Unnatural Amino Acid Incorporation

to Rewrite the Genetic Code

and RNA-peptide Interactions

Thesis by

Xin Qi

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

2005

(Defended May 19, 2005)

Acknowledgments

I like to sincerely thank my advisor, Dr. Richard W. Roberts, for his great mentorship over the course of my graduate career here at Caltech. This thesis would have not been possible without his tremendous scientific vision and a great deal of guidance.

I am also in debt to Dr. Peter Dervan, Dr. Robert Grubbs, and Dr. Stephen Mayo, who serve on my thesis committee. They have provided very valuable help along the way that kept me on track in the pursuit of my degree.

I have been so fortunate to have been working with a number of fantastic colleagues in Roberts' group. Particularly, I want express my gratitude toward Shelley R. Starck, with whom I had a productive collaboration on one of my projects. Members of Roberts group, including Shuwei Li, Terry Takahashi, Ryan J. Austin, Bill Ja, Steve Millward, Christine Ueda, Adam Frankel, and Anders Olson, all have given me a great deal of help on a number of occasions in my research, and they are truly fun people to work with. I also like to thank our secretary, Margot Hoyt, for her support during my time at Caltech. I have also collaborated with people from other laboratories, e.g., Dr. Cory Hu from Prof. Varshavsky's group, and it has been a positive experience to have a nature paper together.

The five years at Caltech have been very pleasant for me, all because of the friendship I have had with a number of friends, including Jessica Mao, Amanda Cashin, Christie Morrill, and many others. I will never forget all the fun time I spent with them. I

sincerely appreciate the support from my postdoc mentor Jennifer Anthony, who gave me a lot of good advice on my presentation skill.

Last but not least, I want to express my gratitude toward my family. My husband Tianbing has always been there for me, and his constant love has been my inspiration for all the years. I am so lucky to have the gifts of my life, my two lovely daughters, Amy and Jennifer, who brought so many joys. My parents Bofu Qi and Yan Pu and my brother Yue Qi were encouraging and loving over so many years. I appreciate their faith in me and their love and support are always there with me.

Abstract

My general research direction is the interface between organic chemistry and biology. Interesting biological systems inspire target-oriented organic synthesis, new methodology development, and molecular design of novel materials. On the other hand, chemical synthesis prepares important biologically active compounds to be used in understanding involved mechanism and engineering biological systems.

We designed and synthesized a series of peptide-acridine conjugates based on the modular design principle to target RNA structures. Some of the peptide-acridine conjugates have substantially improved RNA-binding affinity and specificity relative to the peptide alone. We also generated various high-affinity inhibitors of the tRNA synthetases—aminoacyl sulfamide to create synthetic blanks in our translation extracts. The gaps were filled with chemically aminoacylated orthogonal tRNAs. This unnatural strategy enables peptides and proteins to be constructed containing a single novel residue at specific locations and will facilitate the mRNA display-based protein selection. To test the versatility of ribosome, we have constructed a series of puromycin analogs with natural and unnatural amino acids side chains and tested the effects of side chain characteristics of amino acid moiety on the activity of puromycin analogs. We found that amino acids of different stereo characteristics can be incorporated via puromycin route, and larger hydrophobic amino acids render higher potency.

Overall, these efforts demonstrate the successful use of a combination of molecular design as an efficient and facile method for generating new solutions to biological problems.

Table of Contents

Chapter 3. Probing Flexibility of Protein Synthesis *In Vitro* by the

Puromycin Analogs **66**

Chapter 4. Total synthesis of Hydroxymethylacylfulvene **96**

Appendix.

Chapter 1

Design of Acridine-N Peptide Conjugates with Enhanced

Binding Affinity and Specificity to RNA Targets

Abstract

Arginine-rich peptides and small-molecule intercalating agents provide mechanistically distinct molecular tools for RNA recognition. Here, we have worked to combine these distinct binding modules in an effort to create conjugate ligands with enhanced affinity and specificity using the bacteriophage λ N peptide/boxB RNA interaction as a model system. To do this, we designed and synthesized a series of peptide-acridine conjugates using portions of the N RNA-binding domain (11 and 22 residue peptide segments). We then compared the binding affinities, specificities, salt dependences, and structural properties of the RNA-peptide and RNA-peptide-conjugate complexes using steady-state fluorescence, CD spectroscopy, NMR, and native gel mobility shift assays (GMSA). These analyses revealed that the full-length peptide-acridine conjugate had substantially improved RNA-binding affinity (~80-fold; $K_d \sim$ 20 pM) and specificity (~25-fold) relative to the peptide alone. This binding enhancement was unique to only full length conjugates tested, implying that the structural context of acridine presentation was critical. In line with this view, the specificity enhancement we observe results because binding of the best conjugate to a noncognate P22 RNA hairpin showed only a modest (3-fold) binding enhancement. Our work supports the idea that peptide- and intercalation-based binding can be combined to create a new class of high-affinity, high-specificity RNA-binding ligands.

Introduction

The design of proteins with the capacity to recognize nucleotide sequences with high affinity and specificity has been pursued extensively. The information gained from designing DNA binders include utilizing zinc finger proteins to target sites in the major groove[1-3] and polyamides to target sites in the minor groove.[4-6] The ability of RNA molecules to fold into complex and unique three dimensional shapes and their important roles in many biological processes have generated great interest in designing sequence or structure-specific RNA-binding molecules. Recent successes of sequence-specific RNA binding molecules, the expansion of structural databases for RNA and RNA/protein complexes, together with the development of new synthetic tools, have provided new opportunities for the design of peptide or small molecule RNA binders. One approach to enhance RNA binding is to combine different binding modes using potential interaction sites that coexist in a given RNA molecule. In principle, if two binding sites are in close proximity, a dimeric derivative ligand can bind simultaneously to the two sites, resulting in binding affinity greater than either module.[7,8] This strategy was also demonstrated in our efforts to select peptide-drug conjugates as novel inhibitors of the penicillin binding protein.[9] In vitro selection of peptide-drug conjugates resulted in a penicillin-peptide conjugate that at least 100-fold higher activity than the parent penicillin itself. In this chapter, we discuss the recent successful application of these principles to create novel RNA-binding peptides, with particular emphasis on the advances achieved using intercalating agent acridine.

Recent studies have shown that the tethering of multiple binding modules can enhance RNA-binding affinity and specificity. In particular, acridine derivatives have proved very useful as intercalating modules. In a recent report, a Tat-TAR binding inhibitor consisting of a substituted acridine and a polyamine moiety was demonstrated to inhibit Tat function.[10] Their modular design principle of Tat function impairment was based on the ability of the aromatic moiety of acridine for stacking, and a polycationic anchor for contacts with the TAR RNA phosphate backbone. The two modules are linked by an aliphatic linker. Aside from the compound stacking between two bases, direct hydrogen-bond contacts with a GC base pair were also involved.

In another study, a series of cationic small molecules were synthesized and their binding abilities to defined RNA duplexes with and without bulged bases were investigated.[11] Complex stabilization and selectivity for an RNA duplex containing a single bulged base over a normal RNA duplex have been obtained with a ligand consisting of a chloroacridine moiety covalently attached to 2,6-diaminopurine through an aminoalkyl linker. It is believed that the chloroacridine moiety intercalates into the RNA duplex and the 2,6-diaminopurine interacts with the bulged base.

Another example of this type of binding enhancement is illustrated by a neomycin-acridine conjugate, synthesized by covalently linking neomycin B to 9-aminoacridine via a short spacer, as a potent inhibitor of Rev-RRE binding.[12] Its affinity to the RRE is about 50-fold higher than that of the parent neomycin B and approaches that of the Rev

peptide. These results demonstrate that the combination of different binding modes (e.g., ionic and intercalation) within one ligand is a powerful approach for enhancing the RNA affinity of synthetic binders.

Acridine conjugation has also been used in designing functional molecules. For example, attachment of an acridine moiety to a catalytic tripeptide produced an RNase mimic,[13] and attachment to a nucleic base-linker construct endowed with abasic site recognition generated DNA cleavage function at apurinic sites.[14] Acridine has also been used in conjunction with an oligonucleotide to facilitate site-selective RNA hydrolysis.[15] In this case, acridine was hypothesized to push the bulged base out of the helix and present a scission site.

We reasoned that the same principles may also apply to the N peptide/boxB RNA complex system. Bacteriophage N proteins play an essential role in transcriptional antitermination, which are critical for phage development.[16] The inhibition of transcription termination at intrinsic and Rho-dependent terminators by N protein depends on recognition of a *cis*-acting RNA-element called Nut (N utilization) on the nascent phage transcript. Together with four *Escherichia coli* host factors (NusA, NusB, NusG, and ribosomal protein S10), they form a ribonucleoprotein complex that converts the RNA polymerase into a termination-resistant form. [16,17]

The Nut site consists of a 5'-single-stranded RNA element (boxA) and a 3' hairpin (boxB).[18,19] The boxB from lambda phage is a 15-mer RNA stem-loop hairpin

containing a purine-rich pentaloop (Figure 1). The RNA-binding domain of N protein consists of an arginine-rich motif located at the NH_2-terminus.[20] The 22-residue short peptide recognizes the cognate boxB RNA with similar specificity and affinity as the intact N protein.[21] Upon complex formation, four of the pentaloop nucleotides adopt a canonical GNRA tetraloop fold[22] with the fourth adenine extruded;[23,24] the peptide forms a bent α helix and binds tightly to the major groove of the RNA.[23-25]
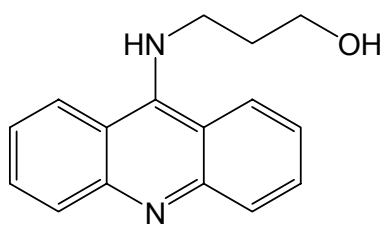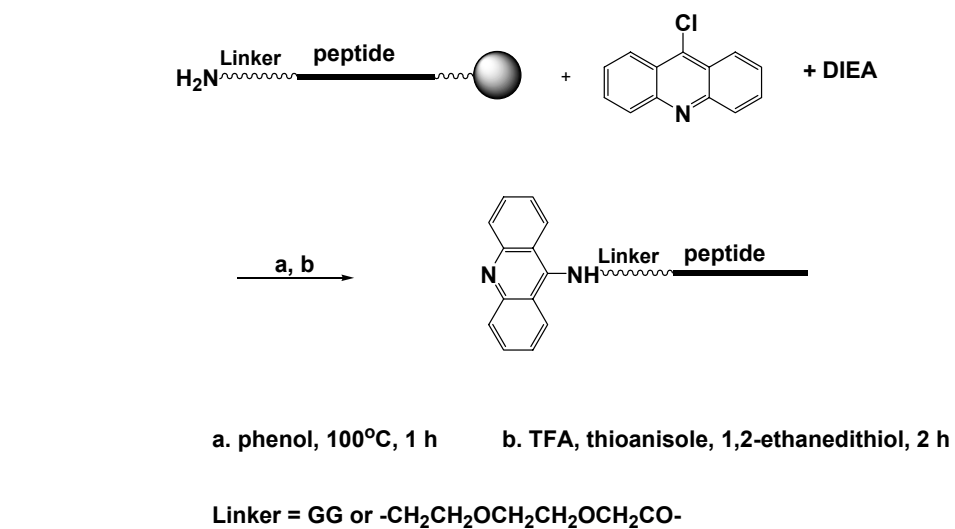
The RNA–protein interface of the N_{22}/boxB complex is dominated by electrostatic interactions and hydrophobic contacts.[23] The five arginines and two lysines of N_{22} create a positively charged surface on one face of the α-helix that interacts with the negatively charged phosphodiester backbone of the boxB RNA. Hydrophobic interactions are also important for boxB recognition. Ala-3 and Trp-18 are involved in crucial hydrophobic interactions. In addition, the roles of arginine and lysine residues are not restricted to ionic interactions, and some of the aliphatic portions of these side chains also contact the RNA bases or sugars.

