Discovery of the first small molecule CsrA-RNA interaction inhibitors using biophysical screening technologies

Christine K Maurer^{1,‡}, Martina Fruth^{1,‡}, Martin Empting¹, Olga Avrutina², Jörn Hoßmann³, Suvd Nadmid⁴, Jan Gorges⁵, Jennifer Herrmann⁴, Uli Kazmaier⁵, Petra Dersch³, Rolf Müller⁴ & Rolf W Hartmann^{1,6,*}

¹ Department of Drug Design and Optimization, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), Saarland University, Campus E8.1, 66123 Saarbrücken, Germany. German Centre for Infection Research (DZIF), Partner Site Hannover-Braunschweig, 30625 Hannover, Germany.

² Clemens-Schöpf-Institute for Organic Chemistry and Biochemistry, Technical University Darmstadt, Alarich-Weiss Straße 4, 64287 Darmstadt, Germany.

³ Department of Molecular Infection Biology, Helmholtz Centre for Infection Research (HZI), Inhoffenstraße 7, 38124 Braunschweig, Germany. German Centre for Infection Research (DZIF), Partner Site Hannover-Braunschweig, 30625 Hannover, Germany.

⁴ Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), Saarland University, Campus E8.1, 66123 Saarbrücken, Germany. German Centre for Infection Research (DZIF), Partner Site Hannover-Braunschweig, 30625 Hannover, Germany.

⁵ Institute of Organic Chemistry I, Saarland University, Campus C4.2, 66123 Saarbrücken, Germany.

⁶ Department of Pharmaceutical and Medicinal Chemistry, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany.

[‡] These authors contributed equally to this work.

* Author for correspondence:

Tel.: +49 681 98806 2000

Fax: +49 681 98806 2009

rolf.hartmann@helmholtz-hzi.de

Structured abstract

Background: The carbon storage regulator A (CsrA) is a global post-transcriptional regulator affecting mRNA translation and/or stability. Widespread among bacteria, CsrA is essential for their full virulence thus representing a promising anti-infective drug target.

Results/Methodology: For discovery of CsrA-RNA interaction inhibitors, we followed two strategies: a screening of small molecules (A) and an RNA ligand-based approach (B). Using surface plasmon resonance-based binding and fluorescence polarization-based competition assays, (A) yielded 7 small molecule inhibitors, among them **MM14** (IC₅₀ of 4 μ M). (B) resulted in RNA-based inhibitor **GGA_{RNA}** (IC₅₀ of 113 μ M).

Conclusions: The first small molecule inhibitors of the CsrA-RNA interaction were discovered exhibiting micromolar affinities. These hits represent tools to investigate the effects of CsrA-RNA interaction inhibition on bacterial virulence.

Keywords

CsrA/RsmA protein family; anti-virulence strategy; protein-nucleic acid interaction; surface plasmon resonance spectroscopy; fluorescence polarization; library screening

Introduction

To secure their survival, bacteria need to sense and react to changing conditions in their environment. Hence, they evolved multi-layered mechanisms to adapt to alterations in their surroundings, which are controlled by complex regulatory networks [1,2]. The carbon storage regulator (Csr) system (alternatively also called regulator system of secondary metabolites Rsm) represents such a network that coordinates a multitude of different physiological, metabolic, and stress response-related processes within bacteria via post-transcriptional regulation mechanisms [3,4]. A key player of this system is the CsrA/RsmA protein. CsrA is widespread and highly conserved among bacteria, but is not present in eukaryotes [3,5,6]. By binding to their 5'-untranslated region, CsrA/RsmA proteins affect translation and stability of mRNA targets [7,8]. Predominantly, this protein-RNA interaction provokes a negative regulation and represses translation of target transcripts by competing with ribosome binding (Figure 1 and refs. [9,10]). Only a few examples of positive regulation mediated by CsrA have been reported in the literature, of which the molecular mechanisms are known [11,12]. For instance, binding of CsrA to the leader region of *Escherichia coli* *flhDC* mRNA appears to stabilize the transcript and to prevent its RNase E-mediated degradation thereby activating the expression of *flhDC* [13].

Previously, the three-dimensional (3D) structures of CsrA from *E. coli* and orthologs of different species have been solved via solution NMR [14-16] and crystallographic studies [6,17,18]. It became apparent that two CsrA monomers, each composed of five β -strands and one α -helix, intertwine to form a symmetrical homodimer comprising a hydrophobic core and two identical RNA-binding surfaces. Thereby, the side chains contributing to sequence-specific RNA recognition are highly conserved among the CsrA protein family with the Arg44 residue being critical for *in vitro* RNA binding [6,15]. Regarding the RNA ligand, *in vitro* selection [19] and NMR [15,16,20] studies have identified a conserved ANGGA(N) core binding motif. This recognition sequence contains the essential GGA motif and is preferentially presented within a hairpin loop [19]. The N-nucleotides as well as those flanking the core motif are variable and allow for modulation of binding affinity over several orders of magnitude [16,20].

Extensive studies over the last years demonstrated that CsrA and its homologs play an important role in coordinating the expression of bacterial virulence factors required for successful host infection (reviewed in [3,4]). For instance, several virulence factors involved in cell attachment and host invasion are mediated by the Csr system in different bacterial species such as enteropathogenic *E. coli*, *Yersinia pseudotuberculosis*, *Helicobacter pylori*, and *Salmonella enterica* serovar Typhimurium [21-24]. In some pathogens such as *Pseudomonas aeruginosa* and *Y. pseudotuberculosis*, this protein family manages the transition between early and late stages of infection [3,22,25]. Accordingly, virulence factors associated with acute infection (e.g. motility, host colonization, invasion) are positively regulated, whereas those involved in chronic infection are negatively regulated (e.g. biofilm formation). Importantly, it has been shown that deficiency of CsrA or its ortholog RsmA leads to significantly attenuated virulence in murine models of *P. aeruginosa* [26], *H. pylori* [23], and *Y. pseudotuberculosis* (Heroven & Dersch, unpublished observations) infection mainly due to impaired colonization.

Hence, CsrA represents a promising target for anti-infective drug development (Figure 1). To date, no small molecule inhibitors of the CsrA-RNA interaction have been reported [20]. In the current work, we present our strategy for the discovery of CsrA-RNA interaction inhibitors, which incorporates a screening of small molecules and an RNA ligand-based design approach. The first step in the discovery of such inhibitors is the establishment of suitable test systems enabling their identification and characterization. Thus, we developed a surface plasmon resonance (SPR)-based binding assay and an orthogonal fluorescence polarization (FP)-based competition assay detecting compounds that interfere with the CsrA-RNA interaction. Utilizing these assays, a set of ~ 1,000 compounds was screened. Complementary to the screening, we applied a ligand-based strategy exploiting the existing knowledge of the CsrA-RNA interaction based on *in vitro* selection data [19] and 3D structures [6,15,16]. From the RNA consensus sequence, oligonucleotides of varying length and nucleic acid backbone were derived and their ability to inhibit the CsrA-RNA interaction was systematically studied.

Materials & methods

Oligonucleotides and chemicals

The DNA primers as well as the following HPLC-purified single-stranded RNA oligomers were purchased from Sigma-Aldrich (Schnelldorf, Germany): RNA_A (5'-UCACGGAUGA-3'), RNA_B (5'-UUCACGGAGAA-3'), and RNA_C (5'-UUCACAGAGAA-3') either unlabeled or fluorescein-labeled (*) at the 3' end with a six-carbon spacer between the RNA and the fluorophore. The HPLC-purified single-stranded oligonucleotides listed below were obtained from biomers.net (Ulm, Germany): GG_{RNA} (5'-GG-3', RNA-backbone), GA_{RNA} (5'-GA-3', RNA-backbone), GGA_{RNA} (5'-GGA-3', RNA-backbone), GGA_{DNA} (5'-GGA-3', DNA-backbone), GGA_{2'OMe-RNA} (5'-GGA-3', 2'-methoxy-RNA backbone), AGGGA_{DNA} (5'-AGGGA-3', DNA-backbone), AGGGA_{2'OMe-RNA} (5'-AGGGA-3', 2'-methoxy-RNA backbone). Kanamycin sulfate, chloramphenicol, biotin, isopropyl-β-D-thiogalactoside (IPTG), cOmplete™ protease inhibitor cocktail tablets, diethylpyrocarbonate (DEPC), Tween 20, bovine serum albumin (BSA), lysozyme from chicken egg white, and salts were purchased from Sigma Aldrich (Schnelldorf, Germany). Dimethylsulfoxide (DMSO) of analytical grade was obtained from VWR (Darmstadt, Germany), peptone from casein from Merck (Darmstadt, Germany), and yeast extract from Fluka (Neu-Ulm, Germany). In general, DEPC-treated water was used for experiments dealing with RNA.

