480–494 *Nucleic Acids Research, 2019, Vol. 47, No. 1* *Published online 12 November 2018*
doi: 10.1093/nar/gky1111

Deducing putative ancestral forms of GNRA/receptor interactions from the ribosome

Erin R. Calkins, Paul Zakrevsky, Vasken L. Keleshian, Eduardo G. Aguilar, Cody Geary and Luc Jaeger*

Department of Chemistry and Biochemistry, Biomolecular Science and Engineering Program, University of California, Santa Barbara, CA 93106–9510, USA

Received July 27, 2018; Revised October 17, 2018; Editorial Decision October 20, 2018; Accepted October 22, 2018

ABSTRACT

Stable RNAs rely on a vast repertoire of long-range interactions to assist in the folding of complex cellular machineries such as the ribosome. The universally conserved L39/H89 interaction is a long-range GNRA-like/receptor interaction localized in proximity to the peptidyl transferase center of the large subunit of the ribosome. Because of its central location, L39/H89 likely originated at an early evolutionary stage of the ribosome and played a significant role in its early function. However, L39/H89 self-assembly is impaired outside the ribosomal context. Herein, we demonstrate that structural modularity principles can be used to re-engineer L39/H89 to self-assemble in vitro. The new versions of L39/H89 improve affinity and loop selectivity by several orders of magnitude and retain the structural and functional features of their natural counterparts. These versions of L39/H89 are proposed to be ancestral forms of L39/H89 that were capable of assembling and folding independently from proteins and post-transcriptional modifications. This work demonstrates that novel RNA modules can be rationally designed by taking advantage of the modular syntax of RNA. It offers the prospect of creating new biochemical models of the ancestral ribosome and increases the tool kit for RNA nanotechnology and synthetic biology.

INTRODUCTION

Ribosomal RNAs, ribozymes and riboswitches take advantage of extended networks of non-covalent tertiary contacts to stabilize their native conformation (1–6). Rather than being idiosyncratic, these contacts are often part of recurrent RNA structural modules (or motifs), which are well-defined three-dimensional (3D) conformers specified by particular sets of nucleotides (e.g. (7–10)). One of these modules is the 'double-locked bulge' (2bp_2x_bulge) module, which forms the binding site of protein S8 in bacterial 16S rRNAs (8,11,12). Interestingly, the same structural module is also found in helix H89 of the large subunit rRNAs from all kingdoms of life, where it forms a universally conserved loop receptor that binds to the universally conserved terminal loop L39 (L39 pentaloop in Eukaryotes and Archaea and L39 hexaloop in Bacteria) (8,13–15) (Figure 1). This long-range interaction participates in the assembly process of the large subunit rRNA by docking domain 2 to domain 5 and likely stabilizes the peptidyl transferase center (PTC) of the ribosome (16–18). The ribosomal L39/H89 interaction has remarkable structural similarities with GNRA tetraloop/receptor interactions, one of the most widespread classes of long-range interactions in structured RNAs (19–25) (Figure 1). However, preliminary investigations of L39/H89 in isolation have indicated that it is thermodynamically metastable and unlikely to promote efficient self-assembly outside its natural structural context (8). We hypothesized that ancestral forms of the ribosome likely relied on more stable RNA/RNA interactions for folding and assembling into their native structures. Indeed, the L39/H89 interaction might have evolved through time to become dependent on additional factors like proteins and/or post-transcriptional modifications within the context of the ribosome (8,13,26). Therefore, an outstanding question is whether stable and more ancestral versions of this universally conserved ribosomal interaction can be recovered by taking advantage of the present structural knowledge of RNA, and more specifically, GNRA/receptor interaction.

*To whom correspondence should be addressed. Tel: +1 805 893 3628; Fax: +1 805 893 4120; Email: jaeger@chem.ucsb.edu
Present addresses:
Erin R. Calkins, Department of Biological Sciences, Moorpark College, 7075 Campus road, Moorpark, CA 93021, USA.
Paul Zakrevsky, RNA Biology Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA.
Vasken L. Keleshian, Mayo Clinic, Baldwin Building 4th Floor, 221 Fourth Avenue SW. Rochester, MN 55905, USA.
Eduardo Aguilar, Laboratory of Virology and Infectious Disease, 230 York Ave, Box 64, New York, NY 10065, USA.
Cody Geary, Bioengineering, Computer Science, and Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125, USA.

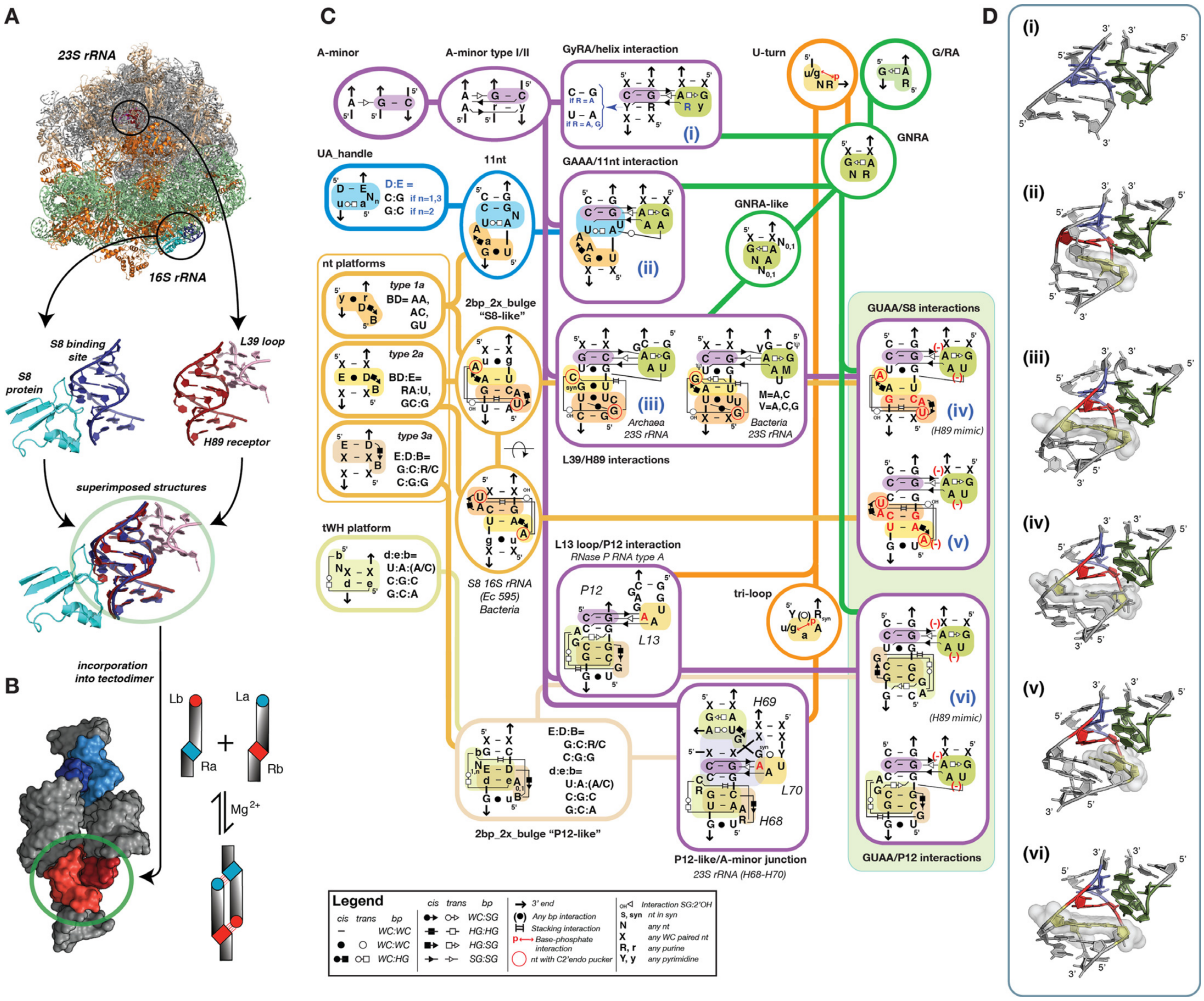

Figure 1. Principles of RNA architectonics and structural syntax for natural and rationally designed GUAA/receptor interactions. (**A**) Through the RNA architectonics approach (39,42), RNA modules such as the L39/H89 (pink/red) and S8 protein/S8 binding site (turquoise/blue) interactions are extracted from the X-ray atomic structure of the ribosome (e.g. *Thermus thermophilus*) and compared for their structural similarities. (**B**) These interactions can be incorporated into the tectoRNA dimer system (PDB ID: 2JYJ) (34) for studying their self-assembly properties outside their natural ribosomal context. Two tectoRNA units, each containing a loop and a receptor, assemble in the presence of Mg^{2+} through loop-receptor interactions (24,30,31). (**C**) Structural syntax of RNA modules of increasing complexity that enters into the composition of natural loop/receptor interactions (e.g.: i-iii) and that can be used for rationally designing new GNRA/receptor interactions (iv-vi). The nomenclature for base pairs and tertiary contacts is according to references (7,8,82). (**D**) 3D structures of natural and engineered GNRA tetraloop/receptor interactions: (i) GYRA/helix interaction (from [pdb_ID: 1s72]); (ii) GAAA/11nt-receptor interaction from [pdb_ID: 1hr2]; (iii) GUAAG/H89-receptor interaction from [pdb_ID: 1s72]; (iv) *type 2a* GUAA/S8-receptor interaction (S8 from [pdb_ID: 2avy]); (v) *type 1a* GUAA/S8-receptor interaction (S8 from [pdb_ID: 2avy]); (vi) *type 3a* GNRA/P12-receptor interaction (P12 from [pdb_ID: 1us9]).

GNRA/receptor interactions typically involve a GNRA terminal loop (with R standing for purine, N for any nucleotide and the first guanine forming a trans shallow-groove/Hoogsteen base pair with the fourth adenine) that forms an 'A-minor' interaction (22) between the two last positions of the tetraloop and two base-pairs (usually two G:C bps) within a helix (19,27) (Figure 1). Moreover, additional non-canonical base pairs and/or stacking interactions often contribute to an increase of stability and loop specificity for the GNRA/receptor interaction (24,28–36). For instance, in addition to the 'A-minor' sub-module, the GAAA/11nt interaction includes a nucleotide (nt) platform and UA_handle sub-module that are responsible for the stacking and specific recognition of the second loop position (24,29,37). Similarly, the GGAA/R1 interaction likely involves a nt platform and forms additional specific contacts with its GGAA tetraloop (24,38). Using the architectonics approach (39–42), we previously developed a tectoRNA dimer system that consists of two units able to self-assemble through two loop/receptor interactions in a synergistic manner (Figure 1B) (30,31,33,34). This system proved extremely useful for characterizing the behaviors of various GNRA loop-receptor interactions (24,30,31,33–35,43,44).

Herein, this tectoRNA system was used as a scaffold to retro-synthesize potential ancestral versions of the L39/H89 ribosomal interaction (Figure 1). Although important RNA structures often display a high degree of sequence conservation, it has actually been found that structure is more conserved in the long run (45,46). Therefore, considering that the L39 pentaloop from Eukaryotes and

Archaea folds as a GUAA-like tetraloop and that R89 adopts the same structural module as the S8 RNA binding site (S8) (Figure 1A), we hypothesized that GUAA-tetraloop/S8-like receptor interactions could be derived from the natural L39/H89 interaction, and promote self-assembly within the context of the tectoRNA dimer. Moreover, by using a structural syntax network that illustrates how small submodules can be combined to generate larger modular loop/receptor interactions, we explored the assembly properties of rationally designed GNRA/receptor interactions with novel submodular combinations not yet identified in nature (Figure 1). Besides demonstrating that retro-synthesized GUAA/receptor interactions can potentially mimic ancestral forms of the universally conserved L39/H89 interaction from the large RNA subunit of the ribosome (47), our study also expands the toolkit of RNA structural building blocks for RNA nanotechnology and RNA synthetic biology applications (e.g. (6,36,39–41,48–50)).

