistent with competition for ComR. The unlabeled P*comS* and P*comX* probes, which differ by three bases, appear to have identical affinity for ComR.

Fig. 6A shows that ComS interacts with ComR and the DNA probe, although its complex with ComR has weaker affinity for the probe than does the XIP/ComR complex. It also shows that at saturating concentrations of ComR, inclusion of ComS produces approximately half of the total fluorescence polarization that is elicited by an equivalent concentration of XIP. As the fluorescent probe is present at very low concentration, this difference may suggest that ComS and XIP induce a qualitatively different interaction between the ComR/peptide complex and the DNA probe. The solid curves in Fig. 6A and 6B are calculated from a two-step, cooperative binding model (*Methods*) in which the dissociation constants for the ComR/peptide complex (k_1) and the complex/promoter (k_2), as well as the order of multimerization n of the ComR/peptide complex, are variables. As discussed in *Methods*, the data are consistent with a range of values for these parameters, but generally we obtain micromolar values for k_1 and nanomolar k_2 for both ComS and XIP, and similar cooperativity n for both peptides. The curves in the figure show the model with a roughly two fold difference in the ComR dissociation constants for ComS (k_1 = 3.2 μM, n = 2.5) and XIP (k_1 = 7.3 μM, n = 2.4). Although the data clearly indicate that both ComS and XIP interact with ComR

to bind the DNA probe, they do not permit a precise determination of the ComS and XIP interaction parameters.

Data are consistent with an intracellular feedback loop in *comS* transcription

The different behavior of the wild-type and *comS* deletion strains in Fig. 1A and the effect of *comS* deletion in Fig. 2A show that transcription from an intact *comS* under native control has an effect on *comX* expression that cannot be duplicated either by exogenous XIP or by a ComS overexpression plasmid. One such effect could be transcriptional positive feedback in which the chromosomal copy of *comS* produces ComS that is retained in the cell and stimulates a high level of *comS* transcription, more than is achieved with a ComS overexpression plasmid alone. Import of extracellular XIP would likely stimulate such a feedback mechanism. The resulting enhanced intracellular levels of ComS, if they activate *comX* efficiently, could account for the important role of natural control of *comS* and the lack of evidence for XIP release by *comX*-active cells. The feedback mechanism, including the interplay between imported XIP and the chromosomal *comS*, is illustrated in Fig. 7A.

A mathematical model of this mechanism is tested in Fig. 7B. The model, which is further described in the Supporting Information, assumes that (1) ComR can form a complex with either ComS or XIP; (2) extracellular XIP is able to enter the cell (in defined growth medium); (3) neither ComS nor XIP is exported, although endogenously produced ComS may be converted to XIP. Motivated by the differences in the saturated fluorescence polarization with ComS and XIP, the model further assumes that the order of multimerization in the ComR/peptide complex that binds DNA is n = 1 (ComS) or n =

2 (XIP) (see *Discussion*). Finally, the maximal rate of transcription (ComX production) was allowed to depend on whether a ComR/XIP or a ComR/ComS complex was bound to the promoter.

The steady states of the dynamical system are found from the nullclines of the differential equations of the model (*Methods*). We found that if the ComR/XIP and ComR/ComS complexes are permitted to give different maximal transcription rates, the model reproduces the different *comX* expression in the Δ*comS* and wild-type backgrounds in Fig. 1A. Further, if only XIP (but not full-length ComS) interacts with ComR to activate *comX*, the model fails to reproduce the Fig. 1A data, giving instead identical *comX* activation in both the Δ*comS* and wild-type backgrounds. Parameter values for the best fit and 90th and 10th percentiles of a bootstrap uncertainty analysis are reported in Supporting Information Table S2. Parameters were found to preserve the relative orders of magnitude of the dissociation constants of the transcriptional activators found in Fig. 6, with the best fit giving a maximal transcription rate that is between 4-fold and 200-fold greater when the ComS-ComR complex is bound than when the XIP-ComR complex is bound.

Discussion

The ComRS system found in mutans, salivarius, pyogenic and bovis streptococci has been described as either a quorum sensing system [20] or a timing mechanism [25] for direct control of *comX*, the master regulator of genetic competence. The ComS-derived peptide XIP is readily imported by *S. mutans* in defined growth medium, where it induces transformability with high efficiency. Some of the key evidence supporting an

intercellular signaling role for XIP include the detection by LC-MS/MS spectroscopy of XIP in supernates of *S. mutans* that were grown to high cell densities [22, 23]. In addition, filtrates of *S. mutans* cultures grown to high density induced P*comX* activity in reporter strains [21], indicating the presence of an active competence signal in the extracellular medium. A recent coculture study verified that XIP is freely diffusible in aqueous media and showed that ComS-overexpressing senders are able to activate *comX* in nearby Δ*comS* receiver mutants, with no cell-cell contact being required [26]. However, this intercellular signaling was suppressed [26] by deletion of *atlA,* which encodes a surface-localized protein associated with envelope biogenesis and autolysis [30]. Taken together these data indicate that the *S. mutans* ComRS system provides intercellular competence signaling when autolysis releases sufficient concentrations of ComS or XIP.

Our data provide several lines of evidence that ComRS can also control *comX* without the accumulation and import of extracellular XIP or ComS. First, the cell's own native regulation of *comS* affects its activation of *comX*, independent of whether it overproduces ComS from a plasmid or imports exogenous XIP via OppA. Although the response of *comX* is different in complex medium supplemented with CSP than in defined medium supplemented with XIP, in both cases the behavior of *comX* is affected by the presence of an intact chromosomal copy of *comS* under the control of its cognate promoter. In complex medium, *comS* is required in order for CSP to elicit any *comX* response; in defined medium, the deletion of *comS* reduces the *comX* response (both average and variance) to exogenous XIP and raises the threshold XIP concentration for a response.

In addition, even if it harbors a *comS* overexpression plasmid, a *comS* deletion strain expresses *comX* much more weakly than does a *comS*-overexpressing strain that retains its chromosomal *comS*. The strain with a chromosomal *comS* also produces substantially more *comS* mRNA transcripts. Both strains activate *comX* in response to extracellular XIP, but only the strain harboring chromosomal *comS* activates *comX* in the absence of extracellular XIP. The striking effect of chromosomal *comS* in the overexpressing strain could possibly result from an interaction between the homologous *comS* sequences on the plasmid and chromosome, in which the chromosomal *comS* is placed transiently under control of the plasmid P23 promoter. However this would have an effect similar to increasing the plasmid copy number by one, and so is unlikely to explain the different behaviors of the strains. Alternatively the native *comS* promoter may simply lead to substantially higher *comS* mRNA levels when activated, leading to high intracellular ComS levels that facilitate further *comS* activation. We note that inactivation of chromosomal *comS* consistently altered the response to XIP/ComS, whether inactivation was accomplished by a deletion-replacement of *comS* or by a change in the *comS* start codon.

Our data also show that even in complex growth medium, which is known to inhibit the uptake of extracellular XIP, ComS-overexpressing (sender) cells activate their own *comX.* This autoactivation is unaffected by very rapid exchange or flow of the medium, strongly suggesting that *comX* activation in these cells does not require accumulation and import of XIP. These behaviors are consistent with an intracellular positive feedback loop in which *comS* stimulates its own expression through its cognate

promoter, enabling the cell to autoactivate *comX* or lowering its threshold of sensitivity to extracellular XIP.

Finally the data show that *comS*-overexpressing cells fail to stimulate Δ*comS* cells (receivers) in cocultures in defined medium, where XIP should be efficiently imported. As the Δ*comS* receivers do respond to exogenously added XIP, these data indicate that the overexpressing cells activate their own *comX* without releasing significant XIP to the medium. The weak intercellular signaling that is observed in cocultures grown to late growth phases is consistent with eventual lysis of sender cells, possibly linked to autoactivation of the lytic pathway driven by ComX and ComDE.

The finding that diffusive signaling by ComS or XIP between *S. mutans* cells is inefficient or lacks spatial range is consistent with the conclusion reached by Gardan *et al*. using *S. thermophilus.* Those authors found that the type I ComS peptide of *S. thermophilus* was not secreted at detectable levels in a strain that produced it naturally, although an overproducing strain did generate detectable ComS in the medium [25]. They argued that ComS does not diffuse through or accumulate in the medium, although it may be able to signal between cells that are in physical contact. This proximity model for ComS resembles a “self-sensing” quorum system [31] in which the secreted signal is retained at elevated concentrations in the immediate surroundings of the cell, possibly associated with the cell surface, so that the cell responds somewhat more strongly to its own secreted signal than to that of the rest of the population.

Our observations suggest more strongly that export or import of ComS or XIP is not essential to ComRS control of *comX* in *S. mutans,* under the conditions examined. The key role of native control of *comS* in our data argues that the more essential

component is the dynamics of the cell's own *comS* transcription. If cells do not export endogenously produced ComS or XIP in the absence of lysis, then ComS or XIP would be available within the cells to activate P*comS* and drive positive feedback in *comS*, leading to strong P*comX* activation. We have previously argued that the bimodal response of *S. mutans* to CSP stimulation, which requires *comS* but not *opp*, suggests that CSP stimulates a noisy, intracellular autofeedback loop of this type. If CSP/ComDE can, through its as-yet-unknown pathway, facilitate *comS* transcriptional feedback, then *comX* expression may occur in at least some cells, leading to the observed bimodal distribution of *comX* activity in a population [11]. Notably the overexpression of *comS* in our study also leads to heterogeneous *comX* activity, suggesting that it plays a role similar to exogenous CSP by facilitating *comS* autofeedback. Transcriptional feedback loops generate heterogeneous, bimodal behaviors at the single-cell level in many other bacterial systems, including competence regulatory pathways. In the regulation of *Bacillus subtilis* competence, an intracellular feedback loop based on ComK activates a subpopulation of cells into the competent state [32].