Cloning of biotin tag into CsrA overexpression vector

The CsrA overexpression vector pAKH172 containing the *csrA* gene from *Y. pseudotuberculosis* YPIII was constructed by PCR of the *csrA* gene using forward primer I68

5'-CGGCGCCTCGAGGTAAGTCGTCGGTTGAGAC-3' containing an *NcoI* site and reverse primer IV783 5'-gcgccCCATGGATGCTTATTCTGACTCG-3' containing an *XhoI* site. The PCR fragment was digested with *NcoI* and *XhoI*, and ligated into pET28a. The generated plasmid pAKH172 was digested with *XhoI* (New England Biolabs, Frankfurt, Germany) and dephosphorylated with FastAP alkaline phosphatase (Thermo Fisher Scientific, Schwerte, Germany). DNA primers comprising the biotin-tag sequence flanked by respective *XhoI* restriction site overhangs (primer sequences (5'-3'): *forward* TCGAGGGCCTGAACGATATTTTTGAAGCGCAGAAAATTGAATGGCATGAGC; *reverse* GCTCATGCCATTCAATTTTCTGCGCTTCAAAAATATCGTTCAGGCCCTCGA) were mixed at equimolar concentrations, heated to 95 °C for 5 min, and allowed to cool down to room temperature for 2 h. These hybridized oligomers were ligated into the *XhoI* restriction site of the processed pAKH172 vector using T4 DNA ligase (Thermo Fisher Scientific, Schwerte, Germany) yielding plasmid pAKH172_biotag.

Expression and purification of biotinylated CsrA

For preparation of biotinylated CsrA, protein expression was coupled to *in vivo* biotinylation using a protocol adapted from ref. [27]. The following plasmids were double-transformed into *E. coli* BL21 (λ DE3): pAKH172_biotag for overexpression of CsrA containing the biotin-tag and a His-tag at the C-terminus and pBirAcm isolated from strain *E. coli* K12 AVB99 (Avidity LLC, Aurora, USA) harboring the IPTG-inducible *birA* gene to overexpress the biotin ligase. The generated *E. coli* strain containing the two plasmids was grown overnight in LB medium (0.5% NaCl [w/v], 0.5% yeast extract [w/v], 1% peptone from casein [w/v]) containing 50 μ g/mL kanamycin and 17 μ g/mL chloramphenicol at 37 °C with agitation. Fresh, antibiotic-supplemented LB medium was inoculated with the overnight culture and incubated at 37 °C until an OD of 0.6 was reached. Then, expression of biotinylated CsrA was induced with 1 mM IPTG, 10 mM MgCl₂, and 50 μ M biotin. The culture was incubated for 17 h at 37 °C and, subsequently, cells were harvested at 3,400 $\times$ g for 15 min at 4 °C. The pellets were resuspended in lysis buffer (50 mM K₂HPO₄, 300 mM NaCl, 10 mM imidazole, 0.2 mg/mL lysozyme, cOmplete™ protease inhibitor cocktail, pH 8.0) and cells were lysed using a microfluidizer (three passages). Cell debris was removed by centrifugation (45 min at 27,000 $\times$ g) and the filtered supernatant was loaded to a 1 mL HisTrap HP column (GE Healthcare) at 0.75 mL/min on an ÄKTExpress system (GE Healthcare). The column was equilibrated with binding buffer (50 mM K₂HPO₄, 300 mM NaCl, 10 mM imidazole, pH 8.0) until the UV signal was stable. Subsequently, it was washed with 15 column volumes of high salt binding buffer (50 mM K₂HPO₄, 1 M NaCl, 10 mM imidazole, pH 8.0) and 15 column volumes of binding buffer containing 30 mM imidazole. The protein was eluted with a linear gradient to 100% elution buffer (50 mM K₂HPO₄, 300 mM NaCl, 250 mM imidazole, pH 8.0) in 80 column volumes at 1 mL/min. Protein-containing fractions, protein identity, and purity were assessed by SDS-PAGE and LC-MS. Combined CsrA fractions were concentrated via centrifugal filtration using Vivaspin 6 columns (MWCO 3,000 Da, Sartorius). Buffer exchange into storage buffer (50 mM K₂HPO₄, 300 mM NaCl, 10% glycerol [v/v], in DEPC-treated H₂O, pH 8.0) was performed using PD-10 columns (GE Healthcare) according to the manufacturer's instructions. Protein concentrations were determined by both UV

spectroscopy ($\epsilon_{280\text{ nm}} = 8480\text{ M}^{-1}\text{ cm}^{-1}$) and Bradford protein assay using Roti®-Quant reagent (Carl Roth, Karlsruhe, Germany). Absence of RNase activity in the protein preparations was assessed by performing functional assay (FP-based saturation assay) with RNA_A in presence and absence of RiboLock RNase inhibitor (Thermo Fisher Scientific, Schwerte, Germany). Protein fractions were accepted as free of relevant RNase activity since no significant difference was observed between both conditions (data not shown). Protein aliquots were stored at -20 °C.

Screening library

The screening library consisted of 478 compounds obtained from AnalytiCon Discovery (Potsdam, Germany) (Class I), 210 compounds from ASINEX (Winston-Salem, NC, USA) (Class II), 20 nucleic acid building blocks (Class III) as listed in the Supplementary Information (Supplementary Table 1), and 259 myxobacterial metabolites derived from the DZIF natural compound library (http://www.dzif.de/en/infrastructures/natural_compound_library/) (Class IV). Class I-III compounds possessing an average molecular weight (MW) of 313 Da were dissolved in DMSO to 5 or 10 mM stocks depending on their solubility. Class IV compounds possessing an average MW of 560 Da were dissolved in DMSO at a concentration of 1 mM. For hit characterization, the compounds NAT31-454537, NAT5-257461, and NAT11-275516 were repurchased as solids from AnalytiCon Discovery (Potsdam, Germany).

Biosensor experiments

Screening experiments were performed with a Reichert SR7000DC SPR biosensor, characterization of the functional binding of CsrA to RNA targets was performed with a Reichert SR7500DC SPR biosensor (Reichert Technologies, Depew, NY, USA). All SPR experiments were performed in SPR running buffer (10 mM Hepes, 150 mM NaCl, 5% [v/v] DMSO, 0.005% [v/v] Tween 20, pH 7.4) at 18 °C using SAD500 sensor chips from XanTec (XanTec bioanalytics, Düsseldorf, Germany).

Immobilization

Biotinylated His₆-CsrA was diluted in running buffer to a concentration of 6 µM and captured on a streptavidin surface at a constant flow rate of 10 µL/min at densities of 5,667-10,800 RU for the screening and initial dose response experiments and at densities of 5,225-6,697 RU for the characterization of the CsrA-RNA interactions. Remaining biotin binding sites on the streptavidin surface in the sample and reference cell were blocked with a 3 min injection of 3 µg/mL biotin. Control experiments with a CsrA-binding RNA sequence demonstrated that the protein was active and stable for at least 5 days under screening conditions.