The N/boxB complex is an ideal system for testing binding enhancement by introducing a new binding mode, e.g., intercalation. In this chapter, we report our efforts using acridine-peptide conjugates to enhance N peptide binding affinity against boxB RNA targets and demonstrate the recent successful application of these principles to create novel RNA-binding peptides, with particular emphasis on the advances achieved using intercalating agent acridine.

Materials and Methods

Synthesis of peptides and acridine-peptide conjugates.

Crude peptides with or without linkers were constructed by automated solid phase peptide synthesis using Fmoc protected monomers (ABI) on an Applied Biosystems 432A peptide synthesizer. A series of full length λ N$_{22}$ and truncated λ N$_{11}$ acridine-peptide conjugates were manually synthesized on the resin using the chemistry outlined in Scheme 1. Two types of flexible linkers, a Gly-Gly linker **1** and an ethylene glycol linker **2**, were used. Crude peptides and crude acridine-peptides conjugates were deprotected and cleaved from the resin and purified by reverse-phase HPLC on a C18 column. The purity of the products was checked by analytical HPLC and their identity confirmed by MALDI–TOF mass spectrometry. Concentrations of peptide stocks were determined by UV absorption at 280 nm for free peptides containing either tryptophan or tyrosine, or at 412/434 nm for acridine-peptide conjugates using extinction coefficients ($\varepsilon=1.32\times10^4$cm^{-1}M^{-1} at 412 nm, and 1.12×10^4cm^{-1}M^{-1} at 434 nm) determined for a water soluble acridine derivative 3-acridin-9-ylamino-propanol synthesized from 9-chloroacridine and 3-amino-1-propanol as described. [12]

a. phenol, 100°C, 1 h   b. TFA, thioanisole, 1,2-ethanedithiol, 2 h

Linker = GG or -CH₂CH₂OCH₂CH₂OCH₂CO-

3-acridin-9-ylamino-propanol

λmax (nm)	ε (cm⁻¹M⁻¹)
220	2.84×10^4
266	6.99×10^4
412	1.32×10^4
434	1.12×10^4

Scheme 1. Synthesis and quantification of acridine-peptide conjugates.

Synthesis of 2AP-labeled RNA oligomers.

Crude RNA oligomers with fluorescent 2- aminopurine (2AP) label substituted for adenine at the 2nd, 3rd, and 4th base positions of the pentaloop (denoted 2AP-2, 2AP-3, and 2AP-4, respectively) were constructed by automated synthesis using 2-aminopurine-TOM-CEphosphoramidite (Glen Research, Sterling, VA). Oligomers were deprotected and purified by 20% urea-PAGE. Purified oligomers were desalted on NAP-25 column and quantified by UV absorption at 260 nm.

Steady-State Fluorescence Measurements.

Steady-state fluorescence measurements were made on a Shimadzu RF-5301PC Spectrofluorophotometer as described. [21,26] Aliquots of concentrated stocks of free peptide or acridine-peptide conjugates were added stepwise to a stirred solution of 2AP-containing RNA maintained at various temperatures in a series of buffer conditions. Fluorescence signal of 2AP was monitored at 370 nm with excitation at 310 nm. Dissociation constants (K_d) were analyzed using the computer program Dynafit. [27]

Band Shift Analysis.

Free *boxB* RNA or complexes of peptide-RNA were preformed in NMR buffer (50 mM NaCl, 10 mM Phosphate, 0.5 mM EDTA, pH 6), and diluted by TBE buffer, before loading to 20% non-denaturing PAGE gel maintained at 10-15 °C. Free RNA and complex bands were visualized by UV shadowing.

One Dimensional NMR Spectroscopy.

Unlabeled 15mer boxB RNA 5'GCCCUGAAAAAGGGC3' (bases in the loop are underlined) was synthesized by in vitro transcription using T7 RNA polymerase.[28] The RNA was purified by 20% urea-PAGE and desalted on a NAP25 column. Purified RNA oligomer was resuspended in NMR buffer (50 mM NaCl, 10 mM Phosphate, 0.5 mM EDTA, pH 6, 90:10 H2O/D2O). Spectra were collected on a Varian INOVA 600 MHz NMR spectrometer at 25 °C. Spectrum of free boxB RNA was collected first; titration of concentrated peptide or acridine peptide conjugate into the boxB RNA was monitored by inspecting the imino proton region of RNA.

CD Spectroscopy.

Spectra were taken on an Aviv 62 DS CD spectrometer at 20°C. The samples contained 5 µM RNA and 6 µM peptide in 10 mM potassium phosphate buffer (pH 7.9). The spectra of the bound peptides were determined by subtracting the spectra for free RNA and excess free peptide from that of the complex.

Results

We envisioned that in the acridine peptide conjugate–RNA complex, the high affinity λ N peptide would bind to the major groove of the boxB RNA with the same structure as in the wild-type complex. [23,24] This would deliver the acridine moiety proximal to the bottom part of the RNA stem (Figure 1) where it can either intercalate into the stem or simply associate with the RNA extrahelically.
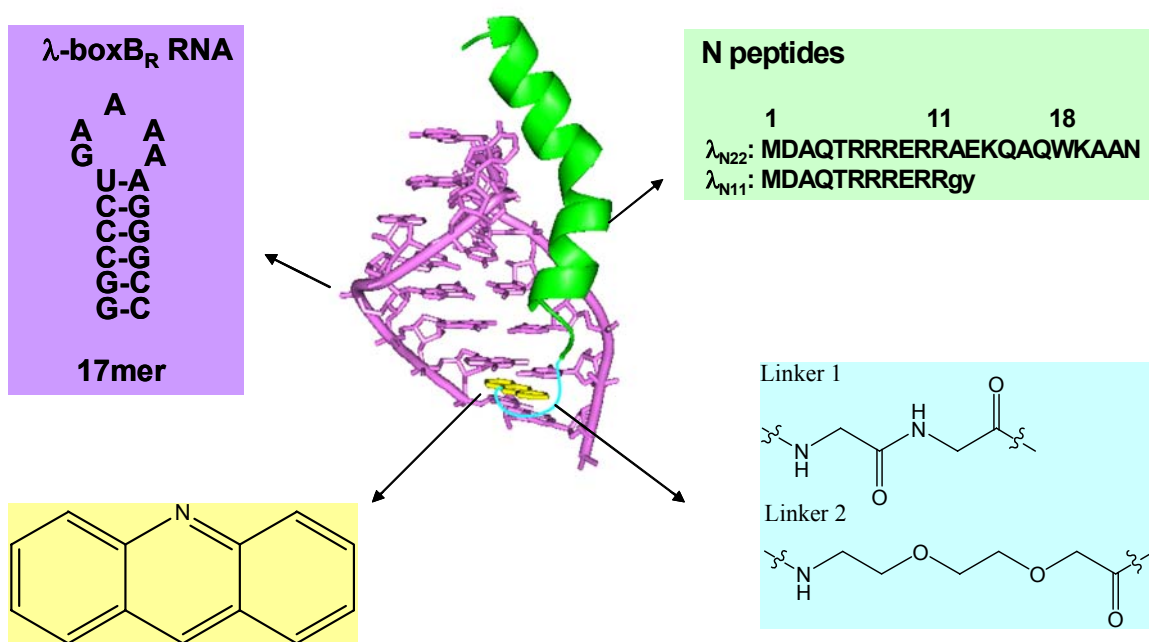
Figure 1. Design of acridine-peptide conjugate binding to boxB RNA. **A**. Sequences and secondary stem-loop structures of boxB RNA from λ bacteriophages used in this study. **B**. Sequences of λ N_{22} full length and truncated λ N_{11} peptides. The 11-mer was tagged with gy for quantification purposes. **C.** Structure of acridine. **D**. Structure of linkers.

K_d values.

Dissociation constants (K_d) were determined by monitoring the change in fluorescence of 2-aminopurine (2AP) incorporated at variable positions within the loops of the RNA hairpins. This base was found to be extremely useful as a probe of nucleic acid structure and dynamics,[29] and RNA-peptide interactions.[26] The fluorescence intensity of 2AP is very sensitive to local environment, decreasing when stacked and increasing when solvent exposed.[30] In most cases, peptide binding can be detected by either a fluorescence increase or decrease of more than 20% from the starting value .[21,31]

Table 1. Dissociation constants (K_d) for peptides and acridine-peptide conjugates against boxB_R RNA[a]

|	λ boxB_R RNA targets (17mers)		
	λ-2AP-2	λ-2AP-3	λ-2AP-4
	A A A G A U A C G C G C G G C g c	A A A G A U A C G C G C G G C g c	A A A G A U A C G C G C G G C g c
Peptide	K_d (nM)	K_d (nM)	K_d (nM)
λ_N11	1126	1480	1000
Acr-Link1-λ_N11	3945	2304	3504
Acr-Link2-λ_N11	3082	2200	3245
λ_N22	1.9	1.0	1.2
Acr-Link1-λ_N22	2.7	1.3	2.0
Acr-Link2-λ_N22	0.025[b]	0.019[b]	0.015[b]

[a] Binding constants were determined for standard condition: 20 ˚C; 50 mM KOAc, 20 mM Tris.OAc, pH 7.5. Individual isotherms were fit to a one-step reaction with less than 10% error. Hairpin base positions substituted with 2AP are underlined. Acr- refers to acridine moiety.

[b] Extrapolated from high salt measurement to 50 mM K$^+$ (see Table 2 and Figure 2).