In the case of *S. mutans*, where it is unknown whether ComS is processed to XIP inside the cell, either XIP or ComS could potentially act as the intracellular feedback signal. Although *S. mutans* competence was shown to be unresponsive to exogenous full-length ComS [20], this finding may reflect either selectivity by ComR or simply inefficient import of full length ComS by OppA. ComS is significantly larger (17 residues) than peptides that are typically transported by ABC transporters. Shanker *et. al.* found that *S. mutans* ComR is unresponsive to the ComS peptides produced by other streptococcal species [33], although an eight residue XIP ($ComS_{10-17}$) did interact

effectively with ComR to bind the *comS* and *comX* promoters [19]. Our fluorescence polarization data confirm that both ComS and XIP can interact with ComR to bind the *comX* and *comS* promoter regions. They also suggest that ComS and XIP may form ComR complexes of different degrees of multimerization, a difference that could have consequences for the nonlinear dynamics of feedback regulation. Our mathematical model for transcriptional autofeedback in the *comRS* system incorporates the data by assuming that endogenously produced ComS is not released to the environment, although extracellular XIP is imported and supplements the endogenous ComS in interacting with ComR. Therefore, the ComS and XIP complexes of ComR may together drive expression of both *comS* and *comX.*

Further work will be needed to verify whether intracellular ComS and XIP can both activate *comX* by modulating ComR binding. Structural studies in *S. pyogenes* have shown that some intracellular Rgg receptor proteins can bind pheromones that differ in length and sequence [34]. Crystallographic structures of homologous ComR proteins [33, 35] show the SHP binding pocket of the ComR C-terminus to fall in the tetratricopeptide repeat domain that is responsible for multimerization, while the N-terminus helix-turn-helix structure binds DNA after an induced structural rearrangement. The location of the SHP binding pocket could allow the longer ComS to hinder multimerization when bound, resulting in a monomer binding to its target, while XIP does not. Supporting Fig. S2 shows preliminary evidence that the ComS N-terminus affects ComR binding in *S. mutans*. Neither ComS nor XIP could induce DNA binding by an N-terminally 6x histidine tagged ComR, whereas XIP (but not ComS) caused DNA

binding activity in a C-terminally tagged ComR. These data indicate that steric effects around the SHP binding pocket may influence DNA binding affinities.

Positive feedback occurs in many quorum sensing systems as the accumulation of the chemical signal in the extracellular environment stimulates the cell to produce additional signal or its cognate receptor. For example, in *Vibrio fischeri* the C8 homoserine lactone autoinducer stimulates expression of *ainS*, which encodes the autoinducer synthase [36]. In *Vibrio cholerae* the CAI-1 signal stimulates production of its receptor CqsS [37]. In these cases the extracellular signal concentration, which is the positive feedback signal, is sensed by large numbers of cells, and so the population responds homogenously. However, if an individual cell responds preferentially to its own signal production then the feedback signal is specific to the individual cell and the behavior is qualitatively different. Individual feedback can convert a graded (or unimodal) population response to a switched or bimodal response [38]. Depending on parameters such as the rate of signal production, noise levels or the cell density, the response of the cells may then span a range from strongly social or quorum behavior to purely autocrine or self-sensing [31] behavior in which cells respond independently and the population becomes heterogeneous [39]. Synthetic biology has exploited this phenomenon in several bacterial quorum sensing systems to amplify the cell's sensitivity to an exogenous signal. This can lower the quorum circuit's threshold sensitivity to the signal, and it can also enhance the amplitude of the cell's full response to that signal. Fig. 1A and 1B suggest that the chromosomal *comS* in *S. mutans* roughly doubles the amplitude of *comX* response and lowers the XIP sensitivity threshold

roughly twofold. This amplification is comparable to what was accomplished in engineered, synthetic systems [40, 41].

As a result the ComRS system may have two modes of function in *S. mutans*. At low population densities, during early growth, ComRS operates through intracellular feedback, leading to population bimodality in *comX* expression. Here only a small subpopulation of cells activate the late competence genes. However, in later growth phases or in mature biofilms, stress mechanisms that drive autolysis allow the release of XIP, providing a diffusible signal that is detected by other cells and amplified through the internal feedback mechanism to elicit a strong competence response. In this sense XIP may serve to broadcast localized stress conditions, stimulating *S. mutans* to scavenge DNA resources opportunistically from nearby lysing cells [42, 43].

Materials and methods

Strains and growth conditions

S. mutans wild type strain UA159 and mutant reporting/gene deletion strains from glycerol freezer stock were grown in BBL BHI (Becton, Dickinson and co.) at 37˚C in 5% CO_2 overnight. Antibiotics were used at the following concentrations where resistance is indicated in Table 1: erythromycin (10 µg ml^{-1}), kanamycin (1 mg ml^{-1}), spectinomycin (1 mg ml^{-1}). For experiments in defined medium, strains were washed twice by centrifugation, removal of supernatant and re-suspension in the defined medium FMC [44]. These were then diluted 20-fold into fresh FMC and allowed to grow in the same incubator conditions until at optical density at 600 nm (OD_{600}) of 0.1.

Synthetic XIP (sequence GLDWWSL) was synthesized and purified to 98% purity by NeoBioSci (Cambridge, MA).

E. coli strains were grown in LB at 37 ˚C shaking in an aerobic incubator overnight. Antibiotics were used at the following concentrations where resistance is indicated: ampicillin (10 µg ml^{-1}). The next day the overnight cultures were diluted 100-fold into LB containing ampicillin at the indicated concentration and grown under the same incubator conditions as overnight.

Mutant strains used

Table 1: list of strains and plasmids used.

Strain or plasmid	Characteristics*	Source or reference
***S. mutans* strains**		
P*comX* GFP	UA159 harboring P*comX* GFP promoter fusion.	[11]
P*comX* RFP	UA159 harboring P*comX* dsRed RFP promoter fusion.	[26]
P*comX* GFP Δ*comS*	UA159 *comS* gene replaced with a non-polar erythromycin resistance cassette. Harboring P*comX* GFP promoter fusion.	[11]

184*comS* P*comX* RFP	UA159 harboring pIB184*comS* and P*comX* dsRed RFP promoter fusion.	[26]
184*comS* P*comX* RFP Δ*comS*	UA159 harboring pIB184*comS* and P*comX* dsRed RFP promoter fusion. *comS* disrupted by point mutation in start codon (ATG to AAG).	This study.
***E. coli* strains**		
BL21(DE3)	Used for recombinant protein expression	New England Biolabs, MA
10-beta	Used for propagating plasmids during cloning	New England Biolabs, MA
Plasmids		
pIB184	Shuttle expression plasmid with the P23 constitutive promoter, Em^R	[45]
pDL278	*E. coli – Streptococcus* shuttle vector, Sp^R	[46]

pET45b(+)his-comR$_{UA159}$	pET45b(+) derivative containing the translational fusion P$_{T7lac}$-*6xhis-comR*$_{UA159}$, Ap^r	This study

*Em = erythromycin, Sp = spectinomycin, Ap = ampicillin.

Construction of mutant strains

<u>Δ*comS* point mutant</u>

The start codon of the *comS* gene was mutated from ATG to AAG. The mutation was introduced directly into the chromosome by site-directed mutagenesis using a PCR product generated by overlap extension PCR [47]. Potential mutants were screened using mismatch amplification mutation analysis (MAMA) PCR [48], as previously described [49, 50]. The point mutation was confirmed by PCR and sequencing to ensure that no further mutations were introduced into the *comS* gene and its flanking regions.

<u>His-ComR expressing *E. coli*</u>

The *comR* gene (SMu.61) was amplified using gene-specific primers (forward, AAAGAATCCTATGTTAAAAGA; reverse, CACCCTAGGAGACCCATCAAA) and cloned into the pET45b(+) vector bearing an N-terminal 6xHis tag. The resulting vector, pET45b(+)his-comR$_{UA159}$, was transformed into *E. coli* 10-beta. After sequencing confirmed the correct insertion (using T7 promoter and T7 terminator primers), the vector was transformed into *E. coli* BL21(DE3) prior to protein purification.

Microfluidic experiments

Microfluidic experiments were performed using a seven-channel PDMS-cast mixing array device as described previously [11, 51]. Cells grown to OD_{600} 0.1 from dilution in FMC were sonicated briefly using a Fisher Scientific FB120 sonic dismembrator probe to split large chains. Sonicated cells were then loaded into the device through a syringe capped with a 5 µm filter to remove any remaining aggregations. FMC containing 1 mg ml^{-1} spectinomycin and a XIP gradient produced from three inlets containing different concentrations of XIP (0 nM, 600 nM and 6 µM XIP inlets) passed through a mixing matrix was pumped through the cell chambers at a steady rate of 0.08 ml/hr to create a constant, different XIP concentration in each cell chamber. Flow stability and XIP concentration were inferred using a fluorescent red dye (sulforhodamine 101, Acros Organics) mixed into the inlet medium in proportion to XIP concentration (i.e. none in 0nM XIP inlet, and 10x the concentration of sulforhodamine in the high XIP inlet as in the middle inlet). This allowed calculation of XIP concentration by measurement of a given channel's red fluorescence relative to the red fluorescence of the channels with known concentration.