SPR-based characterization of CsrA-RNA interactions

For RNA_A a twelve-point concentration series spanning 1.22 nM – 2.5 μ M and for RNA_B an eleven point concentration series in the range from 10 nM to 12.5 μ M was set up in SPR running buffer. Before starting the binding experiments, twelve warm-up blank injections were performed. RNA_A solutions were injected for 150 - 160 s and RNA_B solutions for 120 s, followed by a dissociation of 900 s at a constant flow rate of 35 μ L/min. Each solution of a given concentration was injected in duplicates. Data was processed and analyzed using Scrubber 2 software (version 2.0c, 2008, BioLogic Software). An eight-point DMSO calibration curve (4.25% - 6% DMSO in buffer solution) was generated in order to compensate mismatches in DMSO concentration between running buffer and samples. Buffer blank injections and DMSO calibration were included for standard double-referencing procedure to correct for baseline drifts and solvent effects. To determine binding affinities, binding responses at equilibrium were fit to a 1:1 steady state affinity model. The dissociation constant (K_D) is based on two independent experiments.

SPR-based screening

Compound stocks in DMSO were diluted in SPR running buffer to obtain 100 μ M samples with a final DMSO concentration of 5% [v/v]. The samples were injected for 120 s at a flow rate of 20 μ L/min followed by a dissociation time of 300 s. To monitor the binding capacity of the protein surface over the course of the screening and to ensure an accurate ranking of the screened compounds, a positive control was injected every 24th sample and the binding responses were normalized to the responses of the positive control (Supplementary Figure 1). Up to now, no small molecule inhibitor of the CsrA-RNA interaction has been available that could serve as a reference [20]. Therefore, a library component discovered in a pilot screen and showing reproducible and dose-dependent binding to CsrA was used as positive control. Periodical injections of running buffer containing 5% [v/v] DMSO were used as negative controls. To obtain high-quality data compounds were screened in duplicate and injected in random order in the replicate experiment. Data was processed and analyzed using Scrubber 2 software (version 2.0c, 2008, BioLogic Software). SPR signals were double-referenced and solvent corrected as described before. For the screening experiments the binding responses were normalized to MW. Small molecules with responses $\geq$ 50% compared to that of the positive control in both replicate experiments were defined as 'primary' hits. For hit confirmation, compounds were tested in dose-response binding experiments at concentrations of 62.5 μ M, 125 μ M and 250 μ M.

Development and validation of FP-based screening assay

In general, FP and fluorescence intensity (FI) were recorded in black 384-well microtiter plates (Greiner BioOne, catalog No. 781900) using CLARIOstar® microplate reader (BMG LABTECH, Ortenberg, Germany) with an excitation filter at 485 nm and emission filters at 520 nm. Focal height and gain adjustments were performed before each measurement to

achieve maximum sensitivity [28]. Generally, total FI values deviating more than 30% from those of the controls were excluded from analysis [29]. The respective FP values were reported in millipolarization units (mP). Typically, samples were tested in duplicate and experiments were performed twice independently, unless indicated otherwise. All experiments included a 60 min pre-incubation of CsrA in assay buffer at ambient temperature to mimic the inhibitor screening conditions.

Stock solutions of RNA oligomers were prepared by dissolving the lyophilized RNAs in FP assay buffer (10 mM HEPES, 150 mM NaCl, 0.005% Tween 20 [v/v], pH 7.4), divided into aliquots, frozen in liquid nitrogen, and stored at -20 °C.

To monitor direct binding of the different RNAs to CsrA or BSA, saturation experiments were conducted. Thereby, a constant concentration (15 nM final) of each fluorescein-labeled (*) RNA chosen according to ref. [30] was titrated with increasing concentrations of protein (0.15 nM - 10 µM final) in FP assay buffer. The half-maximal effective concentrations (EC₅₀s) were calculated with GraphPad Prism 5.00 software using nonlinear regression (four-parameter dose-response model). From the saturation experiments, a final CsrA concentration of 400 nM was chosen to be used in competition assays with RNA_B* according to ref. [31].

Competition experiments were performed at fixed concentrations of RNA_B* (15 nM final) and CsrA (400 nM final) by titration with increasing concentrations of unlabeled RNAs (0.92 nM - 120 µM final). The half-maximal inhibitory concentrations (IC₅₀s) were calculated with GraphPad Prism 5.00 software using nonlinear regression (four-parameter dose-response model).

DMSO tolerance was assessed by incubating various concentrations of DMSO (0 – 10% [v/v] final) and a fixed concentration of RNA_B* (15 nM final) either without CsrA or with a medium (150 nM final) or a high (400 nM final) concentration of CsrA (Supplementary Figure 2A). Accordingly, a final DMSO concentration of 5% [v/v] was chosen for the screening. During this experiment, complex stability was investigated additionally by performing measurements after different incubation times (every 30 min for 3 h). As the signal was stable after 1 h of incubation at ambient temperature, an incubation time of 1.5 h was considered suitable for the screening (Supplementary Figure 2B).

Assay robustness was demonstrated by assaying control run plates, i.e. a High Control (HC) plate comprising 382 samples with RNA* (15 nM final), CsrA (400 nM final), and DMSO (5% [v/v] final) in FP assay buffer and a Low Control (LC) plate containing the same samples without CsrA. Low coefficients of variation (CVs) of the controls and a high dynamic range reflected by a Z' factor of 0.9 demonstrated that the assay was robust and, thus, suitable for inhibitor screening (Supplementary Figures 2C and D).

FP-based screening

DMSO stock solutions of validated SPR screening hits out of Classes I-III as well as Class IV compounds were diluted in FP assay buffer to concentrations of 600 µM and 90 µM, respectively. Aliquots of 10 µL of these diluted solutions or of a 15% [v/v] solution of DMSO

in FP assay buffer as vehicle were distributed to 384-well microtiter plates using Eppendorf Xplorer® 12-channel pipette (Eppendorf, Hamburg, Germany). After addition of 10 µL of a 1.2 µM solution of CsrA in FP assay buffer, the samples were pre-incubated for 60 min at room temperature. Thereafter, the final incubation was started by the addition of 10 µL of a 45 nM solution of RNA_B* in FP assay buffer and lasted 1.5 h at room temperature. Accordingly, the final assay concentrations were: 15 nM of RNA_B*, 400 nM of CsrA, 200 µM for Class I-III compounds or 30 µM for Class IV compounds, 5% [v/v] DMSO. On each assay plate, 24 HCs, 24 LCs, and 24 blank samples (only buffer containing 5% [v/v] DMSO) were present to monitor assay performance during the screening. The compounds were tested in duplicate and experiments were performed twice (Class I-III compounds) or once (Class IV compounds) independently. Generally, compounds enhancing or quenching the total FI beyond the above-mentioned threshold were discarded. The percentage inhibition was calculated as follows: % inhibition = (mean of HCs – read value)/ (mean of HCs – mean of LCs) × 100% using the mean values of 24 HCs and LCs present on the respective assay plate.

FP-based dose-response and optical interference studies

The final screening hits were validated by dose-response and optical interference studies using the FP screening assay protocol. Therefore, two-fold serial dilutions of hit compounds were tested up to a concentration of 500 µM in the presence and absence of CsrA. Generally, compounds enhancing or quenching the total FI in a dose-dependent manner and exceeding the above-described threshold value were excluded from further evaluation. The experiments were performed twice independently using duplicates. IC₅₀ values were calculated as described above.

***In silico* modelling**

A model of the interaction between short RNA sequence GGA and CsrA was generated based on the reported X-ray structure of the CsrA ortholog RsmA (PDB 2BTI) [6] and the NMR structure of an RsmE:RNA complex containing a 20-nucleotide nucleic acid ligand (PDB 2JPP) [15] using YASARA structure with the YASARA2 force field [32]. RsmA and CsrA share 100% sequence identity, while RsmE has 81% homology to CsrA (Supplementary Figure 3). First, all solvent molecules were removed and the protein chains of both structures (2BTI chains A and B; 2JPP conformer 1 – chains A and B) were superimposed by use of the implemented MUSTANG algorithm [33]. Then, a three-step energy minimization procedure was performed where (i) all RNA atoms were held constant, (ii) only the RNA backbone atoms were fixed, and (iii) all constraints were removed. This allowed the protein atoms to adopt to the RNA environment and yielded a CsrA/RsmA-RNA(20 nucleotide oligomer) complex structure. All nucleotide atoms except the GGA-motif were deleted maintaining the correct 5' and 3' termini. Finally, a two-step energy minimization was conducted with (i) fixed RNA atoms and (ii) a completely unconstrained system. The generated CsrA/RsmA-GGA complex and the overlay with the original RNA coordinates (from 2JPP, conformer 1) were rendered with PovRay (<http://www.povray.org/>).