Table 1 shows the K_d values determined at 50 mM K^+/20 mM Tris for acridine-λ N_{11}, and acridine-λ N_{22} peptide conjugates, compared to those of WT peptides, against boxB RNA targets from λ phages. Binding affinities for truncated λ N_{11} peptide-acridine conjugates constructed with either a Gly-Gly linker **1** or an ethylene glycol linker **2** are similar to the WT λ N_{11} peptide. However, the binding affinities are enhanced when the longer and more flexible linker **2** is used on the full length peptide. The enhancement depends on the actual construct and RNA targets and the conjugates have increased binding specificity. For example, an acridine-full length peptide conjugate has 0.015 nM affinity to the 17mer boxB RNA with 2AP-4 labeling when extrapolated from high salt measurement to 50 mM K^+ (Table 2, Figure 2), corresponding to a 80-fold enhancement over the wild full length peptide (1.2 nM). The longer and more flexible linker **2** seems to provide affinity enhancement. Full length acridine peptide conjugates constructed with the shorter and more rigid linker **1** exhibited no binding enhancement, therefore further investigations were focused on the linker **2**.

Salt dependence of binding.

Generally protein-nucleic acid interactions become weaker as the salt concentration is increased due to the electrostatic nature of binding. This principle also extends to our system. Since N peptide is arginine-rich, the electrostatic interactions between positively charged peptide side chains and negatively charged RNA phosphate backbone are critical for binding.[23] These electrostatic interactions are mitigated by increasing concentration of salt. Within a range of cation concentrations, plots of log (K_d) vs. log [M^+] give a linear relationship. The slope of these plots corresponds to the number of counterions that are released upon peptide binding. In general, the slope is related to the net charge of the peptide and ranges from 2.5 to 5 for the N peptide sequences that have been tested.[32] Salt dependence of acridine full length N peptide conjugate–boxB RNA complex was analyzed using the 17mer RNA with 2AP-4 labeling (Table 2, Figure 2). Most of the isothermal curves show biphasic transitions with the second transition starting after peptide conjugate reaches 1:1 stoichiometry with the RNA target. The K_d values were fit to the first transition. Indeed, higher salt decreased binding affinity of acridine-peptide conjugate to boxB RNA as observed for the WT peptide. The salt dependence graph showed linear relationship at higher salt and leveled off at low salt (Figure 2) as observed previously.[33,34] The slope has a value of 2.9, slightly higher than the value of 2.4 for wild N peptide, indicating roughly three cations are released upon binding.
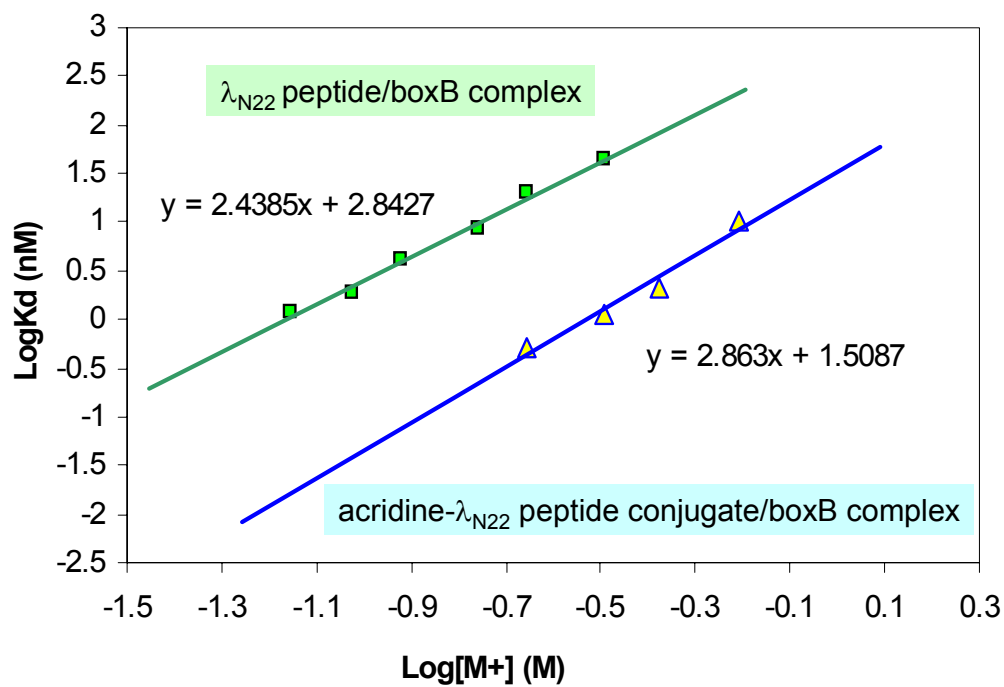
Figure 2. Salt dependence of λ full length peptide (diamond) and acridine-full length peptide conjugate (triangle) binding to 17mer λ boxB RNA (2AP-4). Measurements were made at 20 °C.

The K_d value extrapolated by salt dependence curve to 50 mM K^+, 20mM Tris is 0.015 nM. A roughly 80–fold binding enhancement was observed across a broad range of ionic concentrations. Since the conjugate has slightly higher slope of salt dependence than the free peptide, when extrapolated to lower salt, the enhancement is even higher.

Table 2. Salt dependence*

λ boxB$_R$ RNA target λ-2AP-4(17mer)

```
        A
   A    A
   G    A
   U  A
   C  G
   C  G
   C  G
   G  C
   g  c
```

Total monovalent cation concentration (mM)

Peptide	70	95	120	175	220	320	420	620
					K_d (nM)			
λ$_{N22}$	1.2	1.9	4.2	8.5	19.8	43.5		
Acr-Link2-λ$_{N22}$					0.5	1.1	2.1	10

* Total monovalent cation concentration is sum of K^+ and Tris. Binding constants were determined for standard condition: 20 ˚C; 50-600 mM KOAc, 20 mM Tris.OAc, pH 7.5. Individual isotherms were fit to a one-step reaction with less than 10% error. Hairpin base positions substituted with 2AP are underlined. Acr- refers to acridine moiety.

CD Spectra.

The CD spectra collected for both acridine-full length peptide conjugate and free full length peptide binding to boxB RNA (15mer) are shown in Figure 3. Neither the peptide nor the acridine-peptide conjugate shows any appreciable structure in the absence of RNA. The difference spectra of the two complexes indicate that both peptides fold into α helices when bound to the RNA. Although globally similar, the two complexes display some differences in regions indicative of peptide folding (200-225nm) and RNA folding (260-300nm). The acridine conjugate/RNA complex is less α-helical.
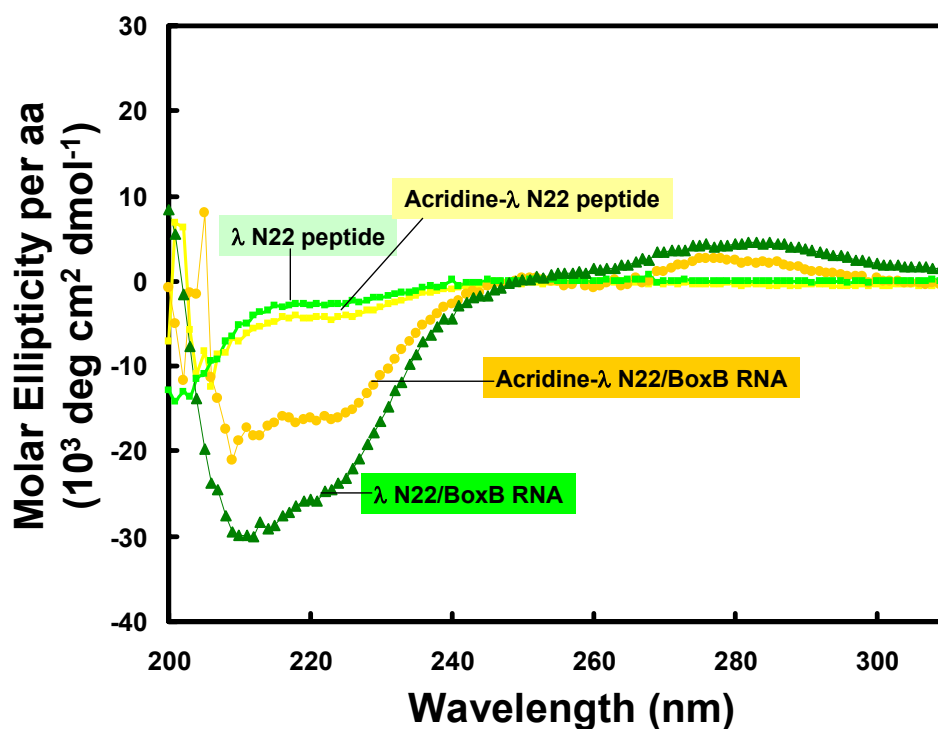

Figure 3. CD spectra of free peptides and peptide/RNA complexes.

NMR Spectroscopy.

NMR experiments can provide useful structural information on the boxB RNA (15mer) and N peptide, especially the structural change upon complexation. In particular, imino protons from RNA base pairs in the stem and Trp18 indole NH proton have unique signals that are easy to monitor. NMR spectra were collected in order to investigate the binding mode between acridine-peptide conjugate and the boxB RNA and compare to that of the free peptide and RNA (Figure 4).