MATERIALS AND METHODS

Rational structural design of new GNRA/receptor interactions and tectoRNA synthesis

The bimolecular tectoRNA system used in this study assembles through two GNRA loop/receptor interactions (24,30) according to the atomic model structures available (Figure 1 and Supplementary Figure S1) (31,34). Using the RNA architectonics guidelines (39,42), 3D structural models for the novel GNRA/receptor interactions described herein were built from RNA modules present in X-ray atomic structures [PDB IDs: 1u9s, 1s72, 2avy] with SwissPdbviewer (51). ModeRNA was used to generate specific, single nucleobase replacements (52) and image rendering was performed in PyMOL (53). Proper folding of the secondary structure of each RNA construct was checked using Unafold (54).

RNA molecules (Supplementary Tables S1 and 2) were transcribed from polymerase chain reaction generated templates with T7 RNA polymerase, purified by denaturing polyacrylamide gel electrophoresis (PAGE) and 3′-[^{32}P]pCp labeled as previously described (30,31,33) (see Supplementary Data). Antisense and primer DNA sequences were ordered from IDT.

TectoRNA self-assembly and determination of equilibrium constants of dissociation (K_d) and receptor specificity

K_d values were derived from the titration experiments performed at 10°C as previously described (9,24,44) (Supplementary Table S3). In brief, self-assembly samples were prepared by mixing equimolar concentrations of receptor (R) and probe (P) tectoRNAs (ranging from 1 nM–20 μM) in water, including 1nM of 3′-[^{32}P]pCp-labeled probe tectoRNA (24). After denaturation (2 min, 95°C; 3 min, 4°C; 5 min, 30°C), RNA samples were incubated 20 min at 30°C in association buffer (15 mM $Mg(OAc)_2$ and 89 mM Tris-borate, pH 8.3) followed by equilibration for 30 min at 10°C. TectoRNA assembly was monitored by native 7% (29:1) polyacrylamide gels at 10°C in association buffer. Monomers (Receptor [R]) and heterodimers (Receptor [R]xProbe [P]) were quantified using ImageQuant software (24,30) (Supplementary Figure S1C). A non-linear fit from the experimental equation: $f = [2\beta M_0 + K_d - (4M_0\beta K_d + K_d^2)^{0.5}]/2M_0$, was used to determine K_d's for the equilibrium reaction $P + R \leftrightarrow PR$, where f is the fraction of RNA heterodimer, defined as the weight-in-weight (w/w) ratio of the dimer (PR) to the total RNA species (P + R+PR), β is the maximum fraction of RNA able to dimerize, and M_0 is the total concentration of the probe (55). With β equal to 1 for most molecules tested, the equation simplifies to $K_d = [(M_0)(1-f)^2]/f$ so that $M_0/2$ is the value at which 50% of heterodimer is formed (24,31). For each set of molecules, K_d values typically correspond to the average calculated from at least two independent experiments. For those values that were calculated from three or more independent experiments the standard error deviation is usually within 30% but can vary from one sample to the other (Supplementary Table S3). K_d measurements below 4000 nM have variations usually within 15% or less whereas those above 4000 nM can be less precise, especially at concentration closer to the upper limit of detection (20 μM). The corresponding free energy of dimerization (ΔG) between tectoRNA receptors and RNA tetraloop probes are determined from the equation, $\Delta G = RTlnK_d$, where R is the gas constant (1.985 cal/K/mol) and T is the temperature (283°K). The apparent free energy variation of dimerization at 10°C (ΔΔG) can be derived from the equation, ΔΔG = ΔG(PR)-ΔG(reference) (Supplementary Table S4). In this study, 4000 nM was arbitrarily chosen as the reference threshold. It corresponds to the nearest 'round' K_d value below which all the tectoRNA helical receptors assemble with their cognate GURA tetraloops. GNRA specificity profiles (utilizing color coded bars) were created for each receptor tested to more easily identify binding patterns (Supplementary Figure S1D and E). At the exception of GAAA, all the other GNRA probes were tested for their affinity to each receptor (Supplementary Tables S3 and 4). An UUCG probe was used as a negative control for indicating the position of the monomer. The UUCG probe was also tested with some of the receptors investigated in this study. It did not assemble at RNA concentrations up to 20 μM (data not shown).

Chemical probing with dimethylsulfate (DMS)

The S8 tectoRNA monomer (2 μM), with an elongated 20 nt 5' tail for primer extension (S8_tail, Supplementary Table S2B), was assembled with or without GUAA probe (2 μM) (dimer contains equimolar amounts of GUAA tetraloop and receptor) as described above in presence of 50 mM HEPES, 100 mM KCl, 15 mM $Mg(OAc)_2$ final. Following incubation, modified RNA samples were directly treated with a solution of dimethylsulfate (DMS) diluted in 100% EtOH (30 mM or 60 mM final) and reacted at room temperature for either 4 min (monomer) or 8 min (dimer) at 18°C. After quenching of the reaction, RNA samples were precipitated and subjected to primer extension using Superscript III RT (Invitrogen) as described in the Supplemental Information and reference (56). Radio-labeled (^{32}P) reverse transcription products of DMS-modified RNA were then separated on 8% denaturing PAGE gel and visualized using a Typhon phosphorimager. Sequencing reactions were

used as markers and DMS experiments were conducted in duplicate.

RESULTS

Natural L39/H89 interactions do not promote efficient self-assembly

In ribosomes from all kingdoms of life, the L39/H89 interaction is structurally similar to GNRA/receptor interactions (Figure 1C and D). L39 is either a pentaloop (GUAAG in Archaea, Eukaryote) or a hexaloop (GMUAAV (with M = A, C and V = G, A, C in Bacteria) that folds like a GNRA tetraloop. In Archaea and Eukaryotes, the pentaloop L39 has the first and fourth positions in a G:A trans Shallow-groove/Hoogsteen (S:H) base pair (bp), with the fifth position in bulge. In bacteria, the G:A trans S:H bp is formed between the first and fifth positions of the hexaloop L39, with the third and sixth nucleotide positions in bulge. Both penta- and hexa-loops L39 have two conserved adenines that form a type I/IIT A-minor interaction with the second and third Watson–Crick (W:W) bps of H89 (Figure 1C). Additionally, the receptor H89, which folds as a 'S8-like' 2bp_2x_bulge module (8,12), forms a type 2a platform and a recognition module for stabilizing the docking of L39 (Figure 1C; Supplementary Figures S2 and 3). Natural L39/H89 interactions from Archaeal (e.g. *Haloarcula marismortui*) and Bacterial (e.g. *Escherichia coli*) ribosomes were both tested within the context of the tectoRNA dimer system (Figure 1B) and did not promote any significant self-assembly at 15 mM Mg^{2+} and 10°C (GUAAG/H89A, 50000 nM; GAUAAG/H89B, 200000 nM; Supplementary Tables S3 and 4). Moreover, substituting the penta- and hexa-loops for more stable classic GNRA tetraloops (excepted GAAA) only resulted in barely noticeable improvements for GURA tetraloops with respect to other tetraloops. As noted previously (8), this is likely due to the inability of natural L39/H89 sequences to fold and assemble into their native structures outside the context of the ribosome. For instance, H89 natural sequences have suboptimal A-minor submotifs and rather than folding as 2x_bulge modules, they can adopt alternative structures that are not conducive to assembly.

According to the structural syntax of RNA modules and submodules from stable RNAs, the metastable sequences of H89 can potentially be substituted by sequences from more stable and structurally equivalent submodules. Several natural 2x_bulges adopt the same tertiary structure as H89 ('S8-like' 2bp_2x_bulges) (Figure 1C ; Supplementary Figure S3A and C). Additionally, other 2x_bulges ('P12-like' 2bp_2x_bulges) have different platforms that could potentially be functionally equivalent to those from H89 (8).

L39/H89 can be repaired with structurally equivalent modules

We first attempted to 'repair' L39/H89 by optimizing the sequence of the A-minor stem module X1X2X3:X11X12X13 (Figure 2A). The X1:X13 W:W bp is usually not crucial for A-minor module assembly. However, the optimal sequence for X2X3:X11X12 is CC:GG in GYRA/helix interactions (Figure 1C) (8,19). Several sequence variants of the helix receptor were tested in the tectoRNA context for their ability to recognize GNRA tetraloops (Figure 2). Whereas variations in binding affinity could be observed among them, these variants had the same GNRA selectivity profile. This observation prompted us to hypothesize that receptors with similar structural features might share similar selectivity profiles. GNRA/helix interactions have a common A minor stem module (CCC:GGG triplet) that is best recognized by GURA and GCRA tetraloops, with GGRA tetraloops binding the least (Figure 2B, left). In the context of L39/H89, changing the natural A minor stem modules of H89a and H89B into the CCC:GGG bp triplet resulted in receptors H89A2 and H89B2 with ∼20- to 100-fold decrease in K_d for GUAA in comparison to the natural H89 receptors (H89A2, 570 nM; H89B2, 2073 nM; H89A, 55000 nM; H89B; 40000 nM) (Figure 2B, left; Supplementary Figure S4A and Table S3). Interestingly, GNRA selectivity profiles for H89A2 and H89B2 were nearly identical to those for GYRA/helix interactions (Figure 2B, left column) suggesting that H89A2 and H89B2 behaved as helical receptors rather than 2x_bulge modules.

In order to test whether the presence of a 2x_bulge module at the level of H89 could enhance binding affinity of GNRA loops, the sequence of H89 was substituted with the sequence of the receptor binding site for protein S8 from the *E. coli* (S8a1, S8a2) and *Thermus thermophilus* (S8Tta) 16S rRNAs. H89 receptors and S8 receptors belong to the 'S8-like' family of 2x_bulge modules (Supplementary Figure S3A). They are structurally identical, with a root mean square deviation of 1.4 Å (Supplementary Figure S3C). The central bp doublets of the 2x_bulges from S8 involve classic A:U and/or C:G Watson–Crick bps and are thermodynamically more stable than those from H89 receptors, which contain U:U and U:G Watson–Crick bps. As shown by RNA modeling (Figure 1D(iii) and (iv)), the tertiary structure of GUAA/S8 (type 2a) is structurally equivalent to the one of L39/H89 despite differing in sequence *(7 positions out of 17 in bacterial L39/H89*).