Results & discussion

Characterization of the functional binding of CsrA to different RNA targets by SPR and FP

For reliable establishment and application of the SPR- and FP-based screening assays, it was necessary to ensure that the protein retained its functional activity under assay conditions. Hence, we first analyzed the CsrA-RNA interaction in both assay formats.

For this purpose, CsrA from the enteropathogen *Y. pseudotuberculosis* was chosen as representative of the CsrA protein family. The RNA sequences RNA_A and RNA_B (Figure 2A) derived from the known RNA targets *hcnA* and *hcnA* Δ U (Figure 2A and ref. [20]), respectively, were selected as interaction partners. The latter have been well studied regarding their molecular interactions with and their binding affinity to the CsrA homolog RsmE from *P. fluorescens* [15,20]. Moreover, RNA_B was able to compete with *Y. pseudotuberculosis* *csrA* RNA, a homolog of the respective multivalent high-affinity RNA from *E. coli* [34], for CsrA binding. This was demonstrated in an RNA electromobility shift experiment, where RNA_B displaced one of the two CsrA-RNA complexes represented by the two band shifts in the gel (Supplementary Figure 4). To increase the dynamic range of the FP assay, we used shortened versions of *hcnA* and *hcnA* Δ U omitting the non-interacting stem nucleotides and retaining the conserved ANGGA(N) core binding motif [15,16,19,20].

For the SPR experiments, the protein was captured on streptavidin-coated chips via a site-specific C-terminally introduced biotin residue. Due to site-directed immobilization of CsrA, the RNA binding sites of the CsrA dimer should be freely accessible. In addition, this method should lead to a homogenous orientation of the ligand on the chip, thus enabling a better comparability of results obtained upon different protein immobilizations. To check the activity of the protein on the chip, RNA_A and RNA_B were tested for binding to immobilized CsrA. For both RNA sequences, we observed dose-dependent binding (Figure 2B and C). The responses at equilibrium were fitted to a 1:1 stoichiometry model representing the interaction of one CsrA dimer with two RNA molecules and yielding a K_D of 222 nM for RNA_A and 1.8 μ M for RNA_B (Figure 2A, B, and C). The obtained K_D values of CsrA with RNA_A and RNA_B were in good agreement with those reported for RsmE with *hcnA* (K_D of 85 nM in references [15] and [20]) and *hcnA* Δ U (K_D of 4.1 μ M in ref. [20]), respectively (Figure 2A). This was expected given the high overall homology of the two proteins of 81% and, especially, the 100% homology of the amino acids involved in RNA recognition (Supplementary Figure 3). Moreover, the SPR-derived binding constant for the CsrA-RNA_A complex was consistent with that determined by ITC (K_D of 180 nM) (Figure 2A and Supplementary Figure 5).

For reasons of comparability, biotinylated CsrA was used for the FP experiments as well. The RNA oligonucleotides were labeled with fluorescein at the freely accessible 3' position to be

used as FP probes. To investigate binding of the RNAs to CsrA, saturation plots were generated by plating increasing concentrations of CsrA with a fixed concentration of the respective probe. As expected from the SPR studies, RNA_A* and RNA_B* bound dose-dependently to CsrA with RNA_A* exhibiting a higher binding affinity (EC_{50} of 29 nM) than RNA_B* (EC_{50} of 303 nM) (Figure 2A and D). Moreover, no binding event was detected for RNA_B* when mutated in the essential GGA motif [19] ($GGA \rightarrow AGA$), referred to as RNA_C*. Furthermore, no binding was observed, when RNA_B* was incubated with the unrelated protein BSA (Figure 2D). These findings indicated specific binding of RNA_B* to CsrA. To investigate reversibility of binding, we conducted a competition assay using the respective non-labeled RNAs as competitors at varying concentrations (Figure 2E). Anticipating an identification of weak inhibitors during compound screening, we chose RNA_B* as probe for the competition assay according to ref. [35]. Indeed, the non-labeled oligomers RNA_A and RNA_B were able to compete off the probe in a dose-dependent manner with IC_{50} values of 147 nM and 955 nM, respectively (Figure 2A and E). In contrast, RNA_C only partially displaced the probe at high concentrations with an approximate IC_{50} of 90.7 μ M. The results from the competition assay in accordance with those from the saturation assay demonstrated a reversible and specific binding of the RNAs to CsrA.

Overall, the results generated by the three different methodologies SPR, FP, and ITC were in good agreement with each other as well as with those reported in the literature [20] (Figure 2A). This demonstrated compatibility and reliability of the established assays. Decisively, the results showed that the protein was active under assay conditions. Thus, the established assays were considered suitable for inhibitor screening and characterization.

Strategy A: Screening of small molecules

General screening procedure

For discovery of the first low-molecular-weight CsrA-RNA interaction inhibitors, we conducted a screening of ~ 1,000 small molecules from synthetic or natural sources. Thereby, hit selection was conducted stepwise using both, protein binding (primary screening) and inhibitory activity (secondary screening), as selection criteria (Figure 3). For reasons of selectivity and ease of drug optimization [36,37], the primary screening should preselect compounds binding dose-dependently to the CsrA protein. This screening was based on SPR technology due to its high sensitivity and particularly low protein consumption. The secondary screening should assess the functional activity of the SPR screening hits, i.e. their potential to inhibit the CsrA-RNA interaction. For efficient and reliable inhibitor screening, the respective competition assay was based on homogenous, rapid, and robust FP technology [38]. Due to low availability, the myxobacterial metabolites from the DZIF natural compound library were directly tested in the functional assay.

SPR-based primary screening

A total of 708 structurally diverse small molecules derived from different libraries (Class I-III) with an average MW of 313 Da were screened at 100 μ M for binding to CsrA using SPR spectroscopy. Relatively high-density surfaces ($\sim$ 6,000-10,000 RU) were used to enhance the sensitivity and enable the identification of even fragment-sized molecules present in the screening library. A CsrA-binding small molecule derived from the screening library was selected as positive control to be included in the screen. To ensure an accurate ranking of the screened compounds, the obtained responses were normalized to that of the positive control. In Figure 4A, the mean values of the normalized responses were plotted for each screened compound. The SPR screen resulted in 83 primary hits with a response $\geq$ 50% compared to that of the positive control in both replicate experiments (hit rate of 11.7%). To eliminate false positives the hit compounds were confirmed in dose-response studies (Figure 3). Out of the 83 primary hits 72 compounds bound to CsrA in a concentration-dependent manner, which was reflected in a hit rate of 10.2%.

FP-based secondary screening

In the secondary screen, the 72 confirmed SPR hits as well as the 259 myxobacterial metabolites were tested for inhibition of the CsrA-RNA interaction using the FP-based competition assay. The SPR hits were screened in duplicate and in two independent experiments at a final concentration of 200 μ M to detect even weak inhibitors. In each screening plate, one row each was used for HCs and LCs to survey data reliability during screening. Control data within and between experiments were reproducible with intra-run CVs $\leq$ 4%, an inter-run CV of 3%, and Z' factors $\geq$ 0.8 [39]. Additionally, 99% of the duplicate values for the compounds were $\leq$ 10% apart. Compounds showing $>$ 50% inhibition for all single values were defined as hits. For elimination of false hits due to interference with FI, compounds showing more or less than 30% of the total FI of the controls, a commonly used cutoff [29], were flagged as enhancers or quenchers. According to these criteria, 3 out of 72 compounds were identified as hits corresponding to a hit rate of 4.2% (Figures 3 and 4B). The high reproducibility of the FP assay as observed during screening of the SPR hits allowed a testing of the myxobacterial metabolites (Class IV compounds) in duplicate within only one experiment reducing consumption of the small-scale natural product library. Given a higher average MW of 560 Da, we chose a lower screening concentration of 30 μ M for these compounds. The hit criteria were fulfilled by 4 out of 259 compounds resulting in a hit rate of 1.5% (Figures 3 and 4B).