The free 15mer boxB RNA (Figure 4A, a) has 5 base pairs in the stem, the 1D imino proton NMR spectrum showed 3 peaks (13.25 ppm, 12.65 ppm, and 12.55 ppm) corresponding to the imino protons of G12, G13, and G14 from the three middle CG base pairs. The terminal GC pair is not observable due to fraying, and the U imino proton from loop closing UA pair is also missing presumably due to the flexibility of the unstructured pentaloop in the free RNA. Binding by either full length λ N_{22} peptide or truncated amino-terminal λ N_{11} induces similar RNA structural changes, evidenced by shifting of the three CG imino protons to higher field (Figure 4A, b and e). Beside these changes, there were additional imino protons at 13.5 ppm corresponding to U5 of the UA pair, and a broader peak at 10.7 ppm attributable to the G6 imino proton from the sheared GA base pair characteristic of GNRA tetraloop type of folding. This indicates that the loop has been stabilized by peptide binding. In addition, the Trp18 indole NH proton in the full length peptide has a large shift from 10 ppm (Figure 4A, d) in the free peptide to 9.2 ppm in the complex (Figure 4A, e), this is due to the stacking interaction between Trp18 and the RNA loop. [24]
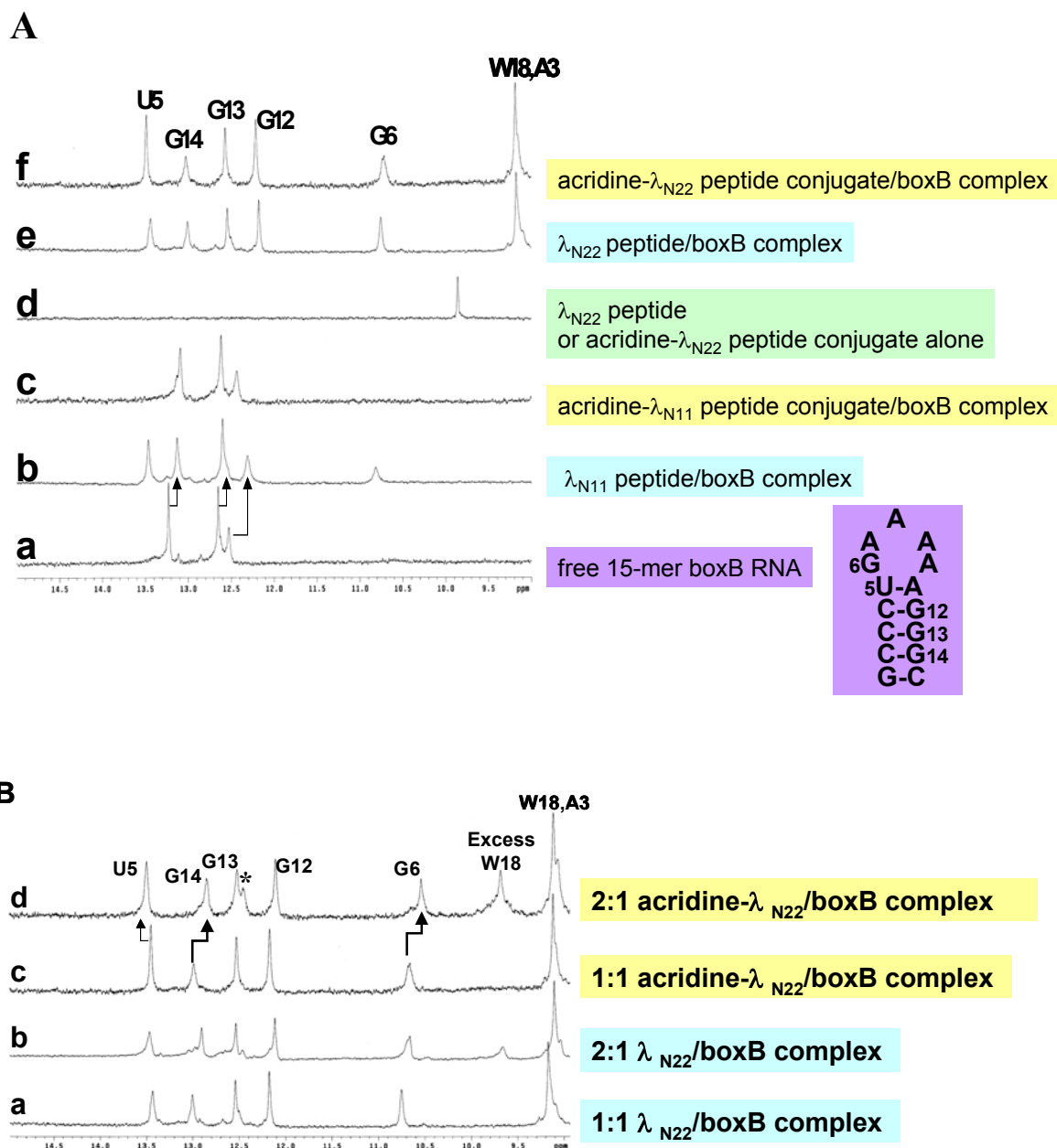
Figure 4. One dimensional NMR spectra for peptide, boxB RNA, and complexes. **A**. Comparison of of λN_{11} and λN_{22} peptide/boxB 1:1 ratio complexes. **B**. Comparison of 1:1 peptide/boxB complex and 2:1 peptide/boxB complex. The asterisk indicates the additional peak.

In comparison, the spectrum for acridine-λ N_{11} conjugate complexed to boxB RNA (Figure 4A, c) did not show these characteristic signal changes as observed in λ N_{11} peptide/RNA complex, except for broadening of the original three peaks. This indicates that the acridine-λ N_{11} conjugate does not bind the RNA in the same mode as the λ N_{11} peptide. Since acridine compound 2-acridin-9-ylamino-propanol binds to boxB RNA with about 15 μM affinity, similar to the affinity between λ N_{11} peptide and boxB RNA, it is possible that the two modules compete for the RNA binding site, and acridine moiety can associate with the RNA non-specifically, disrupting the specific complex structure between λ N_{11} peptide and boxB RNA.

The spectrum for 1:1 acridine-full length peptide conjugate/boxB RNA complex (Figure 4A, f) showed very similar imino proton signals for RNA component and characteristic Trp18 indole proton shift as seen in the spectrum for λ N_{22} peptide complex. Within the acridine-full length peptide conjugate, it seems likely that the peptide portion binds the RNA in the same mode as in the WT λ N_{22} peptide complex. This should deliver the acridine moiety to the lower part of boxB RNA stem.

Since acridine-full length peptide conjugate showed biphasic binding to boxB RNA, we further investigated the potential second binding site on boxB RNA for the conjugate. Over titration of the conjugate produced a 2:1 ratio of conjugate:RNA complex. Comparison of the spectra for 1:1 and 2:1 complexes of both WT peptide and acridine-peptide conjugate revealed interesting discrepancies (Figure 4B). For the case of 2:1 acridine-λ N_{22} peptide conjugate/boxB complex, the imino protons for U5 moved further down field, and G6 and G14 moved up field, respectively; furthermore, there is an

additional peak at 12.4 ppm. These features suggest a specific second binding site for the acridine-peptide conjugate.

Band Shift Analysis.

Observations from NMR experiments are consistent with the band shift analysis shown in Figure 5. Because the λ N$_{11}$ peptide binds the boxB RNA specifically, a band shift can be observed (lane 2 of Figure 5A); while there is only smearing in the acridine-λ N$_{11}$ peptide conjugate with boxB RNA (lane 3 of Figure 5A) indicating non-specific interaction, possibly random intercalation of acridine either in the stem or in the loop region which compete with the binding of peptide portion.
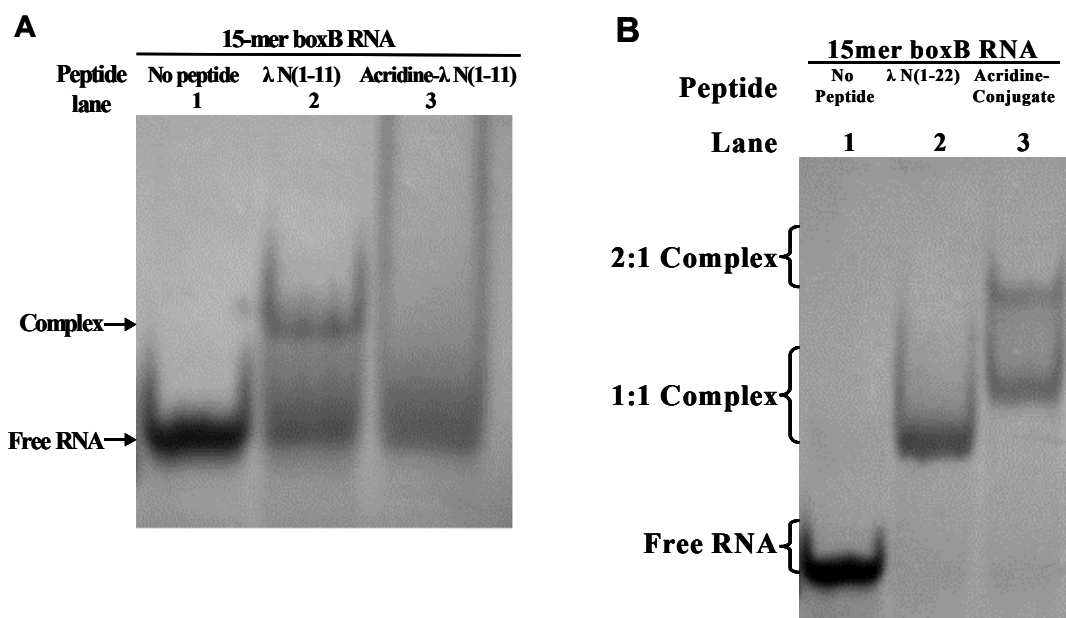
Figure 5. Gel shift analysis of peptide/boxB complexes for WT peptides and acridine-peptide conjugates. **A**. λ N$_{11}$/RNA and acridine-λ N$_{11}$/RNA complexes. **B**. λ N$_{22}$/RNA and acridine-λ N$_{22}$/RNA complexes.

The acridine full length conjugate, however, not only showed a 1:1 complex to boxB RNA 15-mer (lane 3 of Figure 5B) as in peptide-RNA complex (lane 2 of Figure 5B), but a higher stoichiometric complex as well. These results suggest that there is a second binding site for the acridine-peptide conjugate. This may explain the profile of the fluorescence titration curves, which sometimes show biphasic transitions with the second transition apparently following equimolar titration.

Discussions

To achieve higher RNA binding affinities, the design of a ligand consisting of two moieties connected by a linker is a promising approach. [7,8] Higher binding affinities could result from a favorable entropic factor compared with the binding of the two monomeric counterparts. Sometimes the length and nature of the linker is critical. The molecule can be a dimeric form of a known binder [35] or can consist of two distinct moieties that bind to RNA in different modes, e.g., groove binding and intercalation. [10-12] To apply this idea to the N peptide system, a series of λ N peptide-based peptide-acridine conjugates were designed and synthesized. Their binding affinities to boxB RNAs with stem-loop hairpin structures were determined by steady-state fluorescence measurements.

Since acridine compound 2-acridine-9-ylamino-propanol binds boxB RNA with 15 μM affinity, theoretically, it can provide up to three orders of magnitude enhancement when linked to another binding module. The full enhancement may never be realized in real systems. From comparing salt dependence of binding, acridine-peptide conjugation showed 80-fold binding enhancement over a broad range of salt concentrations. This result is similar to binding enhancement observed for a neo-acridine construct over neomycin B alone binding to RRE RNA. [12] The same construct showed 3-fold tighter binding to a related P22 boxB RNA target. Compared to λ boxB, this RNA has orientation of one base pair in the stem flipped, and a cytidine in the loop instead of

adenine (Figure 6). Both λ_{N22} and P22$_{N21}$ full-length peptides discriminate strongly between their cognate RNA hairpins and other boxB targets.[32] The acridine-peptide conjugates showed enhanced binding specificity between these two similar RNA structures (Table 3). Therefore, the actual enhancement by acridine conjugation depends on the targets involved.

Table 3. Dissociation constants (K_d) for peptides and acridine-peptide conjugates against λ boxB and P22 boxB$_R$ RNA[a]

	boxB$_R$ RNA targets (17mers)	
	λ-2AP-4	P22-2AP-4
	A A A G A U A C G C G C G G C g c	C A A G A U A C G G C C G G C g c
Peptide	K_d (nM)	K_d (nM)
λ_{N22}	1.2	257
Acr-Link2-λ_{N22}	0.015[b]	80

[a] Binding constants were determined for standard condition: 20 ˚C; 50 mM KOAc, 20 mM Tris.OAc, pH 7.5. Individual isotherms were fit to a one-step reaction with less than 10% error. Hairpin base positions substituted with 2AP are underlined. Acr- refers to acridine moiety.