Phase contrast and green fluorescence pictures of reporting cells were collected using a Nikon TE2000U inverted microscope (equipped with a 40x objective, CFI Plan Fluor DLL, NA 0.5, Nikon) and a CCD camera (CoolSNAP HQ2, Photometrics) with a green fluorescence filter (Nikon C-FL GFP HC HISN zero shift filter cube) excited by a mercury lamp source (Intensilight C-HGFI, Nikon). Images were taken at 0, 30 and 90 minutes after cell exposure to XIP commenced. Images were analyzed according to the method described previously [52] with fitting performed in Matlab® (The Mathworks

inc.). A red fluorescence filter (C-FL Y-2E/C dsRed Filter Cube, Nikon) was used to visualize sulforhodamine concentration in the channels.

Gamma distributions (a two-parameter probability distribution describing the amount of protein produced in sequential transcription and translation steps) were fit to the single cell fluorescence distributions using Matlab to fit protein production to theoretical description [29]. The fit was applied to cells fluorescing above an arbitrary cutoff of 40 units (around background) in order to prevent turned-off cells from skewing the distribution. Parameters were rounded to three significant figures and reported in Table S1 (see supporting information).

Flow rate dependence experiment

In order to measure the flow rate dependence of XIP signaling we loaded cells into a commercial six-channel microfluidic slide (IBIDI µ-slide VI, IBIDI GmbH). The six channels contained respectively (1) a red fluorescent protein (dsRed) *comX* reporting strain (P*comX* RFP) control channel flowing fresh BHI at 0.1 ml h^{-1}; (2)-(5) four channels containing a *comS* overexpression strain 184*comS* P*comX* RFP (*comS* on plasmid pIB184 under the strong constitutive P23 promoter) with BHI at different flow rates ranging from 0.02 ml h^{-1} up to 1ml h^{-1}, and (6) a 184*comS* P*comX* RFP strain with a point mutation disrupting the chromosomal *comS* gene (Δ*comS*) under flow at 0.1 ml h^{-1}. After 2 hours the plain BHI supplied was replaced with BHI supplemented with 50 µg ml^{-1} chloramphenicol in order to halt further translation and allow any RFP in the cells to fold. This was supplied at 0.1 ml h^{-1} flow rate for all channels. Four hours (the maturation time of our RFP) after chloramphenicol addition final fluorescence images of the cultures were taken. Due to the bimodal *comX* activation in BHI, a fluorescence

cutoff was set as the maximum RFP fluorescence observed in the PcomX RFP negative control. Cells exhibiting RFP fluorescence above this level were collected in an array and the size of this sample as a percentage of the population as well as the median of the above-cutoff fluorescence reported.

Channel coculture experiment

We loaded cocultures of a P*comX* GFP Δ*comS* (responders) with the *comS* overexpressing strain 184*comS* P*comX* RFP (senders) into two commercial microfluidic slides (IBIDI µ-slide VI) using static (not flowing) FMC medium and varying ratios of *comS* overproducers:Δ*comS* responders (percentage by volume of OD_{600} 0.1 cultures vortexed together). P*comX* RFP, P*comX* GFP Δ*comS,* P*comX* RFP + 50 nM XIP and P*comX* GFP Δ*comS* + 50 nM XIP were used as controls. The end ports of the channels were sealed with mineral oil to prevent drying of the medium in the channels. Images were taken as in the microfluidic experiments and analysis performed similarly. In the case of controls XIP was added to planktonic culture and the tube vortexed before pipetting into the slide. Because the population was heterogeneous in both fluorescent reporter type and *comX* expression, a fluorescence threshold was defined as the maximum RFP fluorescence observed in the PcomX RFP negative control as previously. The median of the RFP fluorescence observed above this cutoff in other samples was used as a measure of how strongly the red cells were activating *comX* as a function of their number density.

OD dependence of coculture response

For tests of growth-phase dependence of signaling, cocultures similar to those in microfluidic channel slides were prepared. Overnight cultures were washed and diluted

40x into fresh FMC containing erythromycin (10 µg ml^{-1}) and spectinomycin (1 mg ml^{-1}). Once grown to OD_{600} 0.05, these were mixed in ratios varying from to 0% *comS* overexpressers to 100% overexpressers, defined by volume of *comS* overproducers added divided by the volume of the Δ*comS* culture added. Low initial cell densities were used to ensure that early, mid and late growth phases were probed for XIP release. Every two hours the OD_{600} of the culture and its pH were measured. The pH was corrected back to 7.0 using 2N sodium hydroxide if it had deviated below 6.5, in order to measure reaction to any XIP released at late times into the culture. RFP and GFP fluorescence were measured by pipetting a small amount of the culture onto a glass coverslip and analyzing single cells. 99th percentile GFP fluorescence was then used to determine if XIP was being released to the *comS* mutants in an OD_{600}-dependent manner.

RT-qPCR measurement of *comS* and *comX* transcripts

S. mutans cells were diluted 20-fold into BHI (P*comX-gfp* WT/BHI, 184*comS PcomX-rfp*, 184comS P*comX-rfp* ∆comS samples) or FMC (P*comX-gfp* WT/FMC ± XIP, P*comX*-gfp ∆comS ± XIP samples). Where added, XIP was supplied at OD_{600} = 0.1. Cells were harvested at OD_{600} = 0.5 by centrifugation and re-suspended in RNA protectant buffer for 10 minutes. Samples were then centrifuged, the supernatant removed and the pellets frozen at -80 °C. RNA extraction was performed using the Qiagen RNEasy mini kit (Qiagen, USA). RNA sample concentration and purity were measured using a Thermo Scientific NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Scientific, USA). 1 µg of RNA was then reverse transcribed

to cDNA using the Bio-Rad iScript reverse transcription kit with random primers (Bio-Rad, USA). The qPCR was performed on a Bio-Rad CFX96 Real-Time System using Bio-Rad Sso Advanced Universal SYBR Green Supermix with a 50-fold dilution of the cDNA and 500 nM gene-specific primers. Nine replicates (3 biological × 3 technical) were performed for each condition. Sequences used for the primers are given in Table S3 (Supporting information). A standard curve across 8 orders of magnitude of transcript copies (from 10^8 to 10^1) was used to determine transcript count for each gene. For each sample the *comX* and *comS* transcript counts were normalized by the 16S rRNA count for the same sample. The results show the median of this ratio, with error bars indicating the range from the second lowest to second highest value among the replicates at each condition.

Fluorescence polarization

For fluorescence polarization studies of promoter binding, the *comR* gene was cloned into the 6x-His tagged site on pET-45b(+) vector in *E. coli* 10-beta using standard PCR cloning methods. His-ComR was then expressed in *E. coli* BL21(DE3) by induction with 1mM IPTG at mid-exponential phase in LB. After 4 hours cells were lysed using lysozyme in B-PER lysis buffer (ThermoFisher). Protein was then purified from clarified lysate using Ni-NTA agarose affinity chromatography and the histidine tag cleaved using enterokinase max at 4°C (EKMax, Invitrogen). The resulting protein solution was dialyzed into PBS pH 7.4 for experimental use. Native ComR concentration was measured by the Pierce BCA assay (Thermo Scientific) and purity of the cleaved form verified by SDS-PAGE run against an uncleaved sample.

Fluorescence polarization assays were performed in a 96-well plate with black bottom and black sides in a Biotek Synergy 2 plate reader (Biotek Instruments inc.) in the polarization mode. A 5' Bodipy FL-X labeled self-annealing stem-loop DNA strand with sequence corresponding to P*comX* (sequence 5'-BODIPY FL-X - ATGGGACATTTATGTCCTGTCCCCCACAGGACATAAATGTCCCAT - 3'), synthesized by ThermoFisher) was used as the binding aptamer and a filter set with excitation 485 nm, emission 528 nm was used for fluorescence excitation. 1 nM labeled DNA probe was added to a reaction buffer previously described [47] supplemented with 1 mM EDTA and 0.05 mg ml^{-1} salmon sperm DNA. ComR was titrated in concentration in this buffer alone, in the presence of 10 µM XIP or in the presence of 10 µM comS. The reactions were incubated at 37°C for 20 minutes before reading. Synthetic ComS (sequence MFSILTSILMGLDWWSL) for fluorescence polarization was synthesized and purified to 60% purity by Biomatik (Wilmington, DE).

Competition assays using unlabeled DNA probes were performed with 1.5 µM ComR in the same buffer, containing 1 nM P*comX* fluorescent DNA and 10 µM SHP (either ComS or XIP). An unlabeled probe corresponding to either the P*comS* (sequence 5' - ACG GGACATAAATGTCCTGTCCCCCACAGGACATTTATGTCCCGT - 3'), synthesized by Thermo Fisher) or the above P*comX* probe was titrated into this solution and the fluorescence polarization measured. Reactions were again incubated at 37°C for 20 minutes before measurement. In all FP experiments reading was performed three times on the same plate to estimate instrument error.