The final hit selection comprised 7 structurally diverse molecules with MWs ranging from $\sim$ 200 to $\sim$ 2,000 Da (Figure 5). For some of them, biological activities have been reported. The polyketide (+)-ambruticin S is a potent antifungal agent with oral activity in mice infection models [40]. MM2516 is a congener of the myxovalargin class of peptide antibiotics [41,42]. The tetrapeptide tubulysin Ar-672 belongs to the tubulysin class of antimitotic agents [43-46]. Interestingly, structurally related tubulysin derivatives included in the screening library were found to inhibit the CsrA-RNA interaction as well - although below the hit threshold (Supplementary Table 2 and Supplementary Figure 6). This prompted us to test further

congeners of this compound class, which, however, turned out to be inactive in the FP screening assay (Supplementary Table 2 and Supplementary Figure 6).

Hit characterization

For characterization of the 7 final screening hits, optical interference and dose-response studies were performed using the FP screening assay protocol. Thereby, serial dilutions of each compound were assayed up to a concentration of 500 μ M with and without addition of CsrA. None of the hit compounds systematically enhanced or quenched the FI of the probe neither in presence nor in absence of CsrA. All of them inhibited the CsrA-RNA interaction in a concentration-dependent manner (Figure 6A and B). Interestingly, all hit compounds shared an incomplete inhibition of the interaction (Figure 6A and B) reaching a maximal efficiency of $\sim 80\%$. A similar behavior has also been observed by Roehrl et al. [47] for protein-protein interaction inhibitors identified via FP screening. They explained this incomplete displacement behavior by an anticooperative four-state model [48] allowing the following mechanistic interpretations: compound and RNA bind to neighboring sites and reciprocally change the binding site geometry or they share partially overlapping binding sites restricting the access for the RNA to a subset of its natural binding interface [47]. Nevertheless, these models require experimental confirmation, for example via X-ray analysis. Importantly, the identification of multiple CsrA ligands with diverse molecular scaffolds provides promising starting points for the development of high affinity ligands via fragment linking [47,49].

Generally, IC_{50} values ranged from 4 to ~ 106 μ M (Figure 6C). The most potent inhibitor was the myxobacterial metabolite MM14 with an IC_{50} value of 4 μ M. Given that CsrA proteins can recognize RNA targets with a wide range of affinities *in vitro* including such with μ M or even mM affinity [16], MM14 bears the potential to displace such RNAs from CsrA in cell-based assays. The fragment-sized molecule NAT31-454537 (IC_{50} of 27 μ M) turned out to be the most efficient inhibitor with a ligand efficiency (LE) of 0.38 (Figure 6C). A high LE is generally considered as favorable for affinity optimization without dramatically increasing MW [50]. This is especially important given the low MW and $clogD_{pH\ 7.4}$ observed for compounds known to be active in Gram-negative bacteria [51,52]. Thus, NAT31-454537 represents a promising starting point for the development of high-affinity inhibitors to be used as tool compounds in cell-based assays. It is of great importance to have such tool compounds in hand, since it is not yet known, whether inhibition of the CsrA-RNA interaction by small molecules might result in attenuation of bacterial virulence. It is also worth mentioning that the CsrA-RNA interaction could be inhibited even with fragments such as NAT31-454537 (MW of 229 Da). This cannot be taken for granted given the large interaction interfaces lacking well-defined binding pockets - a scenario, which is usually encountered with macromolecule-macromolecule interactions [53].

Finally, SPR studies demonstrated that all 7 final hit compounds bound dose-dependently and reversibly to CsrA, whereas none of them interacted significantly with RNA_B (Supplementary Figures 7 and 8). This indicated that interaction inhibition occurred due to

protein binding and not due to interference with RNA, which was in accordance with the pursued screening strategy.

Strategy B: RNA ligand-based design approach

Complementary to the screening approach, we applied a rational, ligand-based strategy to discover inhibitors of the CsrA-RNA interaction. Thereby, the conserved ANGGA(N) core binding motif identified via *in vitro* selection [19] and NMR structural studies [15,16,20] served as starting point for the design of RNA-derived inhibitors. Within this sequence, the GGA motif was shown to be critical for CsrA binding [19]. Thus, for derivation of efficient inhibitors, we excised this motif as the ‘active core’ of the RNA ligands and analyzed its capability to inhibit the CsrA-RNA_B interaction using the FP-based competition assay. Interestingly, this short RNA (GGA_{RNA}) fully inhibited the interaction at the highest test concentration of 500 μ M (Figure 7A and B). This result prompted us to further restrict the binding motif to pairs of linked nucleotides, namely to GG and GA motifs. However, these RNA dimers (GG_{RNA} and GA_{RNA}) were inactive at the same concentration (Figure 7A). This was in accordance with the lack of activity observed for the single nucleobases adenine and guanine during FP-based secondary screening. These results confirmed the importance of the triple nucleotide GGA motif for CsrA affinity [19] even in the context of short RNA oligomers. Thus, we chose GGA as the core motif to be subjected to further modifications.

A general issue of RNA molecules is their hydrolytic instability, in which the 2' hydroxyl (2'OH) group of the RNA backbone plays a major role [54,55]. Consequently, we explored the tolerability of different backbones modified towards increased stability. The first modification, a complete elimination of the 2'OH groups as found in DNA backbones, resulted in reduced inhibition of the CsrA-RNA interaction (GGA_{DNA} in Figure 7A). To explain the results, the CsrA-GGA_{RNA} complex was modeled based on reported X-ray [6] and NMR [15] structures of RsmA and of RsmE in complex with RNA, respectively (Figure 7C). Accordingly, reduced inhibition might be explained by a loss of inter- and intramolecular H bonds. The latter might be important to stabilize the bioactive conformation. The second modification, a methylation of the 2'OH functions commonly used to impede enzymatic degradation by nucleases [54,56], strongly diminished the inhibitory activity of the GGA motif (GGA_{2'OMe-RNA} in Figure 7A). According to the structural model (Figure 7C), introduction of the methyl group also prevented inter- and intramolecular H bonds in which the 2'OH group of the RNA backbone is supposed to function as a donor. The third modification involved an exchange of the sugar-phosphate backbone against a pseudopeptide backbone consisting of *N*-(2-aminoethyl) glycine units, to which nucleobases were attached via a methyl carbonyl linker (Supplementary Figure 9). This so-called peptide nucleic acid (PNA) might not only have the advantage of increased hydrolytic stability but also of improved cell permeability due to lack of the negatively charged sugar-phosphate backbone [57]. Moreover, conjugation with cell-penetrating peptides has proven to be a successful strategy to import PNAs into the bacterial cell [58,59]. However, the respective GGA oligomer (GGA_{PNA}) did not inhibit the CsrA-RNA interaction at the test concentration (Figure 7A), which might be due to the drastic structural changes. Overall, the modified GGA motifs

did not retain the inhibitory activity of GGA_{RNA}. Finally, the GGA motif was extended to AGGGA, a pentamer reported to be an effective CsrA target in the context of a hairpin structure [20]. As expected, this enlargement resulted in an increased inhibitory activity in case of AGGGA_{DNA} and AGGGA_{2'OMe-RNA}, whereas AGGGA_{PNA} was inactive (Figure 7A). The activity increase might be due to additional interactions with the protein, e.g. electrostatic interactions caused by the extended sugar-phosphate backbone. However, from a medicinal chemist's point of view, the shorter GGA motifs represent more suitable starting points for development of efficient and cell-permeable inhibitors. Thus, we performed a dose-response study with the most active trimer, GGA_{RNA}, yielding an IC₅₀ value of 113 μ M (Figure 7B). This was a rather low activity compared to those achieved with most screening hits. Moreover, this inhibitor possessed a relatively high MW (985 Da) and, thus, a low LE (0.09). However, in contrast to all of the screening hits, GGA_{RNA} was able to fully inhibit the CsrA-RNA interaction (Figure 7B). This indicated that GGA_{RNA} might occupy a different binding site than the screening hits providing the basis for a possible linking strategy. Additionally, the binding site of GGA_{RNA} is known and a binding pose could be proposed (Figure 7C) that was in accordance with the obtained structure-activity data (Figure 7A). This might allow structure-based optimizations of the inhibitor. Alternatively, GGA_{RNA} could be used as tool compound in competition experiments to screen for inhibitors targeting the conserved CsrA-GGA interaction hotspot.