[b] Extrapolated from high salt measurement (Figure 2)

A

B

```
        A                            C
    A       A                    A       A
    G       A                    G       A
      U-A                          U-A
      C-G                          C-G
      C-G                          G-C
      C-G                          C-G
      G-C                          G-C
      G-C                          G-C
```
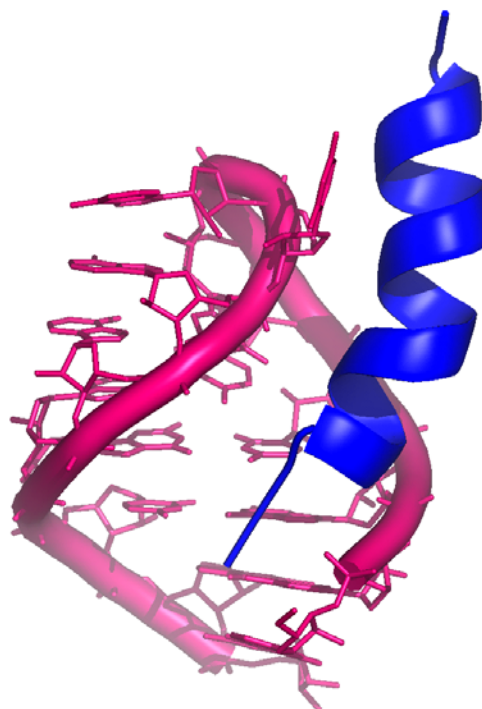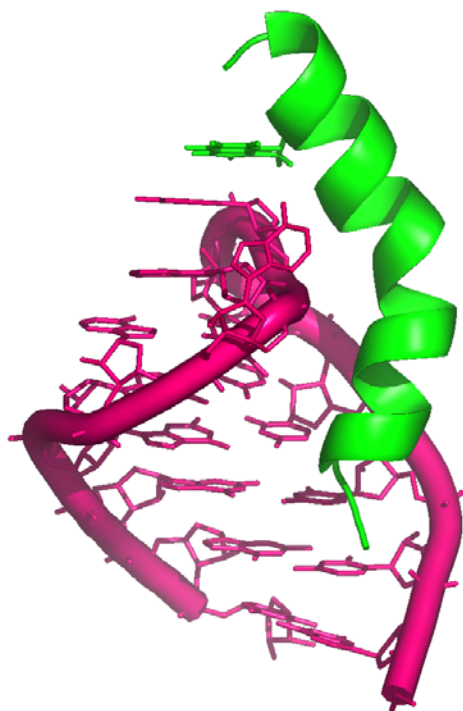
λ-boxB 17mer   **P22-boxB 17mer**

Figure 6. Comparison of λ boxB and P22 boxB RNA. **(A)** Structural models of phage λ and P22 N peptide-boxB hairpin complexes viewed from the major groove. In all complexes the N peptide is shown as a ribbon; **(B)** Schematic representation of λ and P22 boxB hairpin stem-loop structure.

Homologous hydrophobic interactions occur between the boxB hairpin stem and conserved alanine residues within the N peptide amino-terminal module. Distinct hydrophobic interactions appear between the boxB hairpin loops and carboxy-terminal modules of the N peptides. In λ, a tryptophan residue stacks on the boxB loop; in P22, non-polar alanine and isoleucine residues interact with an extruded pyrimidine. In phage λ and P22, the bound pentaloops adopt stable GNRA tetraloop folds by extruding either loop base 4 (4-out) in the λ complex, or loop base 3 (3-out) in the P22 complex.

In contrast, conjugation of acridine to the λ N_{11} peptide did not improve binding. In fact, the conjugate bound about 3-fold weaker than the λ N_{11} peptide alone. The NMR and gel shift data showed that the conjugate does not form a specific complex with boxB RNA (Figure 4A and 5A). Presumably this is due to the fact that the acridine moiety has similar affinity to the boxB RNA compared to the the λ N_{11} peptide, and acridine binding is non-specific. Therefore the acridine moiety competes the binding sites of RNA structure with peptide moiety. This result suggests that when designing a ligand with multiple binding modules, it is important to avoid a non-specific binding module that can compete multiple sites intended for the other modules. A module with higher affinity for one site will help anchor the entire ligand. The specific binding peptide module of the full length acridine-peptide conjugate has much higher affinity for the λboxB RNA, the peptide portion maintains all the native contacts with RNA (Figure 4A), and the acridine moiety can be specifically delivered to the intended site.

Previous structural studies on DNA intercalated with acridine related compounds have consistently shown intercalation site at alternating base steps, particularly at the terminal base steps.[36] This is likely due to the fact that terminal base steps have much greater conformational freedom than internal base steps, which facilitates intercalation. Recent studies on DACA, an antitumor acridine derivatized with variable functional substitutions at position 4, showed that they intercalate into DNA with side chains lying in the major groove.[37-39] However, the crystal structure of an acridine-tetraarginine conjugate with the peptide attached to the position 9 showed intercalation of acridine into the minor groove of central AA/TT base step of a long DNA fragment, leaving the peptide lying in the minor groove.[40]

In these cases, it appeared that specific hydrogen bonding interactions between the lateral chain of acridine and DNA may affect how the acridine moiety intercalates into nucleic acids. Interestingly, a recent study also showed that in peptide-acridine conjugates, the point of attachment on acridine affects the conjugates' affinity for nucleic acid target.[41]

Our acridine-N peptide conjugates feature the peptide portion attaching to the acridine at position 9, similar to the acridine-tetraarginine case. To determine the precise binding mode of acridine in these conjugates is challenging at this time. There may be multiple reasons for this. If the acridine moiety derivatized at position 9 with a peptide

prefers intercalation from minor groove as in the case of acridine-tetraarginine,[40] our construct may preclude intercalation, because full length N peptide binds boxB RNA in the major groove with high affinity. In the case of the acridine-λ N_{11} conjugate, the acridine may be able to intercalate into minor groove, but this intercalation is probably non-site-specific. A second reason is that once the N peptide (especially the full length) binds to boxB RNA peptide induces RNA folding. There are extensive interactions between peptide and RNA stem, an intercalated acridine moiety in the stem may disrupt specific interactions between RNA and peptide. However, the acridine may still be able to intercalate at the terminal base step, but since the imino proton from the terminal base pair is not observable, this possibility has not been confirmed.

Similar to the neo-acridine case[12] where there was a second binding site, our fluorescence, NMR, and gel shift data also suggested a second specific binding site for acridine-N peptide conjugate on boxB RNA. The nature of the second site is not known, but it may be in either the major or minor groove. Changes of the imino proton chemical shifts can be observed throughout the RNA structure (Figure 4B), suggesting that the interactions may be extensive.

Other approaches of enhancing RNA binding affinity may also be available. N peptides bind boxB RNA in an α-helical conformation, but have little helical structure when free in solution.[25] Stabilizing the helical form of such peptides is expected to favor RNA binding by virtue of pre-organization. For example, C^{α}-substituted amino acids has long been recognized as a means of introducing local conformational restriction in

peptides.[42] Another approach to stabilize the α-helical form of peptides is through the incorporation of covalent linkages between constituent amino acid side chains. It was found that substantial helix stabilization was achieved when the linkage was placed between the *i* and *i+4* residue in the peptide backbone by ring closing metathesis.[43-46] Therefore, cyclic helical N peptides, wherein ring closing metathesis is used to incorporate a carbon-carbon tether between appropriate residue side chains, is expected to provide binding enhancement. This approach provides a convenient way to enhance the affinity of known modules and facilitates the discovery of powerful new hybrid ligands with new functionalities.

Conclusion

Designing ligands that consist of multiple modules for binding RNA targets can be a powerful approach to achieve high affinity and specificity. A series of full-length or truncated acridine-peptide conjugates based on the N peptide of bacteriophage λ were designed and synthesized. Their binding affinities to a series of *Box B* RNAs with stem-loop hairpin structures were determined using steady-state fluorescence measurements. Thermodynamic parameters have been characterized for these complexes using salt dependence. An 80-fold binding enhancement of acridine-peptide conjugate with *Box B* RNA was exhibited across a range of salt concentrations. NMR, gel shift, and CD data also indicated the N peptide-acridine conjugate/*Box B* RNA complex formation.

References

1. Klug, A. Zinc finger peptides for the regulation of gene expression. *J Mol Biol* **293**, 215-8 (1999).

2. Segal, D. J. & Barbas, C. F., 3rd. Design of novel sequence-specific DNA-binding proteins. *Curr Opin Chem Biol* **4**, 34-9 (2000).

3. Wolfe, S. A., Nekludova, L. & Pabo, C. O. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* **29**, 183-212 (2000).

4. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E. & Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468-71 (1998).

5. Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg Med Chem* **9**, 2215-35 (2001).

6. Fechter, E. J. & Dervan, P. B. Allosteric inhibition of protein--DNA complexes by polyamide--intercalator conjugates. *J Am Chem Soc* **125**, 8476-85 (2003).

7. Michael, K. & Tor, Y. Designing novel RNA binders. *Chem. Eur. J.* **4**, 2091-2098 (1998).

8. Cheng, A. C., Calabro, V. & Frankel, A. D. Design of RNA-binding proteins and ligands. *Curr Opin Struct Biol* **11**, 478-84 (2001).

9. Li, S. & Roberts, R. W. A novel strategy for in vitro selection of Peptide-drug conjugates. *Chem Biol* **10**, 233-9 (2003).

10. Hamy, F. et al. A new class of HIV-1 Tat antagonist acting through Tat-TAR inhibition. *Biochemistry* **37**, 5086-95 (1998).

11. Wilson, W. D. et al. Bulged-Base nucleic acids as potential targets for antiviral drug action. *New J. Chem.* **18**, 419-23 (1994).

12. Kirk, S. R., Luedtke, N. W. & Tor, Y. Neomycin-acridine conjugate: A potent inhibitor of Rev-RRE binding. *J Am Chem Soc* **122**, 980-981 (2000).

13. Tung, C. H., Wei, Z., Leibowitz, M. J. & Stein, S. Design of peptide-acridine mimics of ribonuclease activity. *Proc Natl Acad Sci U S A* **89**, 7114-8 (1992).

14. Fkyerat, A., Demeunynck, M., Constant, J.-F., Michon, P. & Lhomme, J. A New Class of Artificial Nucleases That Recognize and Cleave Apurinic Sites in DNA with Great Selectivity and Efficiency. *J Am Chem Soc* **115**, 9952-9959 (1993).

15. Kuzuya, A., Mizoguchi, R., Morisawa, F., Machida, K. & Komiyama, M. Metal ion-induced site-selective RNA hydrolysis by use of acridine-bearing oligonucleotide as cofactor. *J Am Chem Soc* **124**, 6887-94 (2002).