We compared the FP data to a two-step binding model in which the peptide ComS or XIP forms a multimeric complex with ComR (with dissociation constant k_1),

and then a single copy of this complex binds to the fluorescent DNA probe (with dissociation constant k_2), increasing its fluorescence anisotropy. The model is summarized by

$$P + R \rightleftharpoons C \qquad C + D \rightleftharpoons D^* \qquad C + U \rightleftharpoons U^*$$

Here *P* is the peptide (ComS or XIP), *R* is ComR, *C* is the peptide-ComR multimeric complex, *D* (*U*) is the free labeled (unlabeled) probe, *D** (*U**) is the labeled (unlabeled) probe with complex bound. The order of multimerization of the complex *C* is *n*. We solved the equilibrium equations for the model using the multivariate Newton-Raphson method in Matlab. We performed separate data analyses for the FP data ComS and XIP, respectively. In each analysis we searched for parameter values (k_1, k_2, n) that simultaneously minimized the sum of squares residuals for both the association (Figure 6A) and competition (Figure 6B) experiments for a given peptide *P*.

In general the FP data are compatible with a range of parameter values. If *n* is constrained to be less than 2.5 then optimal values are in the range $k_1 \sim$ 1-6 μM and $k_2 \sim$ 1-30 nM and $n \simeq$ 2-2.5 for for XIP interacting with ComR, and $k_1 \sim$ 3-20 μM and $k_2 \sim$ 30-200 nM and $n \simeq$ 1.6-2.5 for ComS interacting with ComR.

Mathematical model of *comRS* control of *comX*

Deterministic modeling of *comX* activation by *comRS* was performed by least squares fitting a chemical equilibrium model to the microfluidic data from experiments for each of the wild type background P*comX* GFP strain and the Δ*comS* cells using Matlab. The same offset and multiplicative factor were used to map calculated [ComX] onto the GFP fluorescence curves for both strains, as this is a property of the GFP, not

the gene circuit. In the case of the *comS* deficient strain, relevant parameters representing *comS* feedback and constitutive production were set to zero to obtain the different behavior observed. ComR was assumed to be present at around 15 copies per cell, as only modest changes in its expression were previously observed due to early competence inducing factors [53]. Exogenous XIP was taken to be a non-depleting reservoir. Details of the ODE system are in the supporting information section, and a table of parameters is available in supporting information Table S2.

Robustness of fit was tested through the bootstrap method, using the 90th and 10th percentile behavior of parameters to examine whether the transcriptional efficiency difference hypothesized was preserved in this range. Dependence on the initial parameter guess was checked by 50 iterations of adding a Gaussian-distributed random number with a mean of the best fit parameter and standard deviation half the best fit parameter to the start guess vector components used to find the best fit. New sets of fit parameters for each of these were then generated. It was found that the ComS-ComR complex elicited higher *comX* transcription in 100% of cases than did the XIP-ComR complex, and higher *comS* feedback stimulation (V^* parameters) in 78% of cases. Thus while numerous solutions to the system exist, the *comX* transcriptional efficiency discrepancy hypothesized is a generic property of the fit.

Acknowledgments

This work was supported by 1R01 DE023339 from the National Institute of Dental and Craniofacial Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank Minjun

Son, Chris Browngardt, Natalie Maricic, Lin Zeng, Hey Min Kim and Sang-Joon Ahn for helpful discussions, provision of mutant strains and advice on experimental procedures.

References

1. Loesche WJ. (1986) Role of *Streptococcus mutans* in human dental decay. Microbiol Rev 50(4): 353-380.

2. Senadheera DB, Cordova M, Ayala EA, Chavez de Paz LE, Singh K, et al. (2012) Regulation of bacteriocin production and cell death by the VicRK signaling system in *Streptococcus mutans*. J Bacteriol 194(6): 1307-1316.

3. Ahn SJ, Lemos JA, Burne RA. (2005) Role of HtrA in growth and competence of *Streptococcus mutans* UA159. J Bacteriol 187(9): 3028-3038.

4. Senadheera MD, Lee AW, Hung DC, Spatafora GA, Goodman SD, et al. (2007) The *Streptococcus mutans* vicX gene product modulates gtfB/C expression, biofilm formation, genetic competence, and oxidative stress tolerance. J Bacteriol 189(4): 1451-1458.

5. Tremblay YDN, Lo H, Li Y, Halperin SA, Lee SF. (2009) Expression of the *Streptococcus mutans* essential two-component regulatory system VicRK is pH and growth-phase dependent and controlled by the LiaFSR three-component regulatory system. Microbiology 155(9): 2856-2865.

6. Ahn SJ, Wen ZT, Burne RA. (2006) Multilevel control of competence development and stress tolerance in *Streptococcus mutans* UA159. Infect Immun 74(3): 1631-1642.

7. Qi F, Merritt J, Lux R, Shi W. (2004) Inactivation of the *ciaH* gene in *Streptococcus mutans* diminishes mutacin production and competence development, alters sucrose-dependent biofilm formation, and reduces stress tolerance. Infect Immun 72(8): 4895-4899.

8. Son M, Ghoreishi D, Ahn SJ, Burne RA, Hagen SJ. (2015) Sharply tuned pH response of genetic competence regulation in *Streptococcus mutans*: A microfluidic study of the environmental sensitivity of comX. Appl Environ Microbiol 81(16): 5622-5631.

9. Guo Q, Ahn SJ, Kaspar J, Zhou X, Burne RA. (2014) Growth phase and pH influence peptide signaling for competence development in *Streptococcus mutans*. J Bacteriol 196(2): 227-236.

10. Moye ZD, Son M, Rosa-Alberty AE, Zeng L, Ahn SJ, et al. (2016) Effects of carbohydrate source on genetic competence in *Streptococcus mutans*. Appl Environ Microbiol 82(15): 4821-4834.

11. Son M, Ahn S, Guo Q, Burne RA, Hagen SJ. (2012) Microfluidic study of competence regulation in *Streptococcus mutans*: Environmental inputs modulate bimodal and unimodal expression of comX. Mol Microbiol 86(2): 258-272.

12. Reck M, Tomasch J, Wagner-Döbler I. (2015) The alternative sigma factor SigX controls bacteriocin synthesis and competence, the two quorum sensing regulated traits in *Streptococcus mutans*. PLOS Genetics 11(7): e1005353.

13. Son M, Shields RC, Ahn S, Burne RA, Hagen SJ. (2015) Bidirectional signaling in the competence regulatory pathway of *Streptococcus mutans*. FEMS Microbiol Lett 362(19): fnv159-fnv159.

14. Stephen J Hagen and,Minjun Son. (2017) Origins of heterogeneity in *Streptococcus mutans* competence: Interpreting an environment-sensitive signaling pathway. Physical Biology 14(1): 015001.

15. Khan R, Rukke HV, Høvik H, Åmdal HA, Chen T, et al. (2016) Comprehensive transcriptome profiles of *Streptococcus mutans* UA159 map core streptococcal competence genes. mSystems 1.

16. Fontaine L, Wahl A, Fléchard M, Mignolet J, Hols P. (2015) Regulation of competence for natural transformation in streptococci. Infection, Genetics and Evolution 33: 343-360.

17. Hossain MS, Biswas I. (2012) An extracelluar protease, SepM, generates functional competence-stimulating peptide in *Streptococcus mutans* UA159. J Bacteriol 194(21): 5886-5896.

18. Perry JA, Jones MB, Peterson SN, Cvitkovitch DG, Lévesque CM. (2009) Peptide alarmone signalling triggers an auto-active bacteriocin necessary for genetic competence. Mol Microbiol 72(4): 905-917.

19. Fontaine L, Goffin P, Dubout H, Delplace B, Baulard A, et al. (2013) Mechanism of competence activation by the ComRS signalling system in streptococci. Mol Microbiol 87(6): 1113-1132.

20. Mashburn-Warren L, Morrison DA, Federle MJ. (2010) A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an rgg regulator. Mol Microbiol 78(3): 589-606.

21. Desai K, Mashburn-Warren L, Federle MJ, Morrison DA. (2012) Development of competence for genetic transformation of *Streptococcus mutans* in a chemically defined medium. J Bacteriol 194(15): 3774-3780.

22. Khan R, Rukke HV, Ricomini Filho AP, Fimland G, Arntzen MO, et al. (2012) Extracellular identification of a processed type II ComR/ComS pheromone of *Streptococcus mutans*. J Bacteriol 194(15): 3781-3788.

23. Wenderska IB, Lukenda N, Cordova M, Magarvey N, Cvitkovitch DG, et al. (2012) A novel function for the competence inducing peptide, XIP, as a cell death effector of *Streptococcus mutans*. FEMS Microbiol Lett 336(2): 104-112.

24. Chang JC, Federle MJ. (2016) PptAB exports rgg quorum-sensing peptides in *Streptococcus*. Plos One 11(12): e0168461.

25. Gardan R, Besset C, Gitton C, Guillot A, Fontaine L, et al. (2013) Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. Journal of Bacteriology 195(8): 1845-1855.

26. Kaspar J, Underhill SAM, Shields RC, Reyes A, Rosenzweig S, et al. (2017) Intercellular communication via the *comX*-inducing peptide (XIP) of *Streptococcus mutans*. J Bacteriol.

27. Dubnau D, Losick R. (2006) Bistability in bacteria. Mol Microbiol 61(3): 564-572.

28. Taniguchi Y, Choi PJ, Li G, Chen H, Babu M, et al. (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. Science 329(5991): 533.