Conclusions

In this work, we described the discovery of CsrA-RNA interaction inhibitors by screening of small molecules and RNA ligand-based design. For this purpose, an SPR- and FP-based test system was established enabling their identification and characterization. Overall, the screening approach resulted in 7 structurally diverse hits capable of inhibiting the CsrA-RNA interaction in a dose-dependent manner. Inhibition was achieved by dose-dependent binding to CsrA and not by interfering with RNA, thus facilitating further optimization. None of the screening hits achieved an inhibition of more than ~ 80%. This behavior might hint at possible mechanisms of action such as allosteric modulation of RNA binding or preventing RNA from full interaction with CsrA. The two most promising hit candidates were MM14 with highest activity (IC₅₀ of 4 μ M) and NAT31-454537 with highest LE (0.38) suitable for evaluation in cell-based assays and further development to high-affinity inhibitors, respectively. The RNA ligand-based design strategy yielded short oligonucleotides derived from the conserved binding motif ANGGA(N) that disrupted the CsrA-RNA interaction. The GGA core motif with RNA backbone turned out to be the most promising ligand-derived inhibitor with an IC₅₀ of 113 μ M. Despite its low activity and high MW, this compound bears the advantages of complete interaction inhibition, known binding site, and the possibility of structure-based optimization.

Future perspective

In view of the growing development of antibiotic resistance among pathogens, the search for novel treatment options to combat infectious diseases has acquired major significance. Particularly, addressing new targets with novel modes of action is considered to be a promising strategy to overcome resistance. Given that CsrA is essential for full virulence of bacteria and is conserved among many pathogens, it represents a potential target for broad-spectrum anti-infective drug development. To date, no small molecule inhibitors targeting this protein have been available. Consequently, target validation has relied mainly on bacterial mutant studies not necessarily reflecting a target's suitability for pharmacological intervention [60]. In this work, the first step has been undertaken by identifying hit structures. These might serve as tool compounds to investigate the effects of CsrA-RNA inhibition in bacterial cell-based assays. Furthermore, they can be used as starting points for the development of high-affinity inhibitors, which could then be applied for proof-of-concept studies in *in vivo* models. Finally, they might be used to identify druggable target sites on CsrA, thus enabling virtual screening as an additional drug discovery option. Notably, the developed SPR- and FP-based assay systems represent sophisticated tools not only for inhibitor screening but also for characterization of CsrA-RNA interactions in contrast to time- and material-consuming ITC [15,16,20] or RNA electromobility shift assays [19].

Executive summary

Background:

- Novel anti-infectives are urgently needed due to the growing number of multidrug-resistant pathogens.
- CsrA regulates bacterial virulence and, therefore, represents an attractive, novel target for anti-infective drug development.
- Until now, no small molecule ligands of CsrA have been reported. Hence, screening of small molecules and RNA ligand-based drug design are valuable drug discovery tools for this target protein.

Establishment of a suitable test system enabling hit identification and characterization:

- Aiming at the discovery of small molecule inhibitors of the CsrA-RNA interaction a test system based on surface plasmon resonance and fluorescence polarization technologies has been established and validated in this work.

Discovery of hits by screening of small molecules and RNA ligand-based design:

- Screening of ~ 1,000 small molecules resulted in the identification of 7 structurally diverse hit compounds, which bound to CsrA in a concentration-dependent manner and partially inhibited the CsrA-RNA interaction *in vitro* with low μM activities.

- The RNA ligand-based design strategy led to the discovery of the short RNA oligonucleotide GGA achieving full inhibition of the CsrA-RNA interaction and possessing a known binding site facilitating structure-based optimization.

Conclusion:

- This work provided tool compounds and assay systems, which aid drug discovery efforts targeting the CsrA-RNA interaction.

References

1. Aertsen A, Michiels CW. Stress and how bacteria cope with death and survival. *Crit. Rev. Microbiol.* 30(4), 263-273 (2004).
2. Van Assche E, Van Puyvelde S, Vanderleyden J, Steenackers HP. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front. Microbiol.* 6, 141 (2015).
3. Heroven AK, Bohme K, Dersch P. The Csr/Rsm system of *Yersinia* and related pathogens: a post-transcriptional strategy for managing virulence. *RNA. Biol.* 9(4), 379-391 (2012).
4. Vakulskas CA, Potts AH, Babitzke P, Ahmer BM, Romeo T. Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol. Mol. Biol. Rev.* 79(2), 193-224 (2015).
5. White D, Hart ME, Romeo T. Phylogenetic distribution of the global regulatory gene *csrA* among eubacteria. *Gene.* 182(1-2), 221-223 (1996).
6. Heeb S, Kuehne SA, Bycroft M *et al.* Functional analysis of the post-transcriptional regulator RsmA reveals a novel RNA-binding site. *J. Mol. Biol.* 355(5), 1026-1036 (2006).
7. Romeo T. Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol. Microbiol.* 29(6), 1321-1330 (1998).
8. Babitzke P, Romeo T. CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.* 10(2), 156-163 (2007).
9. Baker CS, Morozov I, Suzuki K, Romeo T, Babitzke P. CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli*. *Mol. Microbiol.* 44(6), 1599-1610 (2002).
10. Dubey AK, Baker CS, Suzuki K *et al.* CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *J. Bacteriol.* 185(15), 4450-4460 (2003).
11. Wei BL, Brun-Zinkernagel AM, Simecka JW, Pruss BM, Babitzke P, Romeo T. Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol. Microbiol.* 40(1), 245-256 (2001).
12. Patterson-Fortin LM, Vakulskas CA, Yakhnin H, Babitzke P, Romeo T. Dual posttranscriptional regulation via a cofactor-responsive mRNA leader. *J. Mol. Biol.* 425(19), 3662-3677 (2013).
13. Yakhnin AV, Baker CS, Vakulskas CA *et al.* CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage. *Mol. Microbiol.* 87(4), 851-866 (2013).
14. Gutierrez P, Li Y, Osborne MJ, Pomerantseva E, Liu Q, Gehring K. Solution structure of the carbon storage regulator protein CsrA from *Escherichia coli*. *J. Bacteriol.* 187(10), 3496-3501 (2005).
15. Schubert M, Lapouge K, Duss O *et al.* Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nat. Struct. Mol. Biol.* 14(9), 807-813 (2007).
16. Duss O, Michel E, dit Konté ND, Schubert M, Allain FH. Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition. *Nucleic Acids Res.* 42(8), 5332-5346 (2014).