Like in L39/H89, the type 2a platform of the 2x_bulge module can serve as a 'ribose' stacking platform for the second nucleotide of the GNRA (or GNRA-like loop) (Supplementary Figures S2 and 3). Moreover, the second position of the GNRA can form additional contacts with the X3:X11 W:W bp of the A-minor stem module. In the tectoRNA dimer context, S8a2 and S8Tta, with a C:G bp for X3:X11, demonstrated dramatic decreases in K_ds for GUAA with respect of natural H89 receptors, in the same order of magnitude as H89A2 (Figure 2 and Supplementary Figure S4). However, their GNRA selectivity profiles were markedly different from GNRA/helix interactions (Figure 2B). S8a2 and S8Tta also had decreased K_ds for all other GNRA loops. For instance, S8a2 was able to assemble with GGAA (290 nM) and GGGA (961 nM) almost as well as with GUAA (220 nM) and GUGA (500 nM), demonstrating that the 2x_bulge was likely to contribute to this change of selectivity. Remarkably, the presence of a U:G in X3:X11 dramatically increases discrimination for GUAA versus the other loops. S8a1, with a U:G bp for X3:X11, improved binding affinity by 75–100 fold with respect to R89A1 and R89B1. It also displayed a much greater selectivity for GUAA versus any other loops: its Kd for GUAA

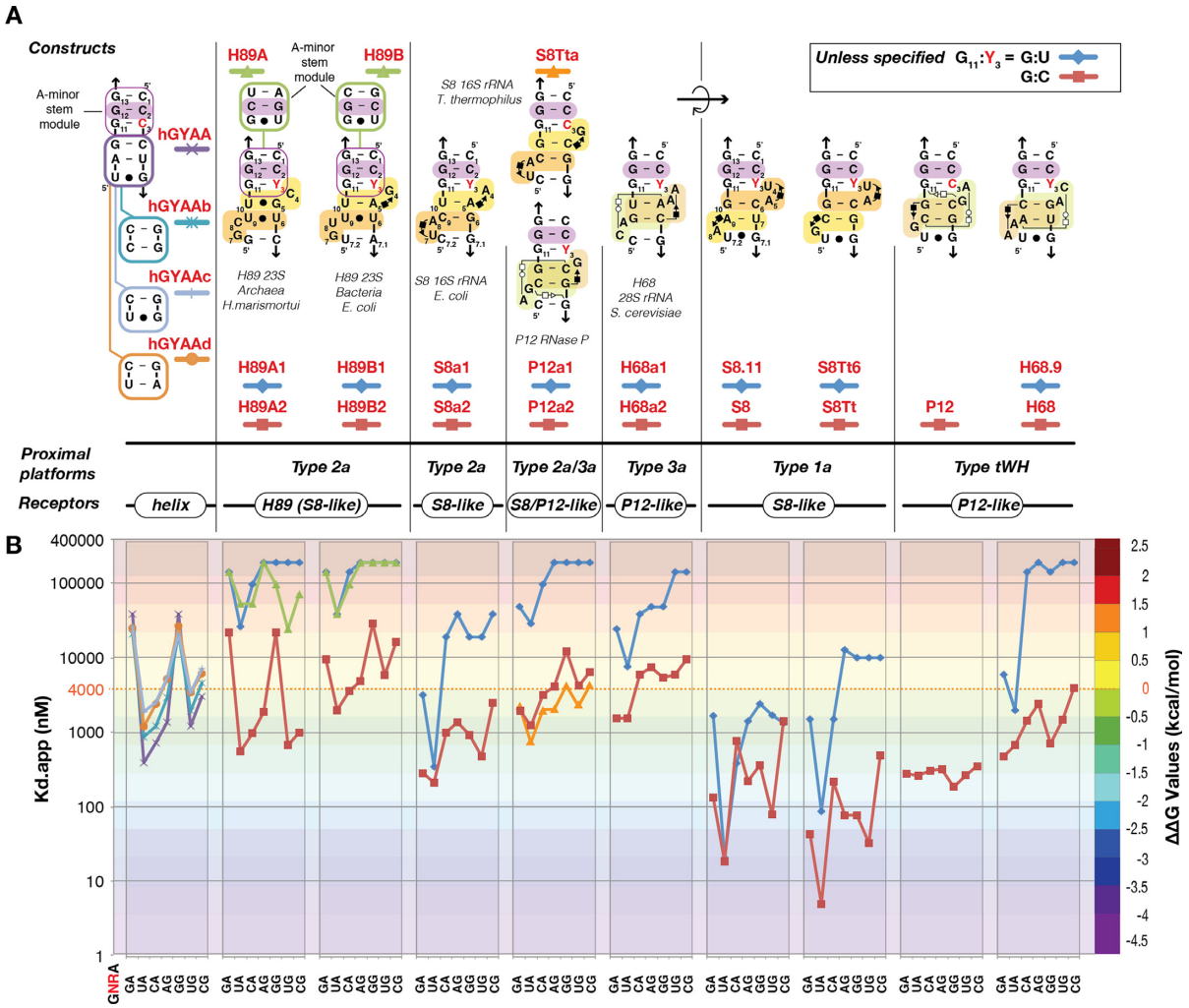

Figure 2. Re-engineering the L39/H89 interaction into more stable variants with S8-like and P12-like receptors. (**A**) 2D diagrams of the different receptor constructs tested. The type of proximal platform from each receptor is indicated. Sub-modules in the receptor include the A minor stem (with A minor G:C in violet) and different platforms (type 1a (orange), 2a (pale yellow), 3a (pale brown) and tWH (pale green)). (**B**) Selectivity profiles of variants of H89, S8-like and P12-like interactions tested against GNRA (except GAAA). Colors are indicative of the relative stability of each loop/receptor interaction at 10°C in presence of 15 mM $Mg(OAc)_2$, with 0 kcal/mol corresponding to a K_d of 4000 nM (see 'Materials and Methods' section). Y3, C (red square); Y3, U (blue diamond).

(358 nM) was approximately nine times lower than the one for GGAA and about 55 to 110 times lower than the Kds for the other GNRA loops.

Novel L39/H89 interactions can be engineered with functionally equivalent nucleotide platform modules

Type 3a platforms as docking sites for GNRA. P12-like 2x_bulge receptors, such as 'P12' from Rnase P RNAs and 'H68' from the large RNA ribosomal subunits, have nucleotide platforms that are structurally different from those observed in S8-like 2x_bulge modules (Figure 1; Supplementary Figures S2 and 3). As exemplified by S8Tt and P12 receptors, sequence variations in no more than 2 nt positions in the 2x-bulge module can specify either for the type 2a and type 1a platforms of S8-like receptors, or the type 3a and tWH platforms of P12-like receptors (Figure 1C; Supplementary Figure S3A and B). As such, P12-like receptors are not structurally equivalent to H89 but could likewise provide functional stacking platforms for docking RNA loops. Type 3a platforms mimic type 2a platforms and provide 'ribose' stacking for docking GNRA tetraloops in a way that resembles most H89/L39 interactions (Figure 1D (vi) and Supplementary Figure S3). In the context of the tectoRNA dimer, P12a2 and H68a2, with a C:G bp for X3:X11, were both able to assemble to GNRA loops but with different patterns of GNRA selectivity and not as well as S8-like receptors (Figure 2 and Supplementary Table S3). Moreover, P12a1 and H68a1, with a U:G bp for X3:X11, did not performed as well as S8a1 (Figure 2B) but nevertheless increased selectivity for GUAA versus most of the other GNRA loops. These data suggest that type 3a platforms are not fully equivalent to type 2a platforms, and emphasize that the structural differences between S8-like and P12-like receptors can have significant effects on the ability to recognize GNRA loops.

Type 1a and tWH platforms as docking sites for GNRA. In bacterial 16S rRNAs, it is not the type 2a but the type 1a platform of the S8 receptor that specifically interacts with protein S8 (Figure 1). This platform is very similar to the type 1a adenosine platform used in the GAAA/11nt interaction (Figure 1) (37). In the P12 and H68 receptors, it is the tWH platform that is used for stacking U-turn containing terminal loops (Figure 1; Supplementary Figures S2 and 3). Both type 1a and tWH platforms allow base-base stacking between the loop and the nucleotide platform rather than ribose-base stacking like in H89/L39 (Supplementary Figure S2). Moreover, both S8-like and P12-like receptors have an internal two-fold pseudo-symmetry that suggests that their platforms could potentially be swapped (Supplementary Figures S2 and 3). Therefore, we hypothesize that type 1a and tWH platforms of S8-like and P12-like receptors, respectively, could potentially be used to promote docking to GNRA tetraloops (Figure 1). According to the 3D model of the GNRA/S8 (type 1a) interaction, type 1a platforms can provide optimal stacking for GNRA loops meanwhile preserving the H-bond contacts occurring between L39 and H89 (Figure 1D(v) and Supplementary Figure S2). By contrast, the tWH platform of P12-like modules, while optimal for docking U-turn containing terminal loops, is not as optimal for docking GNRA tetraloops: for instance, the precise spacing of the loop/receptor interaction with P12-like modules is slightly different from the one with S8-like modules as the tWH platform shifts the stacking nucleotide by approximately half a base pair, leading to a potential steric clash when GNRA loops dock to P12-like receptors (Supplementary Figure S2). Because of their different base-pair patterns, P12-like receptors might not be as easily swappable as S8-like receptors (Supplementary Figure S3). It is however possible that the flexible spring-like nature of helical stems might provide local adjustments for the efficient formation of GNRA/P12-like receptor within the context of tectoRNA self-assembly dimers.

As anticipated, S8 (type 1a) constructs had remarkable self-assembly properties within the context of tectoRNA dimers. S8 and S8Tt had K_ds for GUAA of 19 and 5 nM, respectively. They assembled to GUAA with 11- to 150-fold improvements in binding affinity in regard to their respective constructs with proximal type 2a platform (S8a2 and S8Tta) (Supplementary Figure S4E). Moreover, binding improvements of S8 and S8Tt versus H89 variants (H89A2 and H89B2) range from 30- to ∼400-fold for GUAA and ∼70- to ∼500-fold for GGAA, a tetraloop that is poorly recognized by H89A2 and H89B2 (Supplementary Figure S4E). Both S8 and S8Tt favored GUAA versus all the other loops in the following order GUGA>GGAA>GAGA>GGGA>GCAA>GCGA folds (Figure 2; Supplementary Figure S4 and Table S3). This data demonstrates that the type 1a platform from S8 modules can be used as a docking site for GNRA loops leading to interactions with affinity comparable to the one of the GAAA/11 nt interaction (24,31), albeit with lower selectivity.

P12 and H68 receptors with tWH platforms performed better than their type 3a counterparts (P12a2 and H68a2) by one order of magnitude (Figure 2B). However, P12 displayed poor GNRA selectivity and its Kds for GNRA were significantly higher than those obtained from type 1a S8 receptors. P12 bound all loops within the same range of K_ds (191–367 nM), with a slight preference for GGGA. H68 assembled best with GGAA (487 nM), GUAA (688 nM) and GGGA (732 nM) but did not discriminate the other loops by more than 8-fold.

The Y3:G11 bp contributes to GUAA selectivity

According to the structure of L39/H89 from Archaea and Eukaryotes, U3:G11 bp is responsible for preferentially recognizing the uracil at the second position of the GUAAG loop in L39/H89 (Supplementary Figure S5). This is likely favored by the formation of a specific H-bond between atom O4 of the uracil of L39 and G11 of H89 and might be enhanced by chelation of an ion (possibly a potassium) (Supplementary Figure S5) (15,57). Like type 2a receptors, most of the type 1a, type 3a and tWH receptor variants that were tested with a U:G bp instead of a C:G bp for position Y3:G11, displayed a significant increase of selectivity for GUAA versus the other loops (Figure 2 and Supplementary Figure S6). S8.11 and S8Tt6 significantly improved the receptor selectivity for GUAA but with slightly higher K_ds than their C:G counterparts (S8 and S8Tt) (Figure 2B, Supplementary Figures S4, 6 and Table S3). Among all the receptors with increased selectivity for GUAA, S8.11 had the lowest K_d for GUAA (21 nM): it favored GUAA versus GCAA by ∼19-fold and GUAA versus all the other GNRA loops by 65- to 85-fold. S8Tt6 followed a similar trend. Interestingly, H68.9, a P12-like receptor with a U3:G11 bp, also had a significant increase of selectivity for GUAA in comparison to H68 (with a C3:G11 bp). Despite the fact that H68.9 had an increased K_d for GUAA (2054 nM) with respect of H68 (688 nM), H68.9 recognized better GUAA than GGAA whereas the opposite trend was observed for H68. As a general trend, these data suggest that the U:G bp at position X3:X11 has a structural role for the specific recognition of GUAA versus other GNRA loops. The importance of this base pair for selectivity is also corroborated by the fact that S8.3, with a C:C mismatch bp in X3:X11, had a significant loss of binding affinity for GUAA but only moderate loss for the other loops in comparison to S8 (Supplementary Figure S4 and Table S3). Similarly, in both P12 and H68 contexts, variants with a C:C bp in X3:X11 (P12b and H68.2) had more significant loss of binding affinity for GYRA loops than for GRRA loops.