29. Friedman N, Cai L, Xie XS. (2006) Linking stochastic dynamics to population distribution: An analytical framework of gene expression. Phys Rev Lett 97(16): 168302.

30. Ahn SJ, Burne RA. (2006) The *atlA* operon of *Streptococcus mutans*: Role in autolysin maturation and cell surface biogenesis. J Bacteriol 188(19): 6877-6888.

31. Bareia T, Pollak S, Eldar A. (2018) Self-sensing in *Bacillus subtilis* quorum-sensing systems. Nat. Microbiol. 3: 83-89.

32. Maamar H, Dubnau D. (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. Mol Microbiol 56(3): 615-624.

33. Shanker E, Morrison DA, Talagas A, Nessler S, Federle MJ, et al. (2016) Pheromone recognition and selectivity by ComR proteins among *Streptococcus* species. PLOS Pathogens 12: e1005979.

34. Aggarwal C, Jimenez JC, Nanavati D, Federle MJ. (2014) Multiple length peptide-pheromone variants produced by *Streptococcus pyogenes* directly bind rgg proteins to confer transcriptional regulation. J Biol Chem 289(32): 22427-22436.

35. Talagas A, Fontaine L, Ledesma-Garca L, Mignolet J, Li de lS, et al. (2016) Structural insights into streptococcal competence regulation by the cell-to-cell communication system ComRS. PLOS Pathogens 12(12): e1005980.

36. Lupp C, Ruby EG. (2004) Vibrio fischeri LuxS and AinS: Comparative study of two signal synthases. J Bacteriol 186(12): 3873-3881.

37. Hurley A, Bassler BL. (2017) Asymmetric regulation of quorum-sensing receptors drives autoinducer-specific gene expression programs in *Vibrio cholerae*. PLOS Genetics 13(5): e1006826.

38. Becskei A, Séraphin B, Serrano L. (2001) Positive feedback in eukaryotic gene networks: Cell differentiation by graded to binary response conversion. Embo J 20: 2528.

39. Youk H, Lim WA. (2014) Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. Science 343(6171): 1242782.

40. Sayut DJ, Niu Y, Sun L. (2006) Construction and engineering of positive feedback loops. ACS Chem Biol 1: 692-696. 10.1021/cb6004245.

41. Nistala GJ, Wu K, Rao CV, Bhalerao KD. (2010) A modular positive feedback-based gene amplifier. J Biol Eng 4: 4-1611-4-4.

42. Leung V, Dufour D, Lévesque C,M. (2015) Death and survival in *Streptococcus mutans*: Differing outcomes of a quorum-sensing signaling peptide. Frontiers in Microbiology 6: 1176.

43. Perry JA, Cvitkovitch DG, Lévesque CM. (2009) Cell death in *Streptococcus mutans* biofilms: A link between CSP and extracellular DNA. FEMS Microbiol Lett 299(2): 261-266.

44. Terleckyj B, Willett NP, Shockman GD. (1975) Growth of several cariogenic strains of oral streptococci in a chemically defined medium. Infect Immun 11(4): 649-655.

45. Biswas I, Jha JK, Fromm N. (2008) Shuttle expression plasmids for genetic studies in *Streptococcus mutans*. Microbiology 154: 2275-2282.

46. LeBlanc DJ, Lee LN, Abu-Al-Jaibat A. (1992) Molecular, genetic, and functional analysis of the basic replicon of pVA380-1, a plasmid of oral streptococcal origin. Plasmid 28(2): 130-145.

47. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77(1): 51-59.

48. Cha RS, Zarbl H, Keohavong P, Thilly WG. (1992) Mismatch amplification mutation assay (MAMA): Application to the c-H-ras gene. PCR Methods Appl 2(1): 14-20.

49. Zeng L, Das S, Burne RA. (2010) Utilization of lactose and galactose by *Streptococcus mutans*: Transport, toxicity, and carbon catabolite repression. J Bacteriol 192(9): 2434-2444.

50. Ahn SJ, Kaspar J, Kim JN, Seaton K, Burne RA. (2014) Discovery of novel peptides regulating competence development in *Streptococcus mutans*. J Bacteriol 196(21): 3735-3745.

51. Jeon NL, Dertinger SKW, Chiu DT, Choi IS, Stroock AD, et al. (2000) Generation of solution and surface gradients using microfluidic systems. Langmuir 16(22): 8311-8316.

52. Kwak IH, Son M, Hagen SJ. (2012) Analysis of gene expression levels in individual bacterial cells without image segmentation. Biochem Biophys Res Commun 421(3): 425-430.

53. Lemme A, Grobe L, Reck M, Tomasch J, Wagner-Dobler I. (2011) Subpopulation-specific transcriptome analysis of competence-stimulating-peptide-induced *Streptococcus mutans*. J Bacteriol 193(8): 1863-1877.

Figure captions

Fig. 1: *comS* deletion is not fully complemented by synthetic XIP

(A) Comparison of P*comX-gfp* activity in *S. mutans* cells of the UA159 (wild type) background (magenta) and Δ*comS* mutant background (cyan). The median GFP fluorescence is shown for cells that were supplied with continuous flow of exogenous synthetic XIP in microfluidic chambers. Data are shown at 30 minutes (circles, dash-dot lines) and 90 minutes (squares, dashed lines) of flow. The smooth curves are spline fits to the data. (B) Median GFP levels in the two strains at the 90 minute time point of the flow experiment. Also shown are the histograms of the individual cell P*comX-gfp* reporter activity, versus exogenous XIP concentration, for (C) the UA159 background and (D) the Δ*comS* background. Solid black curves in (C) and (D) show the best fit gamma probability distribution for each histogram. A cutoff of 40 units of P*comX* GFP fluorescence has been applied to exclude background autofluorescence. (E) Parameter *a* of the (two parameter) gamma probability distribution, obtained from fits in (C)-(D), reflecting the ratio of transcription rate to protein degradation rate, and (F) the distribution parameter *b*, reflecting the ratio of translation rate to the rate of mRNA degradation. In (E)-(F) cyan indicates the Δ*comS* mutant, and magenta indicates the UA159 background.

Fig. 2: Activation of ComX in a ComS-overexpressing strain is independent of rate of medium replacement

P*comX-rfp* reporter activity is shown in cells growing in microfluidic chambers supplied with continuously flowing fresh complex medium (BHI). (A) Histograms of individual cell

P*comX-rfp* reporter fluorescence: (Leftmost column) wild type background (negative control) with flow at 0.1 ml h^{-1}; (Second column) ComS-overexpressing 184*comS* Δ*comS* background at 0.1 ml h^{-1}; (Columns 3-5) ComS-overexpressing (184*comS*) background at 0.02 ml h^{-1}, 0.1 ml h^{-1}, 0.5 ml h^{-1} and 1 ml h^{-1}. (At a flow rate of 1 ml h^{-1} the medium in each flow chamber is replaced every 6 s.) (B) RFP fluorescence of cells that exceeded the wild type (negative control) red fluorescence in column 1 of (A). The black bar indicates the median of data in each channel: (Leftmost column) 184*comS* Δ*comS* background; (Columns 2-5) ComS-overexpressing (184*comS*) background. All RFP measurements were made 4 hours after addition of chloramphenicol to the cultures.

Fig. 3: chromosomal *comS* increases the amount of *comS* transcript produced in response to XIP

RT-qPCR measurement of (A) *comS* and (B) *comX* transcripts in cultures harvested at OD_{600} = 0.5. Each bar indicates the ratio of the median transcript count to the median 16S rRNA count, as measured in multiple biological and technical replicates (see *Methods*). WT/BHI, 184*comS*, and 184*comS* Δ*comS* samples were grown in BHI medium. Remaining samples were grown in FMC. XIP was supplied at 100 nM concentration when used. Error bars indicate the range from second lowest to second highest value among the replicates for each condition.

Fig. 4: ComS overexpresser does not induce *comX* of Δ*comS* strain in coculture

Histograms of (A) GFP and (B) RFP fluorescence of individual cells in cocultures of sender (184*comS* P*comX-rfp*) and receiver (P*comX-gfp* Δ*comS*) strains. Strains were grown to equal optical density, mixed in varying proportion, and then incubated in microfluidic chambers containing stationary defined medium. Blue lines indicate population medians. Panels (A) and (B) show fluorescence of: (Leftmost column) UA159 background strain containing P*comX rfp* reporter, without added XIP (negative control); (Second column) P*comX-gfp* Δ*comS* (receiver) with added 50 nM XIP (positive control); (Columns 3-11) Cocultures of sender and receiver, with columns labeled by percentage by volume of 184*comS* (sender) culture in the initial preparation of the coculture. (C-H) Phase contrast images of cocultures, overlaid with red and green fluorescence images: (C) P*comX-rfp* reporter in UA159 background, with no added XIP (negative control); (D) P*comX-gfp* Δ*comS* cells with 50 nM added XIP (positive control); (E) P*comX-gfp* Δ*comS* (receiver) alone, with 0% sender; (F)-(H) cocultures containing 30%, 80%, and 100% sender respectively.