16. Greenblatt, J., Nodwell, J. R. & Mason, S. W. Transcriptional antitermination. *Nature* **364**, 401-6 (1993).

17. Das, A. Control of transcription termination by RNA-binding proteins. *Annu Rev Biochem* **62**, 893-930 (1993).

18. Salstrom, J. S. & Szybalski, W. Coliphage lambdanutL-: a unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. *J Mol Biol* **124**, 195-221 (1978).

19. Olson, E. R., Tomich, C. S. & Friedman, D. I. The nusA recognition site. Alteration in its sequence or position relative to upstream translation interferes with the action of the N antitermination function of phage lambda. *J Mol Biol* **180**, 1053-63 (1984).

20. Lazinski, D., Grzadzielska, E. & Das, A. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59**, 207-18 (1989).

21. Barrick, J. E., Takahashi, T. T., Ren, J., Xia, T. & Roberts, R. W. Large libraries reveal diverse solutions to an RNA recognition problem. *Proc Natl Acad Sci U S A* **98**, 12374-8 (2001).

22. Heus, H. A. & Pardi, A. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253**, 191-4 (1991).

23. Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289-99 (1998).

24. Scharpf, M. et al. Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur J Biochem* **267**, 2397-408 (2000).

25. Su, L. et al. RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda. *Biochemistry* **36**, 12722-32 (1997).

26. Lacourciere, K. A., Stivers, J. T. & Marino, J. P. Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy. *Biochemistry* **39**, 5630-41 (2000).

27. Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal Biochem* **237**, 260-73 (1996).

28. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* **15**, 8783-98 (1987).

29.   Menger, M., Eckstein, F. & Porschke, D. Dynamics of the RNA hairpin GNRA

       tetraloop. *Biochemistry* **39**, 4500-7 (2000).

30.   Rachofsky, E. L., Osman, R. & Ross, J. B. Probing structure and dynamics of

       DNA with 2-aminopurine: effects of local environment on fluorescence.

       Biochemistry **40**, 946-56 (2001).

31.   Austin, R. J., Xia, T., Ren, J., Takahashi, T. T. & Roberts, R. W. Designed

       Arginine-Rich RNA-Binding Peptides with Picomolar Affinity. *J Am Chem Soc*

       124, 10966-7 (2002).

32.   Austin, R. J., Xia, T., Ren, J., Takahashi, T. T. & Roberts, R. W. Differential

       Modes of Recognition in N Peptide-BoxB Complexes. *Biochemistry* **42**, 14957-67

       (2003).

33.   Record, M. T., Jr., Lohman, M. L. & De Haseth, P. Ion effects on ligand-nucleic

       acid interactions. *J Mol Biol* **107**, 145-58 (1976).

34.   Record, M. T., Jr., Ha, J. H. & Fisher, M. A. Analysis of equilibrium and kinetic

       measurements to determine thermodynamic origins of stability and specificity and

       mechanism of formation of site-specific complexes between proteins and helical

       DNA. *Methods Enzymol* **208**, 291-343 (1991).

35.   Campisi, D. M., Calabro, V. & Frankel, A. D. Structure-based design of a dimeric

       RNA-peptide complex. *Embo J* **20**, 178-86 (2001).

36.   Williams, L. D., Egli, M., Gao, Q. & Rich, A. (eds.) *Nucleic Acids* (Adenine

       Press, Schenectady, NY, 1992).

37.   Todd, A. K. et al. Major groove binding and 'DNA-induced' fit in the intercalation

       of a derivative of the mixed topoisomerase I/II poison N-(2-

(dimethylamino)ethyl)acridine-4-carboxamide (DACA) into DNA: X-ray structure complexed to d(CG(5-BrU)ACG)2 at 1.3-A resolution. *J Med Chem* **42**, 536-40 (1999).

38. Adams, A., Guss, J. M., Denny, W. A. & Wakelin, L. P. Crystal structure of 9-amino-N-[2-(4-morpholinyl)ethyl]-4-acridinecarboxamide bound to d(CGTACG)2: implications for structure-activity relationships of acridinecarboxamide topoisomerase poisons. *Nucleic Acids Res* **30**, 719-25 (2002).

39. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A. & Wakelin, L. P. Crystal structure of the topoisomerase II poison 9-amino-[N-(2-dimethylamino)ethyl]acridine-4-carboxamide bound to the DNA hexanucleotide d(CGTACG)2. *Biochemistry* **38**, 9221-33 (1999).

40. Malinina, L., Soler-Lopez, M., Aymami, J. & Subirana, J. A. Intercalation of an Acridine-Peptide Drug in an AA/TT Base Step in the Crystal Structure of [d(CGCGAATTCGCG)](2) with Six Duplexes and Seven Mg(2+) Ions in the Asymmetric Unit. *Biochemistry* **41**, 9341-9348 (2002).

41. Carlson, C. B. & Beal, P. A. Point of attachment and sequence of immobilized Peptide-acridine conjugates control affinity for nucleic acids. *J Am Chem Soc* **124**, 8510-1 (2002).

42. Liff, M. I., Kopple, K. D., Tian, Z. & Roeske, R. W. Effects of C alpha-methyl substitution on the conformation of linear GnRH antagonist analogs. *Int J Pept Protein Res* **43**, 471-6 (1994).

43.   Miller, S. J. & Grubbs, R. H. Synthesis of Conformationally Restricted Amino-Acids and Peptides Employing Olefin Metathesis. *J Am Chem Soc* **117**, 5855-5856 (1995).

44.   Lynn, D. M., Mohr, B. & Grubbs, R. H. Living ring-opening metathesis polymerization in water. *J Am Chem Soc* **120**, 1627-1628 (1998).

45.   Blackwell, H. E. & Grubbs, R. H. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Ange Chem-Int Ed* **37**, 3281-3284 (1998).

46.   Schafmeister, C. E., Po, J. & Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J Am Chem Soc* **122**, 5891-5892 (2000).

Chapter 2

Creating Synthetic "Blanks" in the Genetic Code Using A Novel Class of Aminoacyl Adenylate Mimics as tRNA Synthetase Inhibitors

Abstract

Efforts to reengineer the protein synthesis apparatus should expand our fundamental understanding of the central macromolecular synthesis machinery inside all cells. The UAG stop codon has often been referred to as a "blank" in the genetic code because it can be suppressed by an appropriate tRNA. We would like to eliminate the activity of specifically chosen aminoacyl tRNA synthetase (aaRS), creating our own blanks in the codes. We generated various high-affinity inhibitors of the synthetases—aminoacyl sulfamide to create synthetic blanks in our translation extracts. The gaps were filled with chemically aminoacylated orthogonal tRNAs. This unnatural strategy enables peptides and proteins to be constructed containing a single novel residue at a specific location. Biocytin, a biotin derivative of lysine, can be incorporated into globin, efficiently enriched by using aminoacyl sulfamide to inhibit natural aminoacylated tRNA and adding orthogonal tRNA coupled with biocytin.

Introduction

The protein synthesis machinery translates genetic information into folded, functional polypeptides, but is constrained by the universal genetic code to using the 20 naturally occurring amino acids as building blocks. Each of these 20 amino acids is escorted into the machine by transfer RNA that recognizes a specific codon in a message RNA template (Figure 1). Rewriting genetic code requires reassigning the codons to correspond to new amino acids.
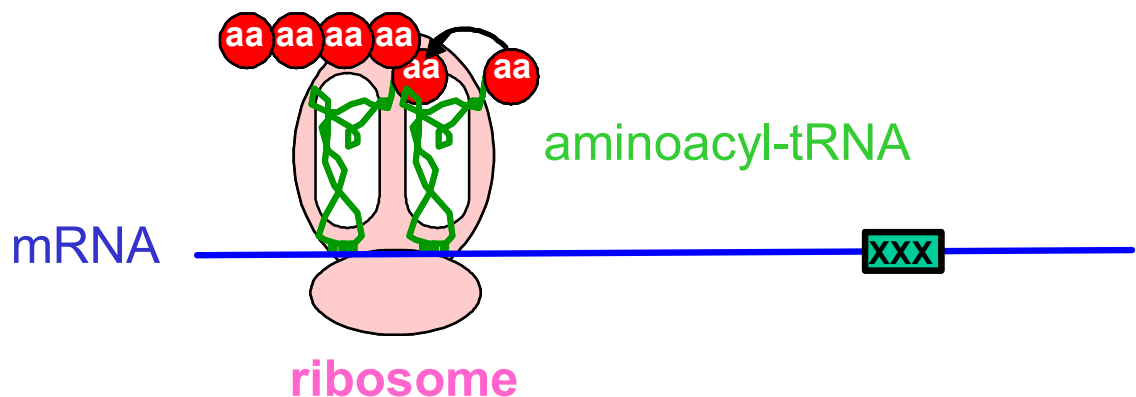
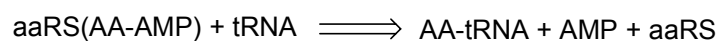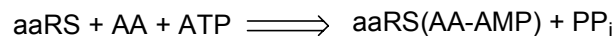

Figure 1. The protein synthesis machinery - ribosome.

The incorporation of unnatural amino acids into peptides and proteins has been pursued intensively over a couple of decades. Several unnatural amino acids have been incorporated into dipeptides through the use of chemically misacylated tRNAs in the

Hecht lab[1]. The UAG stop codon has often been referred to as a "blank" in the genetic code because it can be suppressed by an appropriate suppressor tRNA. Schultz and coworkers demonstrated the insertion of unnatural amino acids into proteins by nonsense UAG suppression[2,3]. While the original description of the method utilized an *E. coli* translation extract, the technique has been applied to eukaryotic translation extracts such as rabbit reticulocyte lysate[4-7] or *in vivo* in *Xenopus* oocytes[8-10]. This unnatural strategy enables peptides and proteins to be constructed containing a single novel residue at a specific location, amenable to the insertion of various entities into proteins including: 1) affinity tags, 2) spectroscopic probes, and 3) various amino acid derivatives for physical organic and mechanistic studies of protein function.
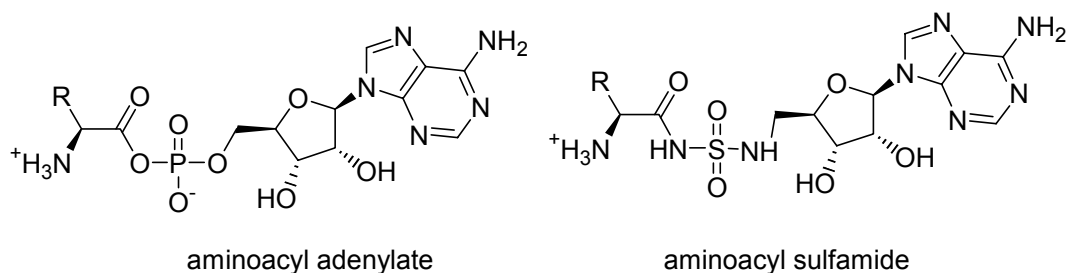
We would like to create our own blanks in the code by eliminating the activity of specific aaRS as we choose. All aaRSs act by using ATP to create an activated form of an amino acid – a mixed carbon-phosphorus anhydride termed an aminoacyl adenylate (Figure 2). Aminoacyl sulfamides are stable structural mimics of aminoacyl adenylates and generally have submicromolar K_i values for their corresponding synthetases[11]. We generated various high-affinity inhibitors of the synthetases to create synthetic blanks in our translation extracts. Aminoacyl tRNA synthetase inhibitors are combined with translation extracts and chemically aminoacylated orthogonal tRNAs to reprogram specific codons as desired. The genetic code will be determined by the chemically acylated tRNAs used, rather than by the natural aminoacyl-tRNA synthetases, the gatekeepers of the genetic code.