The length of the A-minor stem is important for optimal stacking with the platform

In the tectoRNA dimer system, the optimal length of the stems separating the two interacting GNRA/receptor modules is 11 bp (24,30,31). For the S8 and P12 sequences, shortening or increasing the stem length by 1 bp led to a significant loss of binding affinity that was associated to changes of selectivity profiles, corroborating our previous data obtained from this tectoRNA system (24,30,31) (Supplementary Figure S7). Shortening the stem length proved to be much more detrimental in the P12 context than in the S8 context (compare S8.24 and P12a in Supplementary Figure S7B). This substantiates the model where shortening the

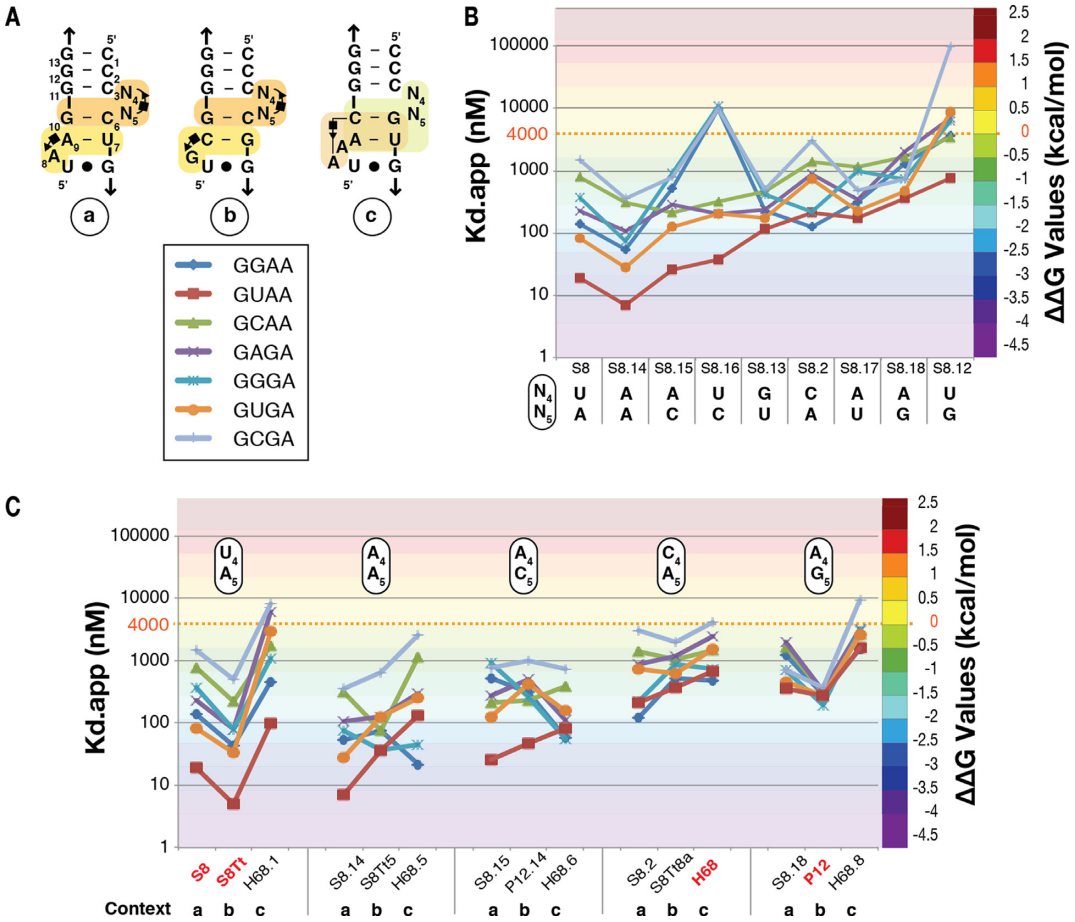

Figure 3. Modulation of binding affinity and loop selectivity through point mutations in S8-like and P12-like nucleotide platforms. (**A**) List of GNRA targets and receptor contexts within which the sequence of the proximal platform was changed: (a) S8; (b) S8Tt/P12; (c) H68. (**B**) Binding affinity and loop selectivity of various platforms sequences in the S8 structural context. (**C**) Variation of binding affinity and loop selectivity is both dependent on the sequence identity of the proximal and distal platforms. Different receptors with the same platform can have different GNRA selectivity profiles. All data were obtained at 15mM Mg^{2+} and 10°C as described in 'Materials and Methods' section.

stem in the P12 context would further create a steric clash between the second position of the GNRA tetraloop and the tWH platform. However, increasing the stem length in both S8 and P12 contexts led to GNRA selectivity profiles resembling the one for GYRA/helix interactions (S8.23 and P12d, Supplementary Figure S7), suggesting that the docking of the GNRA to the receptor might rely on looser stacking interactions with its platform.

Sequence variations within nucleotide platforms modulate binding affinity and GNRA selectivity

In order to demonstrate the role of the platforms in the engineered versions of L39/H89, we tested several variants within the bulging nucleotides involved in the proximal and/or distal platforms of S8-like and P12-like receptors (Figure 3; Supplementary Figures S7 and 8).

The removal of the bulging nucleotides (N4–5 and N8) from S8 led to hGYAA with a typical GYRA/helix selectivity profile similar to the one of H89A2 (Figure 2 and Supplementary Figure S7). Deletion of the proximal type 1a platform of S8 also had a typical GYRA/helix selectivity profile (S8.20). However, deletion of the distal type 2a platform of S8 (S8.19) maintained the selectivity profile of S8, with lower Kds for all loops (Supplementary Figure S7B). In fact, S8.19 turned out to be the best GUAA binder with a K_d of approximately 1 nM. This clearly demonstrates that the distal platform influences binding at the proximal platform. Similarly, removal of the distal type 3a platform of H68 (S8.28) and P12 (P12c) maintained the GNRA selectivity profiles of H68 and P12, respectively, but with higher K_ds (Supplementary Figure S8A, right column, Supplementary Table S3). In good agreement with the atomic structures of 2x_bulges (Supplementary Figure S3A and B), this data suggests that type 3a and type tWH platforms are less independent from one another than type 1a and type 2a platforms.

Interestingly, some S8-like and P12-like receptors differ by no more than two to four nucleotides positions, suggesting that only minor sequence changes are necessary for controlling the folding of 2x_bulge motifs into S8-like or P12-like receptors (e.g. compare P12 and S8Tt and H68 and S8 Figure 2A). According to known X-ray structures, G:U, A:A, U:A, U:C and A:C cis Hoogsteen edge (H): Shallow-groove edge (S) bps are the most abundant base pairs that contribute to the formation of type 1a platforms

(58) (Supplementary Figure S8). Within the context of the S8 receptor, proximal AA, AC and UA and UC platforms at positions N4–5 had Kds for GUAA, ranging between 7 and 38 nM, but they could also display striking changes in loop selectivity (Figure 3B). For instance, the UC platform (S8.16) had a dramatic increase of selectivity for GUAA versus GGRA and GCGA tetraloops in comparison to AA, AC and UA platforms (Figure 3). By contrast, other platforms (CA, GU, AG, AU and UG) displayed weakened loop specificity with higher Kds. Interestingly, the CA platform (S8.2) was the only platform that altered the specificity from GUAA to GGAA within the context of S8 (Figure 3B). S8.2 bound GGGA and GUAA equally well. This pattern of GNRA selectivity was the same as the one of H68, the natural P12-like receptor with a CA tWH platform (Figure 3C). These data strongly suggest that the sequence of the proximal platform can significantly alter GNRA selectivity and can possibly affect the local structure of the 2x_bulge toward either S8-like or P12-like conformations.

We further explored whether the 2x_bulge structural context can affect binding affinity and selectivity profiles of variants with proximal platforms of identical sequence (Figure 3C). UA platform in the S8, S8Tt-P12 and H68 contexts led to receptors with a marked preference for GUAA. While S8 and S8Tt had almost identical GNRA selectivity patterns, H68.1 preferred GGRA versus GUGA and GAGA (Figure 3C). Changing the UA platform into the AA platform had little effect on the loop selectivity in the context of S8 (S8.14). However, the same change in the context of S8Tt-P12 (S8Tt5) and H68 (H68.5) dramatically reduced selectivity for GUAA versus the other loops: S8Tt5 bound GUAA and GGGA loops with the same Kd (37 nM) whereas H68.5 displayed a marked preference for GGAA followed by GGGA, with its GNRA selectivity profile resembling the one of H68 (Figure 3). Furthermore, S8.15, with an AC platform in the S8 context, had increased selectivity for GUAA versus GGRA in comparison to S8 whereas P12.14, with an AC platform in the S8Tt context, had reduced selectivity for GUAA in comparison to S8Tt. The selectivity pattern of H68.6, with an AC platform in the H68 context, was closer to the one of H68, albeit with improved affinity. The AG platform, which is characteristic of P12, led to very significant reduction of selectivity toward all loops in both S8 and H68 contexts (S8.18 and H68.8 in Figure 3C), resembling most the profile of P12. Overall, these data illustrate that different 2x_bulge contexts affect and modulate the selectivity of recognition of GNRA tetraloops by proximal platforms with identical dinucleotide sequence. They also indicate that small changes at receptor positions that are not directly in contact with the loop target can still affect the way the loop is recognized. This is further supported by the fact that a G:U bp is favored over regular Watson Crick bps for the distal closing bp of S8-like receptors (X7.2:X7.1 bp Supplementary Figure S8B). Considering that both classes of 2x_bulge receptors have structural contacts between proximal and distal platforms (Supplementary Figure S3), the distal platform sub-module likely affect the way the proximal platform is folded. This is especially true in the context of P12-like receptors for which tertiary contacts are more pronounced. As the distal and proximal platform sub-modules can act in a more independent fashion in S8-like receptors, our data support the idea that S8-like receptors are somewhat more robust to mutations than P12–like receptors. Moreover, as it is the case for S8Tt-P12 receptors, our data also emphasize that very small variations in sequences lead to the possibility to switch from one class of receptor to another.

S8-like receptors and P12-like receptors are related through sequence/structure space

S8, S8Tt, P12 and H68 are linked through their sequence/structure space (Figure 4), suggesting that receptors with S8-like conformation, such as S8 or H89, could have transiently adopted P12-like conformation through evolution. Remarkably, P12 and S8Tt, which are only two mutations away from one another, displayed very distinct selectivity profiles, which are each characteristic of the class of receptor to which they belong (Figure 4B). S8, which is three mutations away from S8Tt, shares with S8Tt the same characteristic selectivity profile (Figures 2B and 4B). S8Tt5, the sequence intermediate between S8Tt and P12, equally favored GUAA and GGGA, the preferred tetraloops recognized by S8Tt and P12, respectively. Moreover, S8Tt5 did not bind GUGA as well as S8Tt and S8. This suggests that this receptor might eventually have a propensity to switch from one class to another upon docking of a particular GNRA loop. Walking through the sequence space from S8 to H68 indicates that the shift in specificity toward GGAA essentially resulted from the change in platform sequence from UA to CA (Figures 3 and 4). S8, S8.5 and H68.1, with a UA proximal platform, favored GUAA whereas H68, S8.4 and S8.2, with a CA platform, slightly favored GGAA versus GUAA and GGGA (Figure 4B). S8.4, with the same distal platform as S8 but the same proximal platform as H68, had a very similar profile to H68 but with improved K_ds for all loops (Figure 4). This indicates that the additional bulging adenine in the distal platform of H68 is somewhat destabilizing the GNRA/receptor interaction. The presence of this extra adenine in H68.1 is also likely responsible for H68.1 decreased affinity in comparison to S8. The change from G10:C6 bp to C10:G6 bp had little effect on the receptor profile and binding affinity (Figure 4B). The only noticeable trend was that S8 (like S8Tt) preferred GUGA versus GGRA whereas S8.5 and H68.1 had the opposite trend. Overall, these data show that the sequence space associated to 2x_bulges allows possible transitions from S8-like receptors to P12-like receptors (and/or vice versa) without loosing binding affinity for the GUAA tetraloop (Figure 4). Indeed, despite slightly preferring GGRA tetraloops versus GUAA, P12-like receptors still maintain good binding to GUAA. Therefore, our data suggest that GNRA/2x_bulge interactions are modules with high mutational robustness and sequence/structure plasticity.