Fig. 5: Evidence for release of XIP in cocultures late in growth

Histograms of (A) GFP and (B) RFP fluorescence of individual cells in cocultures of receiver (P*comX gfp* Δ*comS*) and sender (P*comX rfp* 184*comS*) strains, following different incubation periods. Labels at top indicate the volume fraction comprised by the sender strain in the preparation of the coculture. Histogram colors indicate incubation times: 2 h (cyan), 8 h (magenta), 12 h (green). The lower panels show the 99th percentile of the individual cell GFP (C) and RFP (D) fluorescence observed in the

culture, versus the cultureOD_{600}. Exogenous XIP was not added, except in the positive (sender-only) control sample indicated by the inverted triangles in (C) and (D).

Fig. 6: ComS and XIP interact with ComR to bind the *comX* promoter

Fluorescence polarization assay testing interaction of ComS and XIP peptides with ComR and the P*comX* transcriptional activation site. The DNA probe is labeled with a Bodipy FL-X fluorophore. (A) Titration of ComR into a solution containing 10 µM of XIP (blue) or full length ComS (green), labeled DNA probe (1 nM), and 0.05 mg ml^{-1} salmon DNA. Negative control (black) contains no ComS or XIP peptide. (B) Competition assay in which unlabeled promoter sequence DNA was titrated into a solution containing ComR (1.5 µM), fluorescent DNA probe (1 nM) and peptide (either ComS or XIP, 10 µM) and 0.05 mg ml^{-1} salmon DNA. Unlabeled P*comX* DNA was used with XIP (blue) and ComS (green). An unlabeled P*comS* probe was also tested for its ability to compete with the fluorescent P*comX* probe in the presence of XIP (red) and ComS (gold). Solid curves indicate binding and competition behavior predicted by the two step model described in *Methods*, in which peptide (ComS or XIP) first forms a multimeric complex with ComR (k_1, n), and a single copy of this complex binds to the (labeled or unlabeled) P*comX* DNA (k_2). For ComS binding/competition (green), the curves represent k_1 = 3.2 µM, k_2 = 2.2 nM, n = 2.5. For XIP binding/competition (blue), the curves represent k_1 = 7.3 µM, k_2 = 33 nM, n = 2.4.

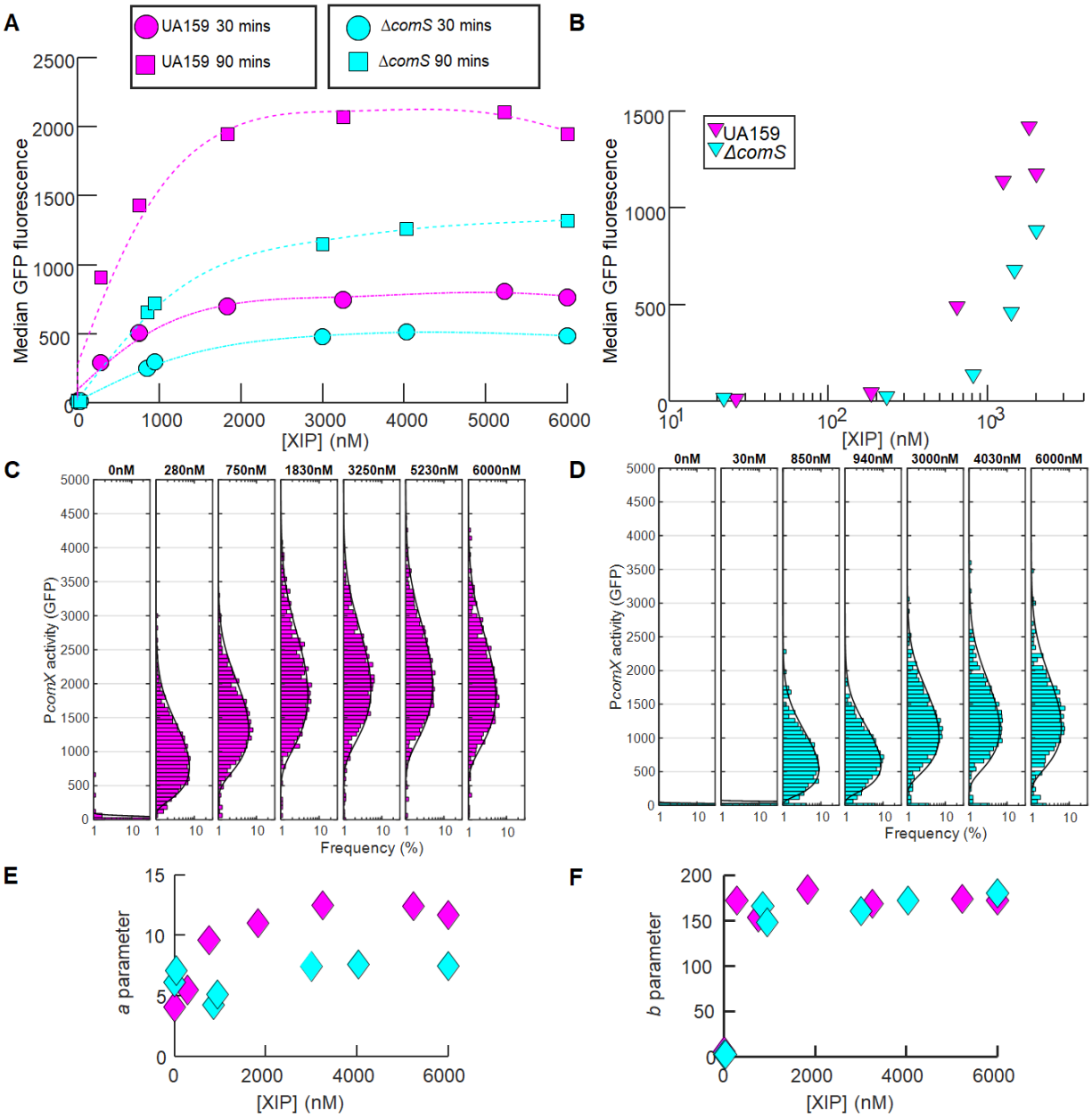
Fig. 7: proposed model for *comRS* regulation of *comX*.

Model for *comS*-feedback enhanced activation of P*comX-gfp* reporter by exogenous synthetic XIP. (A) Illustration of the feedback model and the role of *comS* and extracellular XIP. ComS and XIP both interact with ComR to activate transcription of *comS* and *comX.* At low concentrations of extracellular XIP, *comS* expression is very low and *comX* is not expressed. At high concentrations of extracellular XIP, XIP is imported efficiently by OppA and interacts with ComR to drive expression of both *comS* and *comR*. Endogenously produced ComS is not readily exported in the absence of lysis, and so intracellular accumulation of ComS drives elevated *comS* and *comX* expression. Consequently *comX* expression at any given XIP concentration is boosted by *comS* feedback. Cells lacking native *comS* can respond to synthetic XIP but cannot activate *comX* to the same level as wild type. The figure describes behavior in defined medium; In complex medium extracellular XIP is not imported [11]. (B) Comparison of model simulation with data. Red circles indicate median P*comX-gfp* fluorescence of the UA159 background strain supplied with synthetic XIP in microfluidic flow; blue circles indicate the median P*comX-gfp* fluorescence of the Δ*comS* background. Solid curves represent calculated values from a fit in which 11 parameters were fit to the microfluidic data, as described in *Methods.* The model relates the predicted ComX concentration to the median GFP fluorescence by an offset and scale factor.