17. Rife C, Schwarzenbacher R, McMullan D *et al.* Crystal structure of the global regulatory protein CsrA from *Pseudomonas putida* at 2.05 Å resolution reveals a new fold. *Proteins*. 61(2), 449-453 (2005).
18. Marden JN, Diaz MR, Walton WG *et al.* An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* 110(37), 15055-15060 (2013).
19. Dubey AK, Baker CS, Romeo T, Babitzke P. RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA*. 11(10), 1579-1587 (2005).
20. Lapouge K, Perozzo R, Iwaszkiewicz J *et al.* RNA pentaloop structures as effective targets of regulators belonging to the RsmA/CsrA protein family. *RNA Biol.* 10(6), 1031-1041 (2013).
21. Bhatt S, Edwards AN, Nguyen HT, Merlin D, Romeo T, Kalman D. The RNA binding protein CsrA is a pleiotropic regulator of the locus of enterocyte effacement pathogenicity island of enteropathogenic *Escherichia coli*. *Infect. Immun.* 77(9), 3552-3568 (2009).
22. Heroven AK, Bohme K, Rohde M, Dersch P. A Csr-type regulatory system, including small non-coding RNAs, regulates the global virulence regulator RovA of *Yersinia pseudotuberculosis* through RovM. *Mol. Microbiol.* 68(5), 1179-1195 (2008).
23. Barnard FM, Loughlin MF, Fainberg HP *et al.* Global regulation of virulence and the stress response by CsrA in the highly adapted human gastric pathogen *Helicobacter pylori*. *Mol. Microbiol.* 51(1), 15-32 (2004).
24. Altier C, Suyemoto M, Lawhon SD. Regulation of *Salmonella enterica* serovar typhimurium invasion genes by *csrA*. *Infect. Immun.* 68(12), 6790-6797 (2000).
25. Heurlier K, Williams F, Heeb S *et al.* Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 186(10), 2936-2945 (2004).
26. Mulcahy H, O'Callaghan J, O'Grady EP *et al.* *Pseudomonas aeruginosa* RsmA plays an important role during murine infection by influencing colonization, virulence, persistence, and pulmonary inflammation. *Infect. Immun.* 76(2), 632-638 (2008).
27. Alfrén J, Ottow KE, Hobley TJ. *In vivo* biotinylation of recombinant beta-glucosidase enables simultaneous purification and immobilization on streptavidin coated magnetic particles. *J. Mol. Catal. B: Enzym.* 94, 29-35 (2013).
28. Rossi AM, Taylor CW. Analysis of protein-ligand interactions by fluorescence polarization. *Nat. Protoc.* 6(3), 365-387 (2011).
29. Cho EJ, Xia S, Ma LC *et al.* Identification of influenza virus inhibitors targeting NS1A utilizing fluorescence polarization-based high-throughput assay. *J. Biomol. Screen.* 17(4), 448-459 (2012).
30. Arkin MR, Glicksman MA, Fu H, Havel JJ, Du Y. Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats. In: *Assay Guidance Manual*. Sittampalam GS, Coussens NP, Nelson H *et al.* (Eds.). Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), USA, (2012).
31. Huang X. Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *J. Biomol. Screen.* 8(1), 34-38 (2003).

32. Krieger E, Joo K, Lee J *et al.* Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*. 77(Suppl 9), 114-122 (2009).
33. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM. MUSTANG: a multiple structural alignment algorithm. *Proteins*. 64(3), 559-574 (2006).
34. Yakhnin H, Yakhnin AV, Baker CS *et al.* Complex regulation of the global regulatory gene *csrA*: CsrA-mediated translational repression, transcription from five promoters by σ^{70} and σ^S , and indirect transcriptional activation by CsrA. *Mol. Microbiol.* 81(3), 689-704 (2011).
35. Visco C, Perrera C, Thieffine S, Sirtori FR, D'Alessio R, Magnaghi P. Development of biochemical assays for the identification of eIF4E-specific inhibitors. *J. Biomol. Screen.* 17(5), 581-592 (2012).
36. Kumar A, Park H, Fang P *et al.* Myotonic dystrophy type 1 RNA crystal structures reveal heterogeneous 1 x 1 nucleotide UU internal loop conformations. *Biochemistry*. 50(45), 9928-9935 (2011).
37. Aboul-ela F. Strategies for the design of RNA-binding small molecules. *Future Med. Chem.* 2(1), 93-119 (2010).
38. Lea WA, Simeonov A. Fluorescence polarization assays in small molecule screening. *Expert Opin. Drug Discov.* 6(1), 17-32 (2011).
39. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* 4(2), 67-73 (1999).
40. Hanessian S, Focken T, Mi X *et al.* Total synthesis of (+)-ambruticin S: probing the pharmacophoric subunit. *J. Org. Chem.* 75(16), 5601-5618 (2010).
41. Irschik H, Gerth K, Kemmer T, Steinmetz H, Reichenbach H. The myxovalarginins, new peptide antibiotics from *Myxococcus fulvus* (Myxobacterales). I. Cultivation, isolation, and some chemical and biological properties. *J. Antibiot. (Tokyo)* 36(1), 6-12 (1983).
42. Irschik H, Reichenbach H. The mechanism of action of myxovalargin A, a peptide antibiotic from *Myxococcus fulvus*. *J. Antibiot. (Tokyo)* 38(9), 1237-1245 (1985).
43. Murray BC, Peterson MT, Fecik RA. Chemistry and biology of tubulysins: antimitotic tetrapeptides with activity against drug resistant cancers. *Nat. Prod. Rep.* 32(5), 654-662 (2015).
44. Sasse F, Steinmetz H, Heil J, Hoefle G, Reichenbach H. Tubulysins, new cytostatic peptides from myxobacteria acting on microtubuli. Production, isolation, physico-chemical and biological properties. *J. Antibiot. (Tokyo)* 53(9), 879-885 (2000).
45. Steinmetz H, Glaser N, Herdtweck E, Sasse F, Reichenbach H, Hoefle G. Isolation, crystal and solution structure determination, and biosynthesis of tubulysins - powerful inhibitors of tubulin polymerization from myxobacteria. *Angew. Chem. Int. Ed. Engl.* 43(37), 4888-4892 (2004).
46. Helmholtz Centre for Infection Research, Braunschweig, Germany, EP-1562979-B1 (2005).
47. Roehrl MH, Wang JY, Wagner G. Discovery of small-molecule inhibitors of the NFAT-calcineurin interaction by competitive high-throughput fluorescence polarization screening. *Biochemistry*. 43(51), 16067-16075 (2004).

48. Roehrl MH, Wang JY, Wagner G. A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry*. 43(51), 16056-16066 (2004).
49. Rees DC, Congreve M, Murray CW, Carr R. Fragment-based lead discovery. *Nat. Rev. Drug Discov.* 3(8), 660-672 (2004).
50. Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* 9(10), 430-431 (2004).
51. O'Shea R, Moser HE. Physicochemical properties of antibacterial compounds: implications for drug discovery. *J. Med. Chem.* 51(10), 2871-2878 (2008).
52. Storz MP, Allegretta G, Kirsch B, Empting M, Hartmann RW. From *in vitro* to *in cellulo*: structure-activity relationship of (2-nitrophenyl)methanol derivatives as inhibitors of PqsD in *Pseudomonas aeruginosa*. *Org. Biomol. Chem.* 12(32), 6094-6104 (2014).
53. Arkin MR, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* 3(4), 301-317 (2004).
54. Egli M, Gryaznov SM. Synthetic oligonucleotides as RNA mimetics: 2'-modified RNAs and N3'->P5' phosphoramidates. *Cell Mol. Life Sci.* 57(10), 1440-1456 (2000).
55. Dingermann T, Zündorf I. RNA als Arzneistoff - Faszinierendes Konzept, aber schwer zu realisieren. *Pharmazeutische Zeitung online* 15 (2010). <http://www.pharmazeutische-zeitung.de/index.php?id=33369>
56. Kurreck J. The Role of Backbone Modifications in Oligonucleotide-Based Strategies. In: *Therapeutic Oligonucleotides*. Kurreck J (Ed.). The Royal Society of Chemistry, Cambridge, UK, 1-22 (2008)
57. Paulasova P, Pellestor F. The peptide nucleic acids (PNAs): a new generation of probes for genetic and cytogenetic analyses. *Ann. Genet.* 47(4), 349-358 (2004).
58. Ghosal A, Nielsen PE. Potent antibacterial antisense peptide-peptide nucleic acid conjugates against *Pseudomonas aeruginosa*. *Nucleic Acid Ther.* 22(5), 323-334 (2012).
59. Good L, Nielsen PE. Antisense inhibition of gene expression in bacteria by PNA targeted to mRNA. *Nat. Biotechnol.* 16(4), 355-358 (1998).
60. Sams-Dodd F. Target-based drug discovery: is something wrong? *Drug Discov. Today* 10(2), 139-147 (2005).