Engineered H89 receptors bind to natural penta-loop L39 but recognize hexa-loop L39 poorly

Like natural H89 receptors, helix receptor hGYAA and other S8-like receptors with proximal type 2a platform

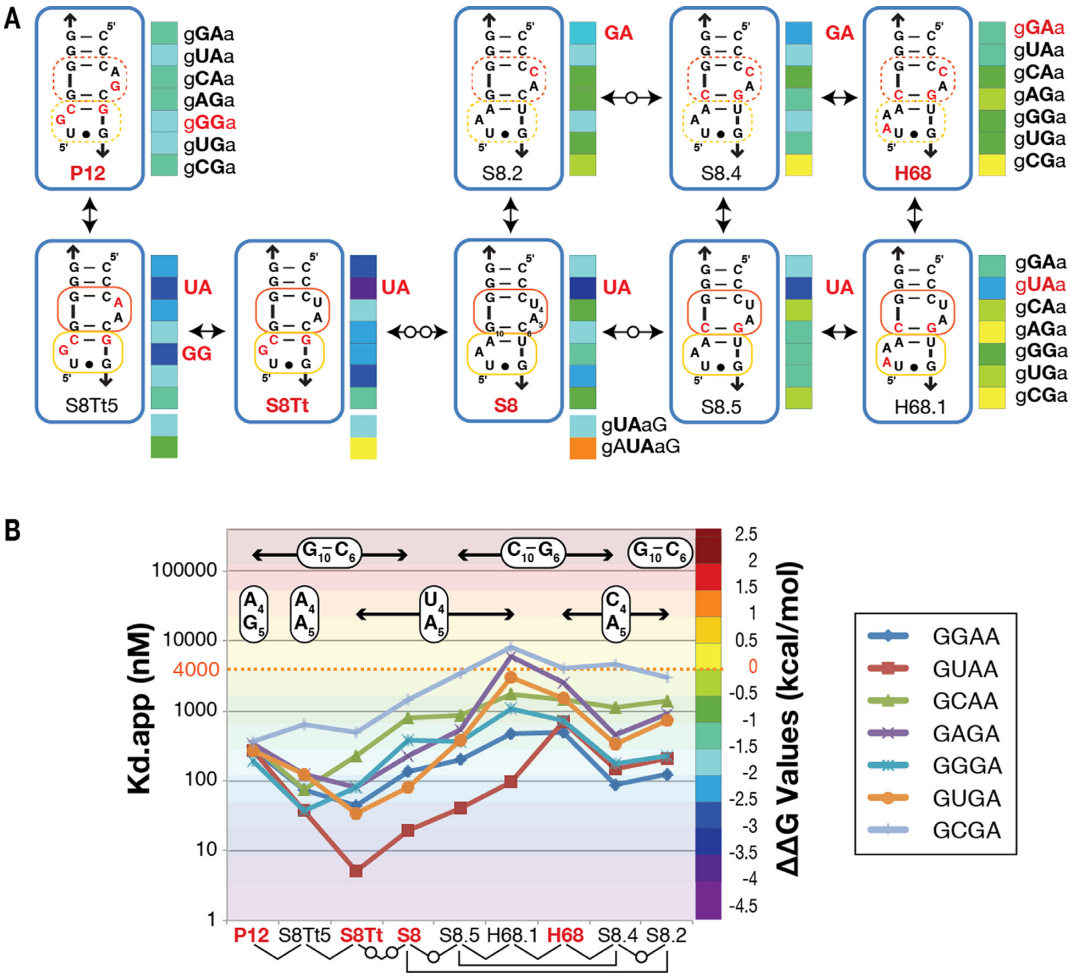

Figure 4. Moving within the sequence/structural space of S8-like and P12-like receptors. (**A**) Mutational pathways linking P12, S8Tt, S8 and H68. Nucleotides in red indicate sequence variations from S8 (reference sequence). Double arrows indicate single point mutation transitions between two receptors. The number of white nodes on arrows indicates the number of intermediates between two receptors. Selectivity profiles are shown as colored barcodes for receptors tested against GNRA (except GAAA), GUAAG and GAUAAG loops (see legend Figure 2). (**B**) Variations of Kds and GNRA selectivity from P12 to H68 via S8Tt and S8. All data were obtained at 15mM Mg^{2+} and 10°C as described in 'Materials and Methods' section.

(S8a1, S8a2) did not promote efficient recognition of the GUAAG pentaloop and GAUAAG hexaloop (Figure 5A). By contrast, all tested receptors with type 1a platforms (S8Tt5, S8.15, S8.11, S8, S8tt2, S8.14, S8Tt, S8.19) and H68.5 recognized GUAAG with K_ds ranging between 122 and 520 nM (Figure 5). For instance, the pentaloop L39 GUAAG bound to S8 and S8Tt with K_ds of 170 and 214 nM, respectively. The best receptor was S8.14, with a K_d of 122 nM. However, most tested receptors did not have significant binding affinity for the hexaloop GAUAAG. The only exception was S8Tt5 that recognized the hexaloop L39 with a K_d of 909 nM. The natural L39 loops were usually not recognized with an affinity stronger than those observed for GNRA tetraloops. Among all the receptors tested, S8.11 proved to be the exception. It displayed a strong selectivity for the pentaloop L39. After GUAA (K_d 21 nM), the second most favored loop recognized by S8.11 was GUAAG (Kd 314 nM) (Figure 5). Considering that the best 2x_bulge receptors tested in this study displayed a marked preference for GUAA versus all the other GNRA (Figure 5B), this data suggests that the ancestral form of L39 was likely a GUAA tetraloop (Figure 5C). As such, the natural penta and hexa-loop L39 can be seen as evolved metastable versions of GUAA (Figure 5C).

2x_bulge receptors are metastable and prone to fold into alternative structures

According to the sequence of natural H89 receptors, it is apparent that these receptors can adopt alternative secondary structures (Figure 6A) (18). For instance, rather than folding into 2x_bulge modules as observed in the ribosome crystallographic structures, natural H89 sequences can form extended helical elements and, as shown previously, they do not promote efficient assembly with L39 (see above). This suggests that the sequence of H89 evolved to be metastable. However, the sequence of other 2x_bulges like the one of S8, B7.8 or S8.2 can also potentially form alternative base pairs (Figure 6C). For instance, the B7.8 receptor previously isolated by Costa and Michel (29) was initially proposed to fold as a 3 nt internal loop (Figure 6C, left side) based on its predicted thermodynamic stability using un-

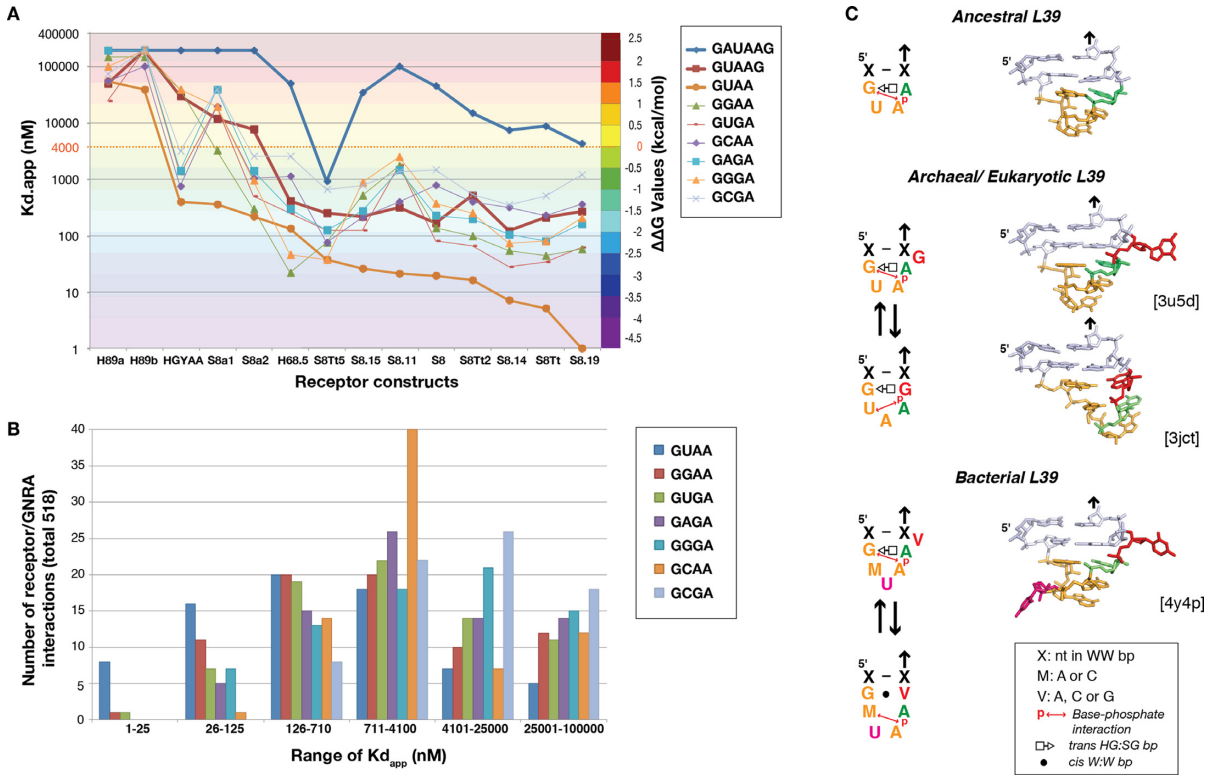

Figure 5. Binding affinity for GUAAG pentaloop and GAUAAG hexaloop for selected receptors and range of Kds for all GNRA/receptor interactions tested in this study. (**A**) Several variants of H89, S8 and H68 were tested at 10°C in presence of 15 mM $Mg(OAc)_2$. Binding affinity for GUAAG (red square) and GAUAAG (blue diamond) are compared to the one for other GNRA tetraloops. (**B**) Range of K_ds observed for all receptor/GNRA interactions tested in this study (total 518). (**C**) Evolution of L39 from an ancestral GUAA tetraloop by insertion of nucleotides: L39 from Archaea and Eukaryote is a metastable pentaloop. In yeast pre-60S ribosomes (Eukaryote), L39 is not folded as a GNRA-like [PDB ID: 3jct]. In the native ribosome, it is folded as a GNRA-like structure once docked to H89 [PDB ID: 3u5d].