Figures

1025

Figure 1



1026

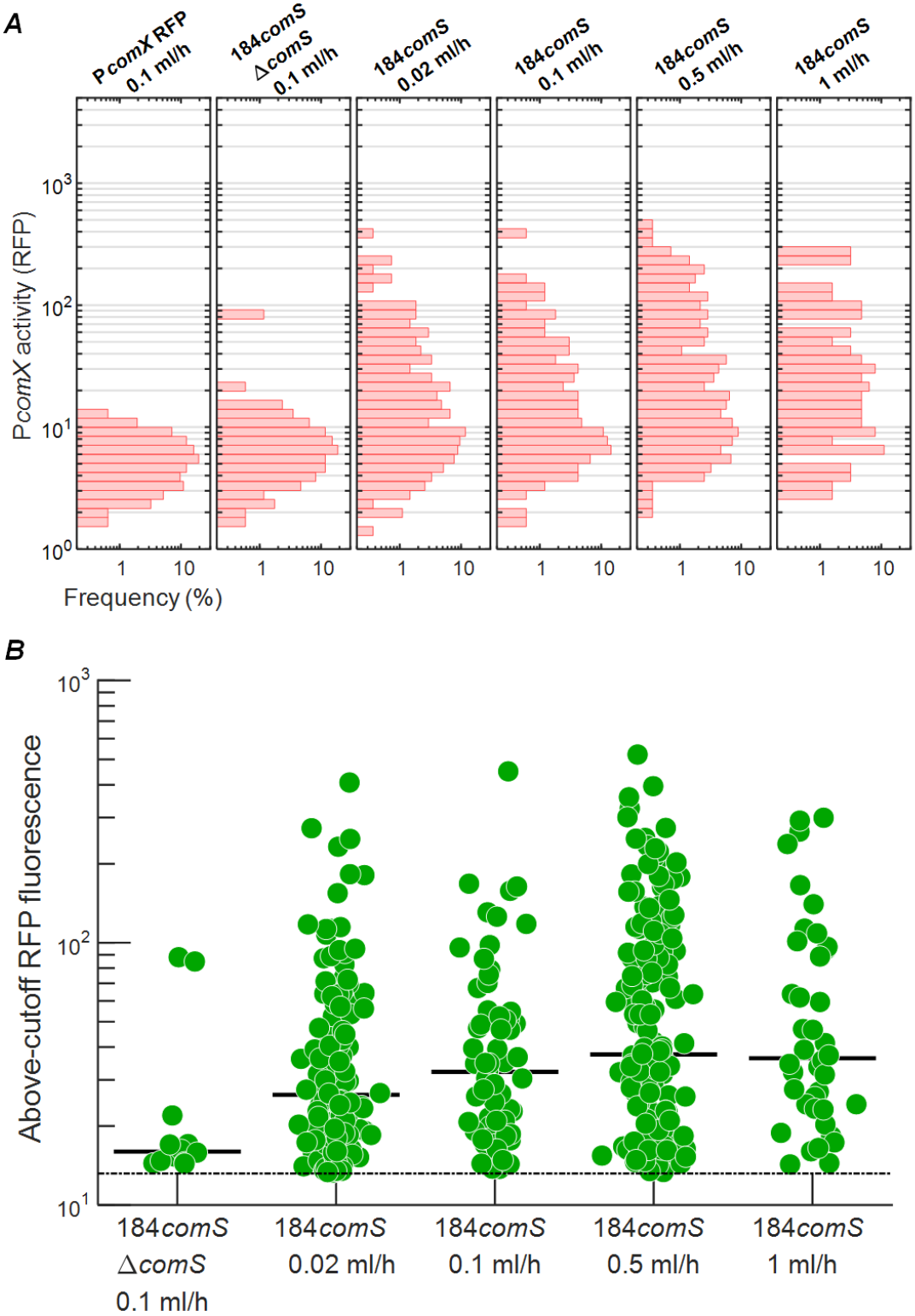
1027

1028

1029

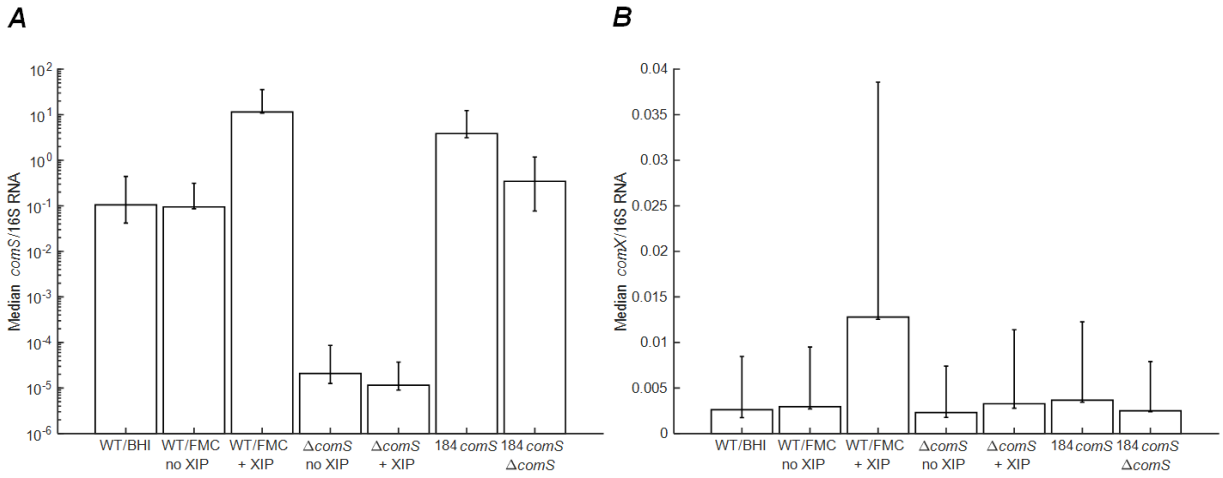
1030

Figure 2


A
P comX RFP 0.1 ml/h
184comS ΔcomS 0.1 ml/h
184comS 0.02 ml/h
184comS 0.1 ml/h
184comS 0.5 ml/h
184comS 1 ml/h
PcomX activity (RFP)
Frequency (%)
B
Above-cutoff RFP fluorescence
184comS ΔcomS 0.1 ml/h
184comS 0.02 ml/h
184comS 0.1 ml/h
184comS 0.5 ml/h
184comS 1 ml/h

1035

Figure 3



1036

1037

Figure 4

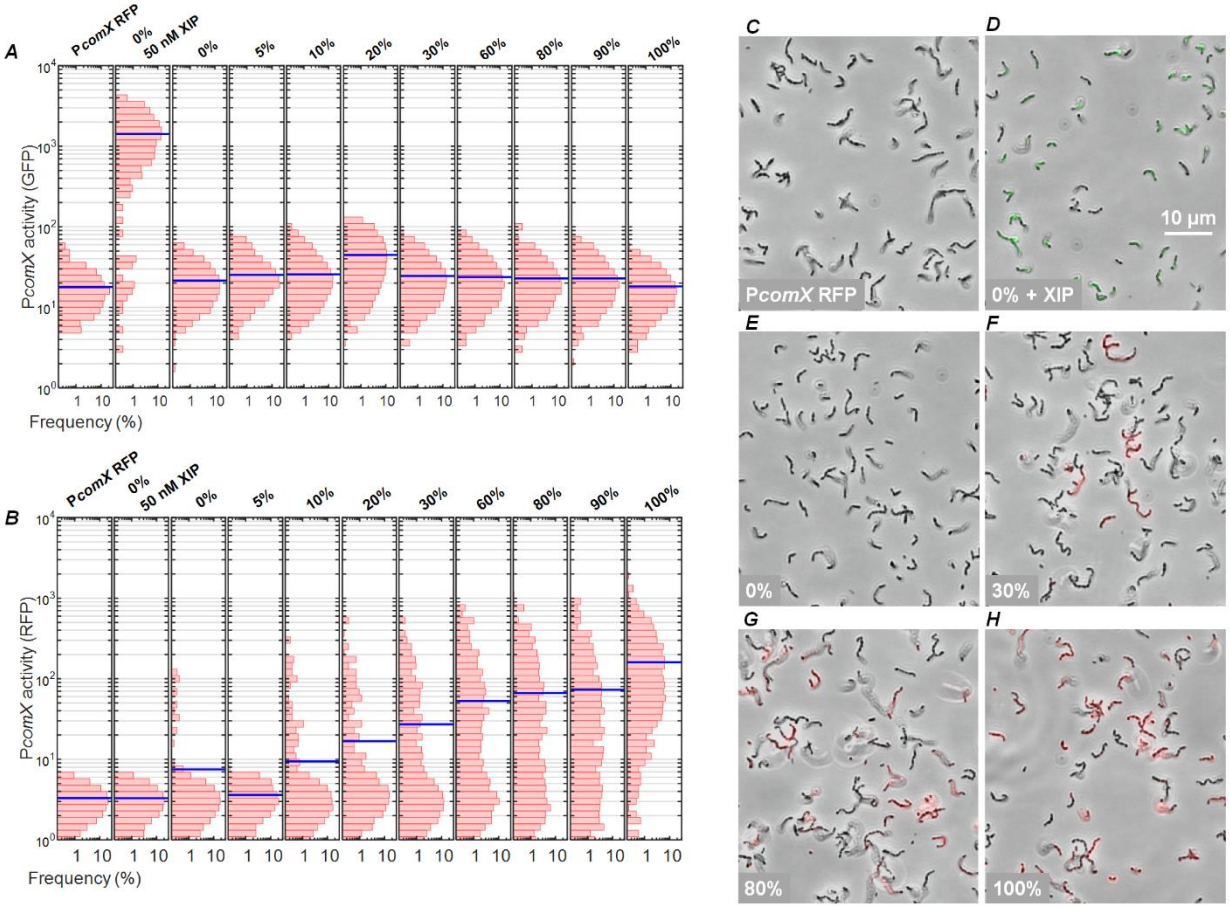
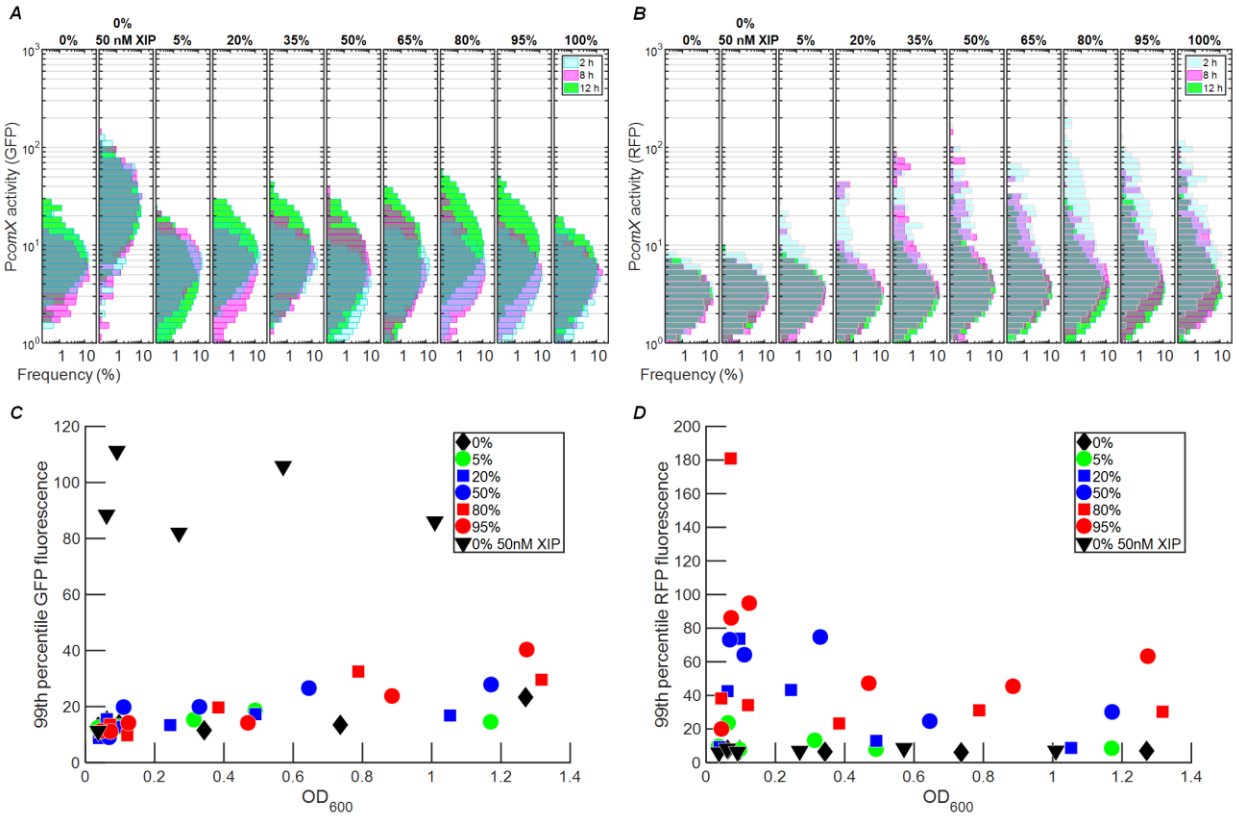
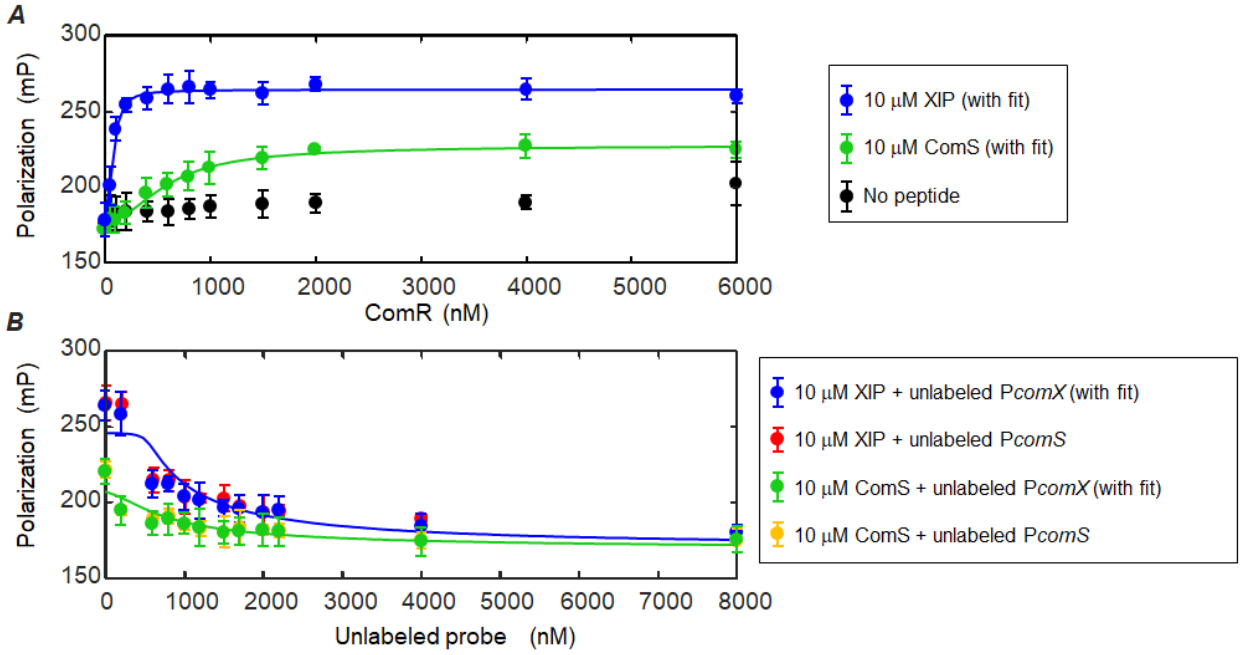