A

$$aaRS + AA + ATP \Longrightarrow aaRS(AA\text{-}AMP) + PP_i$$

$$aaRS(AA\text{-}AMP) + tRNA \Longrightarrow AA\text{-}tRNA + AMP + aaRS$$

B

aminoacyl adenylate   aminoacyl sulfamide

**Figure 2**. (A) Aminoacyl tRNA synthetases (aaRSs) specifically catalyze the aminoacylation of tRNA in a two step reaction. (B) The tRNA synthetases typically catalyze the formation of tightly bound aminoacyl-adenylates in their first reaction step. The analogues of these reaction intermediates (aminoacyl sulfamides) have been synthesized and evaluated for the ability to inhibit aminoacylation.

Application of non-sense and sense-codon suppression in protein selection has been examined in our group[12,13]. First, we reported that biocytin, a biotin derivative of lysine, can be inserted into an mRNA-protein fusion molecule through amber stop codon suppression. We also demonstrated that templates containing the codon corresponding to the biocytin tRNA (a UAG stop codon) can be enriched by iterative cycles of selection against a streptavidin agarose matrix. The universal genetic code links the 20 naturally occurring amino acids to the 61 sense codons. We have developed a selection methodology to investigate whether the unnatural amino acid biocytin could be incorporated into an mRNA display library at sense codons. In these experiments we probed a single randomized NNN codon with a library of 16 orthogonal, biocytin-acylated tRNAs. *In vitro* selection for efficient incorporation of the unnatural amino acid resulted in templates containing the GUA codon at the randomized position. This sense suppression occurs via Watson-Crick pairing with similar efficiency to UAG-mediated nonsense suppression. These experiments suggest that sense codon suppression is a viable means to expand the chemical and functional diversity of the genetic code. We envision that these aaRS inhibitors should facilitate the construction of novel proteins and selection of unnatural amino acids incorporated by sense codon suppression using the mRNA display approach (Figure 3).

mRNA display provides a strategy to make up to 10^{13} unique peptides that fuse with their own genes via a covalent linkage. Compared to other *in vitro* selection methodologies, such as phage-display[14,15] and surface-display[16] libraries, mRNA display[17,18] libraries have several unique features, including the incorporation of unnatural residues, the chemical derivatization of libraries, and the feasibility of mutagenic PCR and DNA recombination.
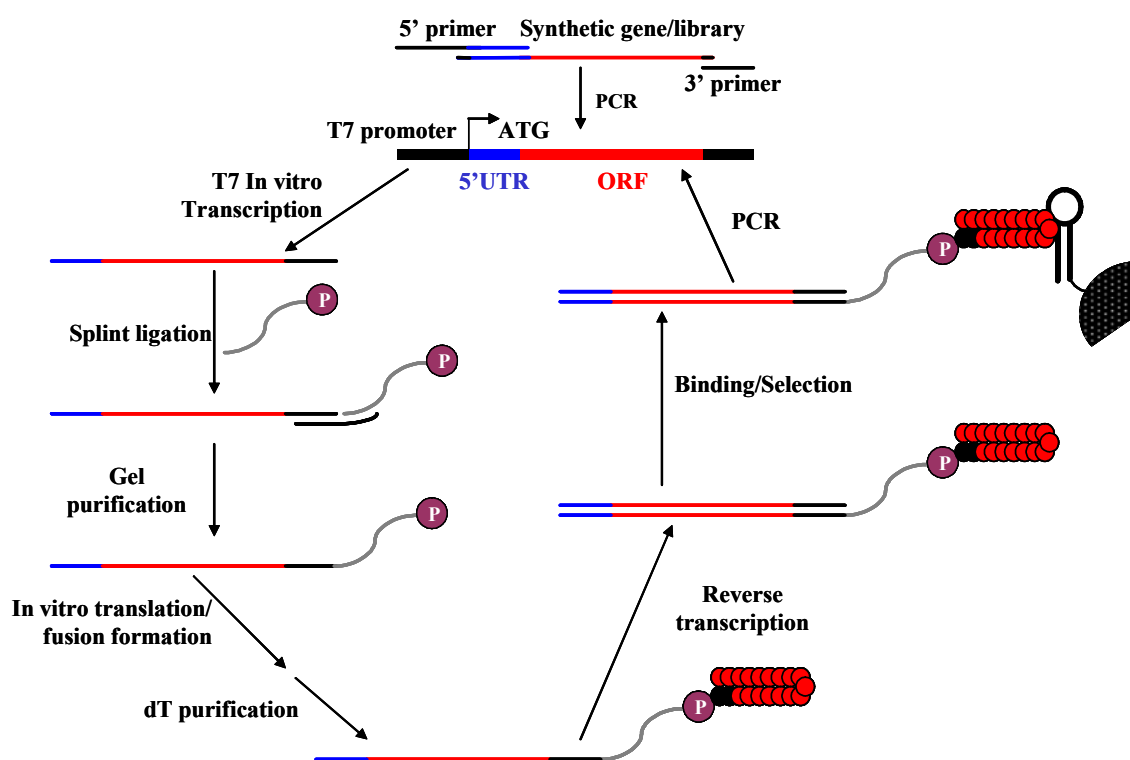
Figure 3. Selection cycle via mRNA display[17].

In affinity selection with mRNA display library, the synthetic DNA library is transcribed to make mRNA, which is modified with a 3'-puromycin. The engineered mRNA template is translated to prepare mRNA-peptide fusion. After a cDNA•mRNA duplex is generated by reverse transcription, the fusion library is subjected to affinity selection against targets of interest. The enriched library is amplified by PCR for the next cycle of selection.

Biocytin can be detected by ECL Plus Western Blotting Detection, and it binds to streptavidin (protein) coupled with Horseradish Peroxidase (HRP enzyme) which catalyzes a chemoluminence reaction.

THG73 tRNA
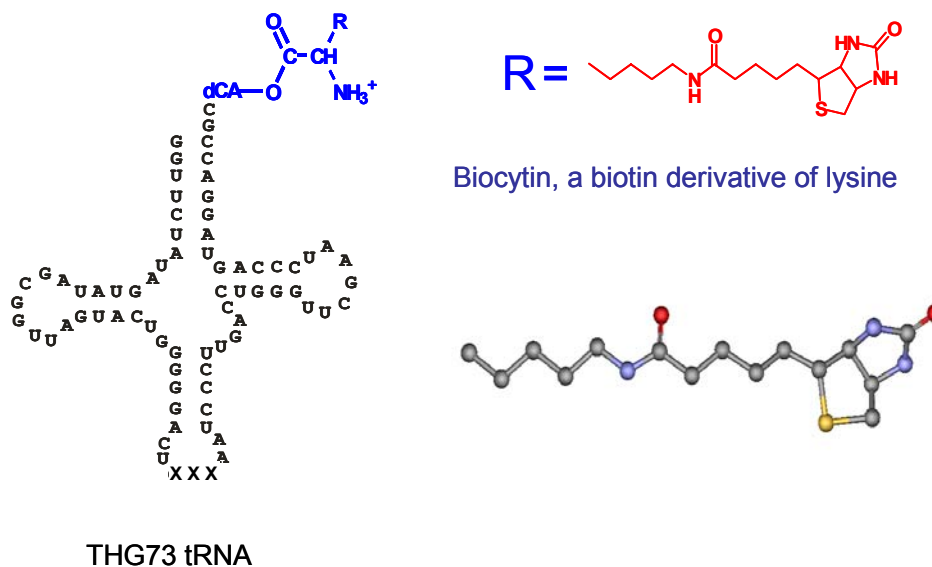
Figure 4. Biocytin charged suppressor tRNA and biocytin structure.

All the tRNAs we used are made from the pUC19-based plasmid harboring the gene for THG73, which was mutated at the tRNA anticodon position by Quikchange (Stratagene). To synthesize chemically aminoacylated tRNA with unnatural amino acid residue, the amino group of amino acid was protected by NVOC group first and the

carboxylic group was activated by acetonitrile. The activated amino acid was then coupled to dinucleotide pdCpA. Using T4-RNA ligase, dinucleotide pdCpA coupled with unnatural amino acid was ligated to the engineered truncated 73-nucleotide tRNA to form intact charged suppressor tRNA. Right before applying it to the translation system, NVOC protecting group was deprotected using UV irradiation (Scheme 1).
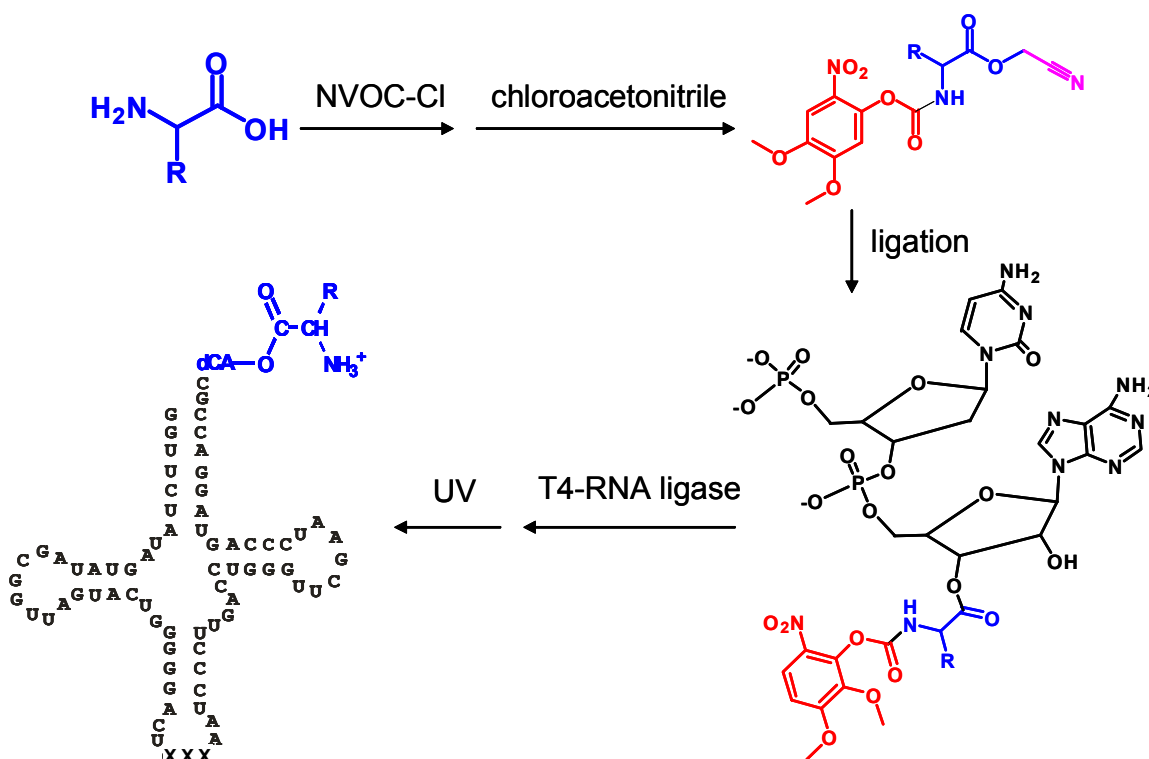
Scheme 1. Preparation of amino acylated THG73 tRNAs.