afold (54). However, its sequence differs from S8 at only 4 nt positions within the A minor submodule. As expected, it can fold into a 2x_bulge module that behaved like S8 (Figure 6). In order to garner more information about the structural state adopted by the S8 receptor in presence and absence of its cognate GUAA target, DMS chemical probing was performed on S8 in the tectoRNA context in absence (monomer) or presence of GUAA probe (dimer) (Figures 6B and Supplementary S9A). In absence of GUAA, adenine positions within the receptor were shown accessible to DMS indicating that the central Watson–Crick bp doublet of the 2x_bulge module did not form. By contrast, in presence of GUAA target, these adenines became protected or partially protected, indicating that the 2x_bulge formed by induce fit when binding to GUAA. Therefore, like natural H89 receptors, the engineered S8 receptor retains metastability in absence of cognate loop. This observation is likely to be a general feature of 2x_bulges. Additional indications for the metastability of 2x_bulges were also provided by unexpected GNRA selectivity profiles for several variants (compare P12h with S8.2 and S8Tt8 with S8Tt). For instance, P12h, which can form an AC platform like P12.14, was initially expected to behave more like P12.14 (Supplementary Figure S7B). However, its GNRA selectivity profile was identical to S8.2, with a CA platform. This prompted a re-evaluation of the local secondary structure of P12h. More likely, P12h adopts an alternative fold with a CA platform (Figure 6C). We then generated S8Tt8, a derivative of S8Tt with a CA platform, to test whether its behavior mimicked that of P12h (Supplementary Figure S9B). Contrastingly, it demonstrated unpredicted behavior that seemed to destabilize the loop/receptor interaction for all loops (Supplementary Figure S9C). S8Tt8 not only had a preference for the GUAA loop, but the binding affinity for GGAA was greatly destabilized, with a K_d of 12.3 μM. Upon closer examination of its primary sequence, S8Tt8 could also potentially adopt an alternative fold to generate a helical stem with an A:C bp mismatch (Supplementary Figure S9B): the stem sequence below the receptor of S8Tt8 likely promoted the formation of a helix over the predicted 2x_bulge secondary structure. Changing the primary sequence of the lower stem to drive the formation of the 2x_bulge structure over the helix (S8Tt8.a) not only restored the interaction between all GNRA tetraloops, but had a profile similar to that of P12, with GUAA and GGAA tetraloops having comparable K_d's (370 and 529 nM, respectively). Interestingly, S8Tt, which can also adopt the same type of alternative fold as S8Tt8, favored the 2x_bulge fold. Accordingly, changing the primary sequence in the lower stem led to S8Tt4a, which behaved like S8Tt. These data illustrate how small variations of sequences can dramatically affect the thermodynamic stability of alternative base pairs localized in the vicinity of 2x_bulges and eventually drive their structure into alternative folds.

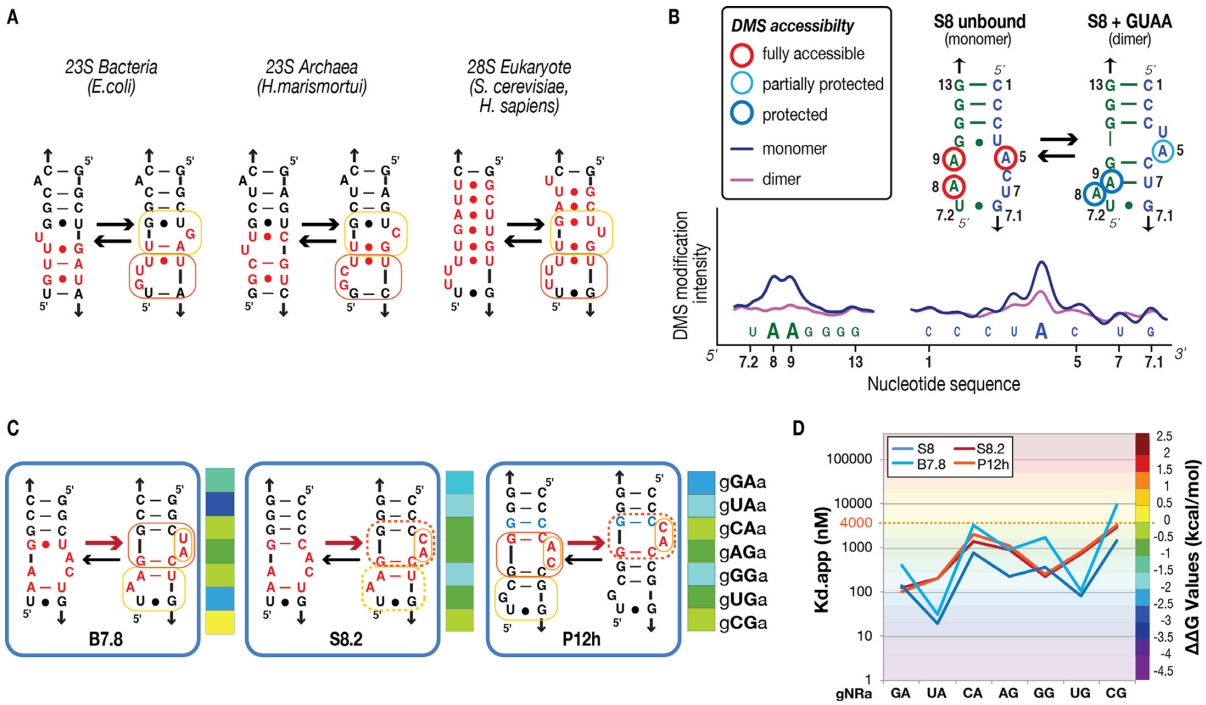

Figure 6. 2x_bulge modules are metastable RNA modules. (**A**) Alternative secondary structures possibly adopted by H89 receptors from Bacteria, Archaea and Eukaryotes. According to unafold (54), the alternative structures (shown on the left) are expected to be thermodynamically more stable than those according to atomic structures (shown on the right). Nucleotide positions in red are those involved in alternate structures. (**B**) DMS chemical probing profiles of the S8 receptor in the context of the tectoRNA monomer and dimer (with GUAA probe) at 15 mM $Mg(OAc)_2$ and 18°C. See also 'Materials and Methods' section and Supplementary Figure S7 for more experimental details. (**C**) The sequence of 2x_bulges B7.8, S8.2 and P12h can fold in two distinct secondary structures. GNRA selectivity profiles are consistent with these receptors folding as 2x_bulges (shown on the right). The B7.8 receptor previously isolated by Costa and Michel (29) was initially proposed to fold as a 3 nt internal loop, which is calculated to be 3.8 kcal more stable than the 2x_bulge module (C, left side). (**D**) K_ds and selectivity profiles of the constructs in (C) in comparison to S8.

DISCUSSION

The evolution of the long-range interaction L39/H89 within the ribosome

It has been proposed that the evolution of the ancestral ribosome into the modern translational apparatus has undergone multiple phases of growth corresponding to the stepwise acquisitions of capabilities for RNA folding, catalysis, subunit association, correlated evolution, decoding, energy-driven translocation and surface proteinization (47,59–62). Consequently, before the apparition of the Last Universal Common Ancestor (LUCA), the ancestral ribosome is likely to have evolved from simpler functional RNA domains that relied more on the intrinsic ability of RNA to fold and assemble into native functional structures through the formation of stable RNA–RNA interactions. Whereas the folding of the modern day ribosome takes advantage of several GNRA/helix interactions (e.g. (14,24,63)), the universally conserved long-range interaction L39/H89 is the only GNRA-like/receptor interaction identified in the ribosome. It is also one of the most central long-range interactions that likely appeared early in the complexification of the large ribosomal subunit (59–61). This interaction has been proposed to originate at an early evolutionary phase of the ribosome corresponding to the stage at which the ancestral small and large subunits of the proto-ribosome started to assemble and work in a concerted fashion (59,60). During this phase (phase 3 out of a total of six), it is possible that small, encoded peptides might have played an increasing role for stabilizing the ribosomal RNA structure but it is also a stage where stable long-range RNA–RNA interactions were most needed for promoting efficient RNA self-assembly. In the modern ribosome, L39/H89 directs the assembly of domain 2 with domain 5 within which the PTC is localized. Because of its central location in close proximity to the PTC, it is hypothesized that L39/H89 likely played a very significant role in the evolution of the ribosome. For instance, the mutation (UU2492–3 → C2493), which prevents H89 to fold as a 2x_bulge module in the ribosomal context of *E. coli*, proved to be lethal in *E. coli* cells by inhibiting the peptidyl transferase activity of the ribosome (18). Chemical probing analysis of the *E. coli* 23S rRNA indicated that the H89 variant prevented the PTC to fold into its native structure, emphasizing the crucial role of the L39/H89 interaction for ribosome folding and activity (18). L39/H89 has structural features resembling those of the GAAA/11nt interaction and tri-loop/P12-like interactions commonly found in stable ribozymes like group I and group II introns, RNase P RNAs and riboswitches (8,19,20,24,25,36,64,65) (Figure 1). However, unlike GAAA/11nt interactions, the L39/H89 interaction does not assemble well outside its natural structural context (Figure 2). As shown herein, this is likely due to the structural metastability of natural L39 and H89 modules that necessitate additional factors for stabi-

lizing their final native conformation. In *E. coli* 23S rRNA, position 2457 of H89 (position 3) and position 955 of the closing base pair of L39 are pseudo-uridines. This suggests that the conformational state of L39/H89 is under the control of post-transcriptional modification enzymes (66). Additionally, loop L39 is in close proximity to the universally conserved ribosomal protein uL13 as well as the loop D of the 5S rRNA, suggesting that uL13 and the 5S rRNA have a possible regulatory role in the stabilization of L39/H89 (16). For instance, in the *E. coli* ribosome, a mutation in loop L39 influences the structures of H89 and the loop D, which are both likely to participate in signal transmission between the ribosome centers responsible for peptide bond formation and translocation (16,67). In fact, during the late assembly and maturation of the yeast nuclear pre-60S ribosomes, L39/H89 is prevented from forming by the regulatory GTPases nug1 and nog1 (17). Nog1 serves to remodel the PTC and acts as a scaffold for recruiting many other factors and r-proteins to participate in quality control of the polypeptide tunnel (17). The PTC, which is localized a few angstroms from H89 and depends on H89 for its folding, cannot fold in its active structure until L39 is docked to H89 (17). Therefore, L39/H89 is a key structural interaction for the assembly of the 23S/28S rRNA. Based on these considerations (16–18,67), we propose that H39/H89 evolved as a structural conformational switch for regulating the assembly and functions of the proto-ribosome from an early stage of ribosome evolution.

Herein, we have generated by rational design possible ancestral versions of the L39/H89 interaction that can self-assemble *in vitro* without the need of any additional factors. Because natural L39 loops are GNRA-like loops, they could have easily originated from GNRA tetraloops, and most especially the GUAA tetraloop, by insertion of one to two extra nucleotides (Figure 5C). Additionally, in the context of the ribosome, H89 folds as a S8-like 2x_bulge with a structure identical to the RNA binding site of the ribosomal protein S8 of the bacterial ribosomal SSU. Because of its central location at few Ångströms from the PTC, it is unlikely that the overall structure of the H89 receptor could be drastically remodeled through evolution without affecting the structural integrity of the PTC. As such, the highly metastable H89 receptor could have evolved from thermodynamically more stable 'S8-like' receptor modules. Interestingly, our novel engineered GUAA/S8-like receptor interactions have retained not only the structural but also the functional features of natural L39/H89 interactions, qualifying them as possible ancestral forms of L39/H89. DMS modification patterns of the S8 RNA binding site in the context of the naked *E. coli* 16S rRNA and within the 30S subunit (68), or within the context of a small rRNA fragment in absence and presence of the S8 ribosomal protein (69,70), are similar to those observed in the context of the tectoRNA system (Figure 6 and Supplementary S9A): the bulging adenines of the S8 2x_bulge are accessible to DMS in the naked 16S rRNA but protected when bound to the S8 ribosomal protein. This corroborates that stable 'S8-like' receptors are still prone to structural metastability as they fold by induce fit upon binding to either the S8 protein or GUAA tetraloop (Figure 6 and Supplementary Figure S9). The presence of such type of RNA modules in the proto-ribosome could have been beneficial for controlling the folding and assembly pathway of the proto-ribosome from an early stage of its evolution. Indeed, local structural RNA metastability resulting from the formation of local alternative structures can prevent formation of long-range misfolded structures whereas still destabilizing locally adjacent regions in absence of the right loop trigger. Consequently, the intrinsic metastability of the ancestral L39/H89 interaction might have been harnessed for regulatory purposes that conditioned the evolution of L39/H89 toward even more metastable structures, which ended being dependent on protein factors for assuring the proper folding and assembly of the ribosome (17). In a first step, increased metastability could have easily been achieved by mutation of the ancestral GUAA L39 into the Archaeal/Eukaryotic GUAAG pentaloop and bacterial GAUAAG L39 (Figure 5C). In the late nuclear pre-60S ribosomes from yeast, the unbound L39 pentaloop is not folded as a GNRA-like loop (17), rationalizing why the L39 pentaloop is not recognized as well as GUAA by S8-like receptors (Figure 5). Interestingly, bacterial L39, which is recognized even more poorly by S8-like receptors and unlikely to fold into a GNRA-like loop by itself, might be less ancestral than its Archaeal/Eukaryotic counterparts.