A
PcomX RFP
0%
50 nM XIP
0%
5%
10%
20%
30%
60%
80%
90%
100%
PcomX activity (GFP)
Frequency (%)
B
PcomX activity (RFP)
Frequency (%)
C
PcomX RFP
D
10 µm
0% + XIP
E
0%
F
30%
G
80%
H
100%

Figure 5


A
B
C
D
0%
50 nM XIP
5%
20%
35%
50%
65%
80%
95%
100%
2 h
8 h
12 h
PcomX activity (GFP)
PcomX activity (RFP)
Frequency (%)
99th percentile GFP fluorescence
99th percentile RFP fluorescence
OD600
0% 50nM XIP

1045

Figure 6

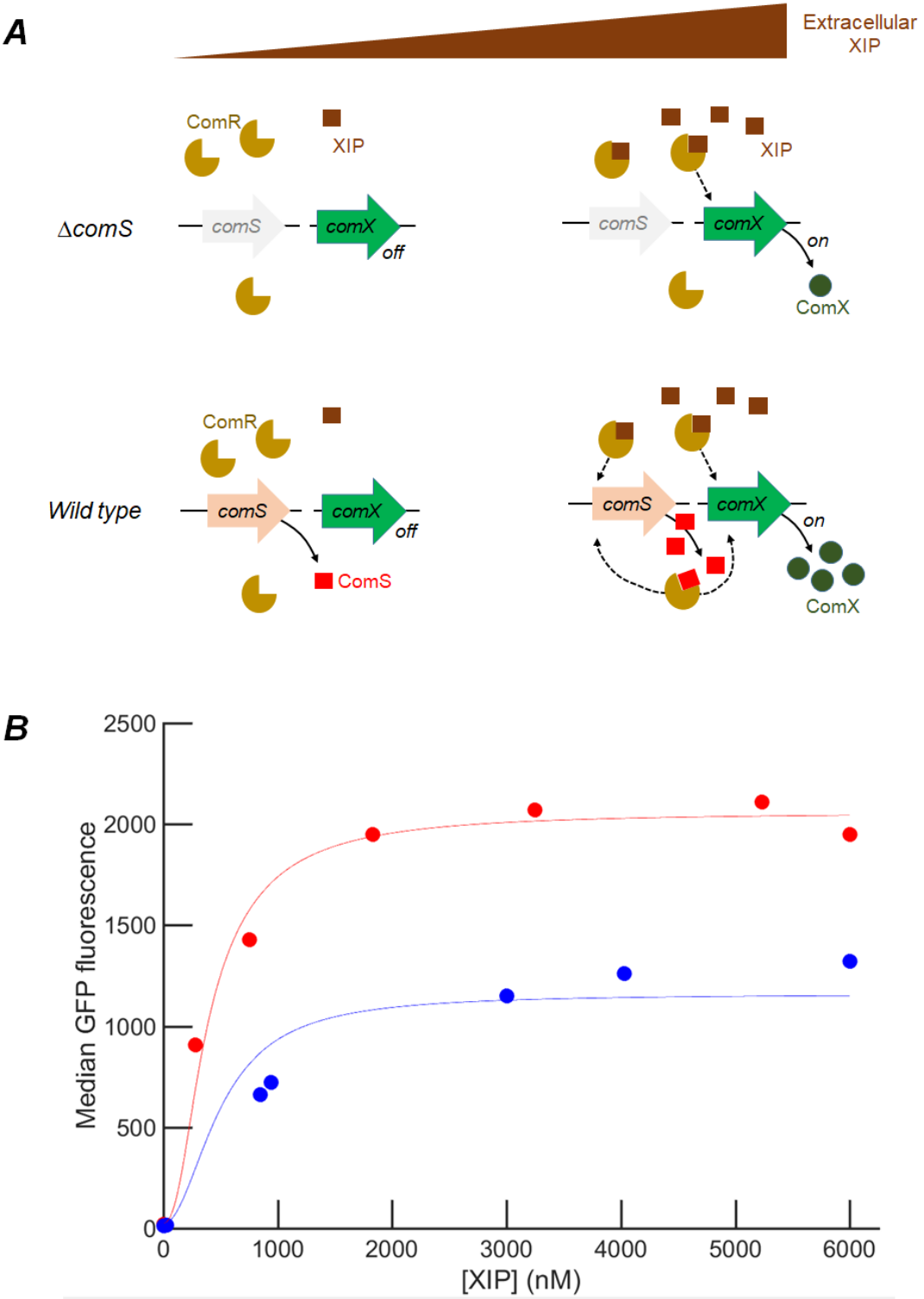


1046

1047

1048

Figure 7


A
Extracellular XIP
ΔcomS
ComR
XIP
comS
comX
off
on
ComX
Wild type
ComS
B
Median GFP fluorescence
2500
2000
1500
1000
500
0
0
1000
2000
3000
4000
5000
6000
[XIP] (nM)