Reference annotations

** Ref. [3]: Comprehensive description of the Csr system in *Y. pseudotuberculosis* in comparison to other bacterial species

** Ref. [15]: Determination of the NMR solution structure of the CsrA homolog RsmE in complex with a 20-nucleotide RNA sequence – molecular basis for RNA recognition

* Ref. [16]: Explanation of the wide range of affinities occurring in CsrA/RsmA-RNA recognition on a molecular level

* Ref. [19]: Identification of RNA consensus sequence for high-affinity CsrA binding via SELEX experiments

* Ref. [20]: Identification of RNA pentaloop structure as effective targets of CsrA/RsmA proteins

** Ref. [4]: Review on the role of the Csr system in bacterial virulence networks of different pathogens

** Ref. [48]: Provision of a general framework for development and data analysis of FP-based competition assays for small molecule screening

Acknowledgments

The authors would like to thank Jeannine Jung for technical assistance with inhibitor screening, Benjamin Kirsch for support regarding vector cloning, Andrea Braunshausen for advice, Michael Hoffmann for LC-MS determination of CsrA mass, Ann Kathrin Heroven for construction of plasmid pAKH172, Ullrich Scheid for re-isolation of compound MM2516, Wolfgang Kessler and Stefan Bernecker for re-isolation of compounds M14 and (+)-ambruticin S, and Carsten Börger and Lorenz Siebenbürger for their help with performing SPR experiments on Biacore instrument.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending or royalties. No writing assistance was utilized in the production of this manuscript.

Figures and Legends to Figures

Figure 1. Molecular function of CsrA and its interruption by small molecule inhibitors. The protein CsrA binds to the ribosome binding site (RBS) of target mRNAs repressing translation by competing with the ribosome for RNA binding. That way, CsrA modulates expression of multiple virulence-relevant processes. Small molecule inhibitors prevent the interaction of CsrA with the mRNA leading to a dysregulation of virulence.

A

Code	Sequence (5'-3')	K_D SPR	EC_{50} FP Sat.	IC_{50} FP Comp.	K_D ITC
<i>hcnA</i>	GGGCUUCACGGAUGAAGCCC	-	-	-	85 ± 35 nM [#]
RNA_A	UCACGGAUGA	222 ± 28 nM	29 ± 2 nM	147 ± 8 nM	180 ± 5 nM
<i>hcnA</i> ΔU	GGGCUUCACGGA_GAAGCCC	-	-	-	4.1 ± 2.5 μM [#]
RNA_B	UUCACGGA_GAA	1.8 ± 0.1 μM	303 ± 19 nM	955 ± 3 nM	-
RNA_C	UUCACAGA_GAA	-	n.d.	$\sim 90.7 \pm 8.1$ μM*	-

B

C

D

E

Figure 2. Characterization of the functional binding of CsrA to different RNA targets by SPR, FP, and ITC. (A) RNA sequences and binding data of different CsrA-RNA complexes determined in SPR, FP, and ITC experiments. Data are given as mean values of two independent experiments $\pm$ standard deviation. #: K_D values reported in ref. [20] using the CsrA homolog RsmE from *P. fluorescens*. *: IC_{50} extrapolated with the bottom plateau fixed to that of the low control (34 mP). (B+C) For SPR studies, varying concentrations of RNA_A (B) and RNA_B (C), respectively, were injected in duplicates over the sensor surface. Both, RNA_A and RNA_B, bound to CsrA in a dose-dependent manner. Equilibrium response data were fit to a 1:1 binding model to determine equilibrium dissociation constants. Representative sensorgrams of one independent experiment for both RNA targets are depicted. (D) Fluorescence polarization (FP)-based saturation experiments were performed by titrating 15 nM of fluorescein-labeled (*) RNAs with a dilution series of target protein CsrA or unrelated protein BSA. Representative curves of one independent experiment are shown with data points representing averaged FP values of duplicates $\pm$ standard deviation. EC_{50} values were determined using non-linear regression (four-parameter dose-response model). (E) FP-based competition experiments were conducted using 15 nM of RNA_B* as probe, 400 nM of CsrA, and varying concentrations of non-labeled RNAs as competitors. Representative curves of one independent experiment are given with data points representing averaged FP values of duplicates $\pm$ standard deviation. IC_{50} values were determined using non-linear regression (four-parameter dose-response model). Abbreviations: K_D , equilibrium dissociation constant; SPR, surface plasmon resonance; EC_{50} , half maximal excitatory concentration; FP, fluorescence polarization; Sat., saturation experiment; IC_{50} ,

half maximal inhibitory concentration; Comp., competition experiment; ITC, isothermal titration calorimetry; n.d., not determinable. RU, response unit; mP, millipolarization; BSA, bovine serum albumin.

Figure 3. Flowchart representing the screening procedure. The SPR-based primary screening should select compounds binding to CsrA, the FP-based secondary screening compounds inhibiting the CsrA-RNA interaction.

Figure 4. Scatter plots of SPR-based primary screening (A) and FP-based secondary screening (B). (A) The scatter plot illustrates the mean values of SPR responses normalized to the response of the positive control (y-axis) for each small molecule screened at 100 μ M (x-axis). The dashed line represents the threshold level. (B) The dots of the scatter plot represent averaged data of screening compounds. The dashed line indicates the hit threshold.

Figure 5. Structures of final screening hits.

Figure 6. Characterization of final screening hits. Dose-response studies with Class I (**A**) and Class IV (**B**) hit compounds performed using the FP-based competition assay. Representative curves of one independent experiment are depicted with data points representing averaged FP values of duplicates $\pm$ standard deviation. Curves were fit to a four-parameter dose-response model to calculate IC₅₀ values. For MM2516, the bottom plateau was fixed to the respective mean value of the screening hits (53 mP). (**C**) IC₅₀ values, LEs, MW, and clogD_{pH 7.4} values of hit compounds. IC₅₀ values are given as mean values of two independent experiments $\pm$ the standard deviation. LE was calculated according to the equation $LE = -1.4 \log[\text{IC}_{50} (\text{mol/L})]/N_{\text{non-hydrogen atoms}}$. The clogD_{pH 7.4} values were calculated with ACD/Percepta 14.0.0 software. Abbreviations: mP, millipolarization; IC₅₀, half maximal inhibitory concentration; LE, ligand efficiency; MW, molecular weight, clogD_{pH 7.4}, the calculated logarithm of the pH-dependent octanol-water partition coefficient at pH 7.4; FP, fluorescence polarization.

A

Code	Sequence (5'-3')	Backbone	% Inhibition
GG _{RNA}	GG	RNA	n.i. [#]
GA _{RNA}	GA	RNA	n.i. [#]
GGA _{RNA}	GGA	RNA	96 ± 1
GGA _{DNA}	GGA	DNA	42 [*]
GGA _{2'-OMe-RNA}	GGA	2'-methoxy-RNA	27 ± 6
GGA _{PNA}	GGA	PNA	n.i. [#]
AGGGA _{DNA}	AGGGA	DNA	74 ^{*‡}
AGGGA _{2'-OMe-RNA}	AGGGA	2'-methoxy-RNA	56 ± 6
AGGGA _{PNA}	AGGGA	PNA	n.i. [#]

B

C

Figure 7. RNA ligand-based design approach. (A) Sequence, backbone and percentage inhibition of oligonucleotides derived from core binding motif ANGGA(N). Compounds were tested at 500 μM in the FP-based competition assay. Data represent mean values of two independent experiments $\pm$ the standard deviation. #: n.i. = no inhibition (i.e., $\leq 20\%$ inhibition). *: values obtained from one experiment. ‡: tested at 333 μM . Abbreviation: PNA, peptide nucleic acid. (B) Dose-response studies of GGA_{RNA} were conducted using the FP-base competition assay. The dose-response curves were fit to a four-parameter dose-response model to determine IC_{50} . The depicted curve consists of averaged data points from two experiments based on single values. Data

points are presented $\pm$ standard deviation. **(C)** Proposed binding mode of isolated GGA_{RNA} (white) modeled using reported X-ray [6] and NMR [15] structures of CsrA homologs RsmA from *Y. enterocolitica* and RsmE from *P. fluorescens* in complex with RNA. The two chains of the homodimer are shown as white and green surfaces, the RNA molecules binding to them as black and green sticks. The RNA-trimer (GGA) is colored as follows: carbon white, nitrogen blue, oxygen red, phosphorus orange, and hydrogens were omitted for clarity. Hydrogen bonds are shown as blue dashed lines, while those arising from the 2'OH group of the RNA backbone are highlighted in yellow.