Rationally designed 2x_bulge receptors have similar features as natural L39/H89

The basis for our rational design is posited on the notion that structural homology can lead to similar phenotypic behaviors. In our approach, GNRA loop specificity and affinity was used as a tool to define different classes of receptors sharing similar secondary and tertiary structures. By examining the 3D structures of H89 and other 2x_bulge modules, we were able to isolate structural subcomponents (or submodules) and systematically alter the sequence of these submodules, while maintaining the overall organization of the receptor. GUAA/'S8-like' receptor interactions turned out to have the most remarkable self-assembly properties. They are the best candidates for ancestral versions of L39/H89.

With few exceptions, most 2x_bulge receptors have a preference for GUAA. Out of a total of 74 receptors synthesized in this study, 33% of them have a K_d below 120 nM for GUAA (83% have K_ds below 4000 nM for GUAA). Overall, S8-like receptors with proximal type 1a platforms have better affinity for GUAA than the other 2x_bulge receptors with proximal type 2a, type 3a or tWH platform. S8-like receptors (e.g. S8 and S8Tt variants) form interactions that are 4 to 5 kcal/mol more stable than their natural H89 counterparts (up to 10 000-fold improvement of K_ds). Interestingly, the universally conserved G:U bp at position X11:X3 of H89 receptors is responsible for significantly increasing selectivity toward GUAA versus all the other loops and more particularly versus GUGA. Noteworthy is receptor S8.11, which not only recognize best GUAA, but prefers GUAAG to all the other GNRA tetraloops. According to the S. cerevisiae ribosome structure (15), the selectivity for GUAA can be rationalized by specific tertiary contacts between the G11:U3 bp of H89 and the uracil at the second position of L39 (GUAAG) through the coordination of an ion (possibly a potassium (57)) (Supplementary Figure S5).

Considering the fact that 2x_bulge receptors have modular platform subunits with a certain propensity for interchangeability, the ancestral L39/H89 might have taken advantage of a type 1a platform for stacking L39, instead of the type 2a nucleotide platform observed in natural H89 receptors. We have demonstrated that it is possible to walk through the sequence/structure space of 2x_bulges without significant loss of binding affinity. Upon few mutations, it is possible to interchange P12-like into S8-like receptors (Figure 4). Likewise, we can envision evolutionary pathways that could have led ancestral stable versions of L39/H89 to evolve into the less stable (more metastable) interactions presently found in the ribosome.

2x_bulge receptors in comparison to other GNRA receptors

Based on their high affinity and specificity for GNRA tetraloops, several engineered S8-like receptors rival natural and *in vitro* selected receptors for GNRA tetraloops, such as the 11nt, C7.2, C7.10, R1, IC3 and Vc2 receptors among others (20,24,31,32,36,44,64,71). For instance, S8.19, S8Tt, S8M14, S8Tt2, S8 and S8.11 have K_d's for GUAA which are comparable to those of the 11 nt for GAAA (2 nM) or R1 for GGAA (4 nM). Like our best S8-like receptors, the 11 nt, C7.2, C7.10 and R1 receptors take advantage of proximal type 1a platforms for stacking their cognate tetraloops (38). However, S8-like receptors are overall less selective for their cognate loop than the naturally occurring 11 nt receptor and the artificially selected R1, which bind GAAA and GGAA, respectively (24). Lower GNRA selectivity is also observed for the natural IC3 and Vc2 receptors, albeit these receptors have overall reduced binding affinities for GNRA (24,32,71). Among a total of 518 combinations of GNRA/2x_bulge receptors engineered in this study, up to 33% have K_ds below 700 nM, and 65% have K_ds below 4000 nM (Figure 5B). Therefore, the fact that the S8-like receptors are somewhat more promiscuous binders than the highly selective 11 nt and R1 receptors suggest that they are more evolvable and mutationally more robust than the 11 nt and R1 receptors.

Interestingly, other 2x_bulge receptors have been isolated by SELEX. An *in vitro* selection for GUGA binders led to the isolation of the B7.8 receptor by Costa and Michel (29). While initially not recognized as a 2x_bulge, the sequence and phenotype of B7.8 revealed that it folds as a S8-like module. Additionally, a recent selection for GUAA receptors led to the isolation of a new class of receptors folding as S8-like receptors (E. R. Calkins, P. Zakrevsky and L. Jaeger, in preparation). While the sequence of class R5.58 differs from the one of S8 at 8 nucleotide positions, it has the same phenotype as S8 (E. R. Calkins, P. Zakrevsky and L. Jaeger, in preparation). This suggests that 2x_bulge receptors for GNRA tetraloops are likely to constitute a large family of variants that might have multiple exemplars in nature.

RNA metastability as a general feature for other RNA modules

Besides 2x_bulge modules, many other loop/receptor interactions are likely metastable when the receptors and/or loops are in their unbound state. This has been observed for the T-loop (72), Z-turn loop motifs (73), GAAA/11nt interaction (21,25) and several other loop/receptor interactions isolated by SELEX such as R1 (24). Moreover, suboptimality in terms of folding and assembly has been noticed for other recurrent RNA modules identified in the ribosome. For instance, the sequence of some A minor junction modules (8) and right angle modules (9,10,74) do not fold and assemble optimally outside the context of the ribosome. Like L39/H89, they can require additional RNA elements, ribosomal proteins or post-transcriptional modifications for efficient folding and assembly in their natural context. RNA metastability plays an important role for RNA remodeling during the regulated and controlled maturation of the ribosome (17,73–75). As a matter of fact, the intrinsic ability of RNA to form metastable structures is at the root of many remarkable functions of RNA, including riboswitches (76).

New self-assembling building blocks for RNA synthetic biology and nanotechnology

As more and more RNA modules are characterized, our understanding of the fundamental nature and hierarchical organization of RNA is rapidly broadening. The knowledge gained from this work and similar studies contributes to elucidate the principles that dictate RNA modular interactions and increases the repertoire of rational design rules that could ultimately lead to an infinite number of engineering possibilities for RNA nanotechnology and RNA synthetic biology (e.g. (6,36,40–42,48,77)). Our engineered receptors increase the tool kit of loop/receptor interactions presently available (36,40,78). This work also demonstrates that our present understanding of the syntax and modularity of RNA is amenable to the precise rational engineering of novel RNA tertiary interactions relying on tertiary contacts and non-canonical base pairs. Rational design strategies as the one used herein, can be applied to other suboptimal structures to engineer novel stable, ordered structures that can rival and complement those isolated by in vitro evolution techniques. S8-like receptors are excellent candidates for re-engineering long-range interactions within novel functional RNAs, like ribozymes or riboswitches. For instance, it would be of particular interest to re-engineer the ribosome with more stable versions of L39/H89 so as to re-create a functional biochemical model of the ancestral ribosome (47) or develop new orthologous ribosomal systems (79–81) less dependent on ribosomal proteins and post-transcriptional modifications for their assembly and functions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Maria del Carmen Jaeger and Dr Pascal Auffinger for fruitful discussions. This work is dedicated to the memory of Prof. Albert Jaeger and Prof. Jérome Lejeune.

FUNDING

National Aeronautics and Space Administration (NASA) [80NSSC17K0031 to L.J., in part]; UCSB Academic Senate, Intramural Research Grants (to L.J.). Funding for open access charge: NASA [80NSSC17K0031].
Conflict of interest statement. None declared.

REFERENCES

1. Noller,H.F. (2005) RNA structure: reading the ribosome. *Science*, **309**, 1508–1514.
2. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA*, **12**, 83–93.
3. Holbrook,S.R. (2008) Structural principles from large RNAs. *Annu. Rev. Biophys.*, **37**, 445–464.
4. Jaeger,L. (2009) Defining the syntax for self-assembling RNA tertiary architectures. *Nucleic Acids Symp. Ser. (Off)*, **53**, 83–84.
5. Butcher,S.E. and Pyle,A.M. (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.*, **44**, 1302–1311.
6. Grabow,W. and Jaeger,L. (2013) RNA modularity for synthetic biology. *F1000Prime Rep.*, **5**, 46.
7. Jaeger,L., Verzemnieks,E.J. and Geary,C. (2009) The UA_handle: a versatile submotif in stable RNA architectures. *Nucleic Acids Res.*, **37**, 215–230.
8. Geary,C., Chworos,A. and Jaeger,L. (2011) Promoting RNA helical stacking via A-minor junctions. *Nucleic Acids Res.*, **39**, 1066–1080.
9. Grabow,W.W., Zhuang,Z., Swank,Z.N., Shea,J.E. and Jaeger,L. (2012) The Right Angle (RA) Motif: a prevalent ribosomal RNA structural pattern found in group I introns. *J. Mol. Biol.*, **424**, 54–67.
10. Grabow,W.W., Zhuang,Z., Shea,J.E. and Jaeger,L. (2013) The GA-minor submotif as a case study of RNA modularity, prediction, and design. *Wiley Interdiscip. Rev. RNA*, **4**, 181–203.
11. Tishchenko,S., Nikulin,A., Fomenkova,N., Nevskaya,N., Nikonov,O., Dumas,P., Moine,H., Ehresmann,B., Ehresmann,C., Piendl,W. *et al.* (2001) Detailed analysis of RNA-protein interactions within the ribosomal protein S8-rRNA complex from the archaeon Methanococcus jannaschii. *J. Mol. Biol.*, **311**, 311–324.
12. Boutorine,Y.I. and Steinberg,S.V. (2012) Twist-joints and double twist-joints in RNA structure. *RNA*, **18**, 2287–2298.
13. Polikanov,Y.S., Melnikov,S.V., Soll,D. and Steitz,T.A. (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat. Struct. Mol. Biol.*, **22**, 342–344.
14. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905–920.
15. Ben-Shem,A., Garreau de Loubresse,N., Melnikov,S., Jenner,L., Yusupova,G. and Yusupov,M. (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. *Science*, **334**, 1524–1529.
16. Sergiev,P.V., Bogdanov,A.A., Dahlberg,A.E. and Dontsova,O. (2000) Mutations at position A960 of E. coli 23 S ribosomal RNA influence the structure of 5 S ribosomal RNA and the peptidyltransferase region of 23 S ribosomal RNA. *J. Mol. Biol.*, **299**, 379–389.
17. Wu,S., Tutuncuoglu,B., Yan,K., Brown,H., Zhang,Y., Tan,D., Gamalinda,M., Yuan,Y., Li,Z., Jakovljevic,J. *et al.* (2016) Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature*, **534**, 133–137.
18. Burakovsky,D.E., Sergiev,P.V., Steblyanko,M.A., Konevega,A.L., Bogdanov,A.A. and Dontsova,O.A. (2011) The structure of helix 89 of 23S rRNA is important for peptidyl transferase function of Escherichia coli ribosome. *FEBS Lett.*, **585**, 3073–3078.
19. Jaeger,L., Michel,F. and Westhof,E. (1994) Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J. Mol. Biol.*, **236**, 1271–1276.
20. Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276–1285.
21. Butcher,S.E., Dieckmann,T. and Feigon,J. (1997) Solution structure of a GAAA tetraloop receptor RNA. *EMBO J.*, **16**, 7490–7499.
22. Nissen,P., Ippolito,J.A., Ban,N., Moore,P.B. and Steitz,T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 4899–4903.
23. Xin,Y., Laing,C., Leontis,N.B. and Schlick,T. (2008) Annotation of tertiary interactions in RNA structures reveals variations and correlations. *RNA*, **14**, 2465–2477.
24. Geary,C., Baudrey,S. and Jaeger,L. (2008) Comprehensive features of natural and in vitro selected GNRA tetraloop-binding receptors. *Nucleic Acids Res.*, **36**, 1138–1152.
25. Fiore,J.L. and Nesbitt,D.J. (2013) An RNA folding motif: GNRA tetraloop-receptor interactions. *Q. Rev. Biophys.*, **46**, 223–264.
26. Lanier,K.A., Roy,P., Schneider,D.M. and Williams,L.D. (2017) Ancestral interactions of ribosomal RNA and ribosomal proteins. *Biophys. J.*, **113**, 268–276.
27. Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature*, **372**, 111–113.
28. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
29. Costa,M. and Michel,F. (1997) Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: Comparison with in vivo evolution. *EMBO J.*, **16**, 3289–3302.
30. Jaeger,L. and Leontis,N.B. (2000) Tecto-RNA: one-dimensional self-assembly through tertiary interactions. *Angew. Chem. Int. Ed. Engl.*, **14**, 2521–2524.
31. Jaeger,L., Westhof,E. and Leontis,N.B. (2001) TectoRNA: modular assembly units for the construction of RNA nano-objects. *Nucleic Acids Res.*, **29**, 455–463.
32. Ikawa,Y., Nohmi,K., Atsumi,S., Shiraishi,H. and Inoue,T. (2001) A comparative study on two GNRA-tetraloop receptors: 11-nt and IC3 motifs. *J. Biochem. (Tokyo)*, **130**, 251–255.
33. Liu,B., Baudrey,S., Jaeger,L. and Bazan,G.C. (2004) Characterization of tectoRNA assembly with cationic conjugated polymers. *J. Am. Chem. Soc.*, **126**, 4076–4077.
34. Davis,J.H., Tonelli,M., Scott,L.G., Jaeger,L., Williamson,J.R. and Butcher,S.E. (2005) RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J. Mol. Biol.*, **351**, 371–382.
35. Nasalean,L., Baudrey,S., Leontis,N.B. and Jaeger,L. (2006) Controlling RNA self-assembly to form filaments. *Nucleic Acids Res.*, **34**, 1381–1392.
36. Ishikawa,J., Fujita,Y., Maeda,Y., Furuta,H. and Ikawa,Y. (2011) GNRA/receptor interacting modules: versatile modular units for natural and artificial RNA architectures. *Methods*, **54**, 226–238.
37. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Szewczak,A.A., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) RNA tertiary structure mediation by adenosine platforms. *Science*, **273**, 1696–1699.
38. Watkins,A.M., Geniesse,C., Kladwang,W., Zakrevsky,P., Jaeger,L. and Das,R. (2018) Blind prediction of noncanonical RNA structure at atomic accuracy. *Sci. Adv.*, **4**, eaar5316.
39. Jaeger,L. and Chworos,A. (2006) The architectonics of programmable RNA and DNA nanostructures. *Curr. Opin. Struct. Biol.*, **16**, 531–543.
40. Ishikawa,J., Furuta,H. and Ikawa,Y. (2013) RNA tectonics (tectoRNA) for RNA nanostructure design and its application in synthetic biology. *Wiley Interdiscip. Rev. RNA*, **4**, 651–664.
41. Grabow,W.W. and Jaeger,L. (2014) RNA self-assembly and RNA nanotechnology. *Acc. Chem. Res.*, **47**, 1871–1880.
42. Geary,C., Chworos,A., Verzemnieks,E., Voss,N.R. and Jaeger,L. (2017) Composing RNA nanostructures from a syntax of RNA structural modules. *Nano Lett.*, **17**, 7095–7101.
43. Davis,J.H., Foster,T.R., Tonelli,M. and Butcher,S.E. (2007) Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. *RNA*, **13**, 76–86.
44. Afonin,K.A., Lin,Y.P., Calkins,E.R. and Jaeger,L. (2012) Attenuation of loop-receptor interactions with pseudoknot formation. *Nucleic Acids Res.*, **40**, 2168–2180.
45. Hsiao,C., Mohan,S., Kalahar,B.K. and Williams,L.D. (2009) Peeling the onion: ribosomes are ancient molecular fossils. *Mol. Biol. Evol.*, **26**, 2415–2425.
46. Johnsson,P., Lipovich,L., Grander,D. and Morris,K.V. (2014) Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta*, **1840**, 1063–1071.

47. Hsiao,C., Lenz,T.K., Peters,J.K., Fang,P.Y., Schneider,D.M., Anderson,E.J., Preeprem,T., Bowman,J.C., O'Neill,E.B., Lie,L. *et al.* (2013) Molecular paleontology: a biochemical model of the ancestral ribosome. *Nucleic Acids Res.*, **41**, 3373–3385.
48. Afonin,K.A., Kasprzak,W.K., Bindewald,E., Kireeva,M., Viard,M., Kashlev,M. and Shapiro,B.A. (2014) In silico design and enzymatic synthesis of functional RNA nanoparticles. *Acc. Chem. Res.*, **47**, 1731–1741.
49. Guo,P. (2010) The emerging field of RNA nanotechnology. *Nat. Nanotechnol.*, **5**, 833–842.
50. Chakraborty,S., Mehtab,S. and Krishnan,Y. (2014) The predictive power of synthetic nucleic acid technologies in RNA biology. *Acc. Chem. Res.*, **47**, 1710–1719.
51. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
52. Rother,M., Rother,K., Puton,T. and Bujnicki,J.M. (2011) ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res.*, **39**, 4007–4022.
53. DeLano,W.L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, Palo Alto.
54. Markham,N.R. and Zuker,M. (2008) UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.*, **453**, 3–31.
55. Paillart,J.C., Skripkin,E., Ehresmann,B., Ehresmann,C. and Marquet,R. (1996) A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 5572–5577.
56. Tijerina,P., Mohr,S. and Russell,R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.*, **2**, 2608–2623.
57. Leonarski,F., D'Ascenzo,L. and Auffinger,P. (2017) Mg2+ ions: do they bind to nucleobase nitrogens? *Nucleic Acids Res.*, **45**, 987–1004.
58. Stombaugh,J., Zirbel,C.L., Westhof,E. and Leontis,N.B. (2009) Frequency and isostericity of RNA base pairs. *Nucleic Acids Res.*, **37**, 2294–2312.
59. Petrov,A.S., Bernier,C.R., Hsiao,C., Norris,A.M., Kovacs,N.A., Waterbury,C.C., Stepanov,V.G., Harvey,S.C., Fox,G.E., Wartell,R.M. *et al.* (2014) Evolution of the ribosome at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 10251–10256.
60. Petrov,A.S., Gulen,B., Norris,A.M., Kovacs,N.A., Bernier,C.R., Lanier,K.A., Fox,G.E., Harvey,S.C., Wartell,R.M., Hud,N.V. *et al.* (2015) History of the ribosome and the origin of translation. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 15396–15401.
61. Bokov,K. and Steinberg,S.V. (2009) A hierarchical model for evolution of 23S ribosomal RNA. *Nature*, **457**, 977–980.
62. Belousoff,M.J., Davidovich,C., Zimmerman,E., Caspi,Y., Wekselman,I., Rozenszajn,L., Shapira,T., Sade-Falk,O., Taha,L., Bashan,A. *et al.* (2010) Ancient machinery embedded in the contemporary ribosome. *Biochem. Soc. Trans.*, **38**, 422–427.
63. Wimberly,B.T., Brodersen,D.E., Clemons,W.M. Jr, Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
64. Costa,M., Deme,E., Jacquier,A. and Michel,F. (1997) Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.*, **267**, 520–536.
65. Abramovitz,D.L. and Pyle,A.M. (1997) Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. *J. Mol. Biol.*, **266**, 493–506.
66. Helm,M. (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.*, **34**, 721–733.
67. Scripture,J.B. and Huber,P.W. (2011) Binding site for Xenopus ribosomal protein L5 and accompanying structural changes in 5S rRNA. *Biochemistry*, **50**, 3827–3839.
68. Moazed,D., Stern,S. and Noller,H.F. (1986) Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. *J. Mol. Biol.*, **187**, 399–416.
69. Allmang,C., Mougel,M., Westhof,E., Ehresmann,B. and Ehresmann,C. (1994) Role of conserved nucleotides in building the 16S rRNA binding site of E. coli ribosomal protein S8. *Nucleic Acids Res.*, **22**, 3708–3714.
70. Mougel,M., Allmang,C., Eyermann,F., Cachia,C., Ehresmann,B. and Ehresmann,C. (1993) Minimal 16S rRNA binding site and role of conserved nucleotides in Escherichia coli ribosomal protein S8 recognition. *Eur. J. Biochem.*, **215**, 787–792.
71. Fujita,Y., Tanaka,T., Furuta,H. and Ikawa,Y. (2012) Functional roles of a tetraloop/receptor interacting module in a cyclic di-GMP riboswitch. *J. Biosci. Bioeng.*, **113**, 141–145.
72. Zhuang,Z., Jaeger,L. and Shea,J.-E. (2007) Probing the structural hierarchy and energy landscape of an RNA T-loop hairpin. *Nucleic Acids Res.*, **35**, 6995–7002.
73. D'Ascenzo,L., Vicens,Q. and Auffinger,P. (2018) Identification of receptors for UNCG and GNRA Z-turns and their occurrence in rRNA. *Nucleic Acids Res.*, **46**, 7989–7997.
74. Sharma,I.M., Rappe,M.C., Addepalli,B., Grabow,W.W., Zhuang,Z., Abeysirigunawardena,S.C., Limbach,P.A., Jaeger,L. and Woodson,S.A. (2018) A metastable rRNA junction essential for bacterial 30S biogenesis. *Nucleic Acids Res.*, **46**, 5182–5194.
75. Kater,L., Thoms,M., Barrio-Garcia,C., Cheng,J., Ismail,S., Ahmed,Y.L., Bange,G., Kressler,D., Berninghausen,O., Sinning,I. *et al.* (2017) Visualizing the assembly pathway of nucleolar Pre-60S ribosomes. *Cell*, **171**, 1599–1610.
76. McCown,P.J., Corbino,K.A., Stav,S., Sherlock,M.E. and Breaker,R.R. (2017) Riboswitch diversity and distribution. *RNA*, **23**, 995–1011.
77. Jasinski,D., Haque,F., Binzel,D.W. and Guo,P. (2017) Advancement of the emerging field of RNA nanotechnology. *ACS Nano*, **11**, 1142–1164.
78. Furukawa,A., Maejima,T., Matsumura,S. and Ikawa,Y. (2016) Characterization of an RNA receptor motif that recognizes a GCGA tetraloop. *Biosci. Biotechnol. Biochem.*, **80**, 1386–1389.
79. Orelle,C., Carlson,E.D., Szal,T., Florin,T., Jewett,M.C. and Mankin,A.S. (2015) Protein synthesis by ribosomes with tethered subunits. *Nature*, **524**, 119–124.
80. d'Aquino,A.E., Kim,D.S. and Jewett,M.C. (2018) Engineered ribosomes for basic science and synthetic biology. *Annu. Rev. Chem. Biomol. Eng.*, **9**, 311–340.
81. Liu,C.C., Jewett,M.C., Chin,J.W. and Voigt,C.A. (2018) Toward an orthogonal central dogma. *Nat. Chem. Biol.*, **14**, 103–106.
82. Leontis,N.B. and Westhof,E. (1998) Conserved geometrical base-pairing patterns in RNA. *Q. Rev. Biophys.*, **31**, 399–455.