Since various high-affinity inhibitors of the synthetases suppress the formation of natural amino acylated tRNA and create synthetic blanks in our translation extracts, the selection rounds will be shortened for the discovery of novel ligands with functionalities beyond those provided by the 20 naturally occurring amino acids.

Materials and Methods

General Information. ^{1}H and ^{13}C NMR spectra were recorded on a UNITY INOVA instrument (Varian, Inc.) operating at 500 MHz using D_2O or DMSO-*d6* as the solvent. ^{1}H NMR data are reported as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br s, broad singlet; dd, doublet of doublets. High-resolution mass spectra (FAB) were recorded on a JMS-600H double-focusing, high-resolution, magnetic sector mass spectrometer (Mass Spectrometry Laboratory, Division of Chemistry and Chemical Engineering, California Institute of Technology). Column chromatography was carried out on silica gel (40-63 μm, EM Science). All reagents were of highest available commercial quality and were used without further purification. Triphenolphosphine, phthalimide, 2',3'-O-isopropylideneadenosine, hydrazine, benzyl alcohol, diethyl azodicarboxylate, chlorosulfonyl isocyanate, Pd 10% on activated carbon wet Degussa type, carbonyldiimidazole and diazabicycloundecene were purchased from Aldrich. Sodium azide was purchased from Sigma Chemical Co. Boc-Ala-OH, Boc-Val-OH, Boc-Leu-OH, Boc-ILe-OH, Boc-Phe-OH, Boc-Trp-OH, Boc-Tyr(2-Br-Z)-OH, Boc-Met-OH, Boc-Pro-OH and Boc-Gly-OH were purchased from Novabiochem.

Rabbit reticulocyte lysate was purchased from Novagen. Rabbit globin mRNA was obtained from Sigma Chemical Co. Ras mRNA was prepared by using two DNA primers complementary to the 5'- and 3'- ends of the coding region for H-Ras (pProEX HTb

vector, a kind gift from Dafna Bar-Sagi)[19] to amplify the gene using PCR.   mRNA was produced by T7 runoff transcription[20] of the H-Ras DNA in the presence of RNAsecure (Ambion) followed by gel purification via denaturing urea-PAGE and 'crush and soak' RNA isolation.   L-[^{35}S]methionine (1,175 Ci/mmol) was purchased from NEN Life Science Products.   GF/A glass microfiber filters were from Whatman.   Scintillation counting was carried out using a Beckman LS-6500 liquid scintillation counter.

General Procedure for Preparation of Aminoacyl-Sulfamides. 5'-Amino-2', 3'-O-isopropylidene -5'-deoxyadenosine was prepared from commercially available 2',3'-O-isopropylideneadenosine via the procedure disclosed previously[21].   It was then converted   to   5'-deoxy-2',   **3'**-O-isopropylideneadenosine-5'-N-[(phenylmethoxy) carbonyl]-sulfamide by treatment with N-carbobenzyloxysulfamoyl chloride and a base such as triethylamine in dichloromethane at 0 °C to ambient temperature. CBZ protected sulfamide was reduced with Pd 10% on activated carbon wet Degussa type and hydrogen gas in EtOH/MeOH to afford 5'-deoxy -2', **3'**-O-isopropylideneadenosine-5'-N-sulfamide, which was further reacted with BOC protected amino acids catalyzed by carbonyldiimidazole and base diazabicycloundecene in acetonitrile.   The obtained 5'-deoxy-2',3'**-**O-isopropylideneadenosine-5'-N-(N-tertbutoxycarbonyl-aminoacyl) sulfamides were fully deprotected according to standard procedures and purified by trituration[22].

IC$_{50}$ determination. Translation reactions containing [^{35}S]Met were mixed in batch on ice and added in aliquots to microcentrifuge tubes containing an appropriate amount of aminoacyl-sulfamide (2.5 μL of different concentrations). Typically, a 25 μl translation mixture consisted of 1.0 μL of 2.5 M KCl, 0.5 μL of 25 mM MgOAc, 2.0 μL of 12.5X translation mixture without methionine, (25 mM dithiothreitol (DTT), 250 mM HEPES (pH 7.6), 100 mM creatine phosphate, and 312.5 μM of 19 amino acids, except methionine) (Novagen), nuclease-free water, 2.0 μL (6.1 μCi) of [^{35}S]Met (1175 Ci/mmol), 10 μL of Red Nova nuclease-treated lysate (Novagen), and 5 μL of 0.05 μg/μL globin mRNA (Sigma). Inhibitor, lysate preparation (including all components except template), and globin mRNA were mixed simultaneously and incubated at 30 ºC for 60 min. Biocytin-tRNAs were added for translation restoration test. Then 2 μL of each reaction was combined with 8 μL of tricine loading buffer (80 mM Tris-Cl (pH 6.8), 200 mM DTT, 24% (v/v) glycerol, 8% sodium dodecyl sulfate (SDS), and 0.02 % (w/v) Coomassie blue G-250), heated to 90 ºC for 5 min, and applied entirely to a 4% stacking portion of a 16% tricine-SDS-polyacrylamide gel containing 20% (v/v) glycerol (30 mA for 1h, 30 min)[23]. Gels were fixed in 10% acetic acid (v/v) and 50% (v/v) methanol, dried, exposed overnight on a PhosphorImager screen, and analyzed using a Storm PhosphorImager (Molecular Dynamics). Inhibition of translation to a variety mRNA templates was tested in a similar condition.

Preparation of biocytin-acylated tRNAs. The pUC19-based plasmid harboring the gene for THG73 was mutated at the tRNA anticodon position by Quikchange (Stratagene). Resulting clones were verified by DNA sequencing before synthesizing individual tRNAs by *in vitro* transcription with T7 polymerase from *Fok I*-linearized plasmids. Transcribed tRNAs were then gel-purified and desalted by ethanol precipitation. Purified tRNAs were ligated to NVOC protected biocytin-dCA with T4 RNA ligase (New England Biolabs). Reaction mixtures were extracted in an equal volume of phenol:$CHCl_3$:isoamyl alcohol (25:24:1, pH 5.2, Fisherbiotech), and precipitated with 3.0 volumes of cold ethanol (-20 °C). After drying, the pellets were resuspended in 1.0 mM sodium acetate, pH 5.2, and adjusted to 4.0 mg/mL and 1.0 mg/mL for each biocytin-acylated tRNA. Biocytin-tRNA concentrations were determined by UV absorbance at 260 nm. Before adding to translation reactions, biocytin-tRNAs were deprotected by a xenon lamp outfitted with a 315 nm cut-off filter for 5 min to remove the NVOC group.

Restoration of Translation Probed with Streptavidin-capture of Biocytin-containing Globin. A portion of streptavidin-agarose (0.75 mL) [Pierce, 50% slurry (v/v)] was washed 3 times with a buffer (1X PBS containing 0.1% Triton X-100) and resuspended in 0.75 mL of buffer. To 100 μL of this suspension, 10 μL of the translation

reaction and up to 0.7 mL of buffer were added. The samples were rotated at 4 ºC for 1.5 h and washed with buffer until the cpm of [^{35}S]Met were <500 in the wash. Immobilized [^{35}S]Met-biocytin-containing-globin was eluted by adding tricine loading buffer (80 mM Tris-Cl (pH 6.8), 200 mM DTT, 24% (v/v) glycerol, 8% sodium dodecyl sulfate (SDS), and 0.02 % (w/v) Coomassie blue G-250) (50 µL for each sample) directly to the streptavidin-agarose beads with heating at 90 ºC for 10 min. The amount of immobilized [^{35}S]Met-biocytin containing globin was determined by scintillation counting of the eluent (Figure 10).

Restoration of Translation Probed by Western Blot Analysis. A portion of each translation reaction (2 µL) was combined with 8 µL of tricine loading buffer, heated to 90 ºC for 5 min, and applied entirely to a 4% stacking portion of a 15% SDS-polyacrylamide gel. The protein was transferred to nitrocellulose membrane in 1X transfer buffer w/ 10% MeOH (10X transfer buffer: Glycine 290g, Tris base 58g to 2L H$_2$O) using Miniprotean II wet transfer apparatus (Bio-Rad). The nitrocellulose membrane was blocked w/ 1% BSA in 1X PBS +0.1% Tween-20 for 1h and probed with Streptavidin-HRP conjugate (1:2000) for 1h. The membrane was washed four times with 30ml 1X PBS-T and visualized using ECL plus Western Blotting Detection System (Amersham Biosciences). Luminescence was then detected using Hyperfilm^{TM} (Amersham) with various exposure times.

Results

We generated various high-affinity inhibitors of the tRNA synthetases—aminoacyl sulfamides to create synthetic blanks in our translation extracts. We demonstrated that we can create and reprogram synthetic "blanks" in the genetic code. Aminoacyl tRNA synthetase inhibitors were combined with translation extracts and chemically aminoacylated orthogonal tRNAs to reprogram specific codons as desired.

A series of aminoacyl sulfamides were synthesized starting from 2', 3'-O-isopropylidene adenosine[21,22] (Scheme 2). 5'-hydroxy to amino group transformation was achieved by reacting with phthalimide followed by reducation using hydrazine. The first step reaction was catalyzed by Diethyl azodicarboxylate $EtO_2CN=NCO_2Et$ and triphenylphosphine. 5' free amino group reacted with N-Benzyloxycarbonylsulfamoyl chloride to afford protected aminoacylsulfamide. Deprotection using Pd catalyzed hydrogenation gave adenosine-5'-N-sulfamide which was coupled with Boc-protected amino acids catalyzed by CDI (carbonyldiimidazole) and DBU (diazabicycloundecene). After Boc protecting group was deprotected using TFA condition, the residue was triturated with anhydrous ether to give aminoacyl sulfamide as a white solid. There are additional procedures according to different amino acids with more complex side chains.
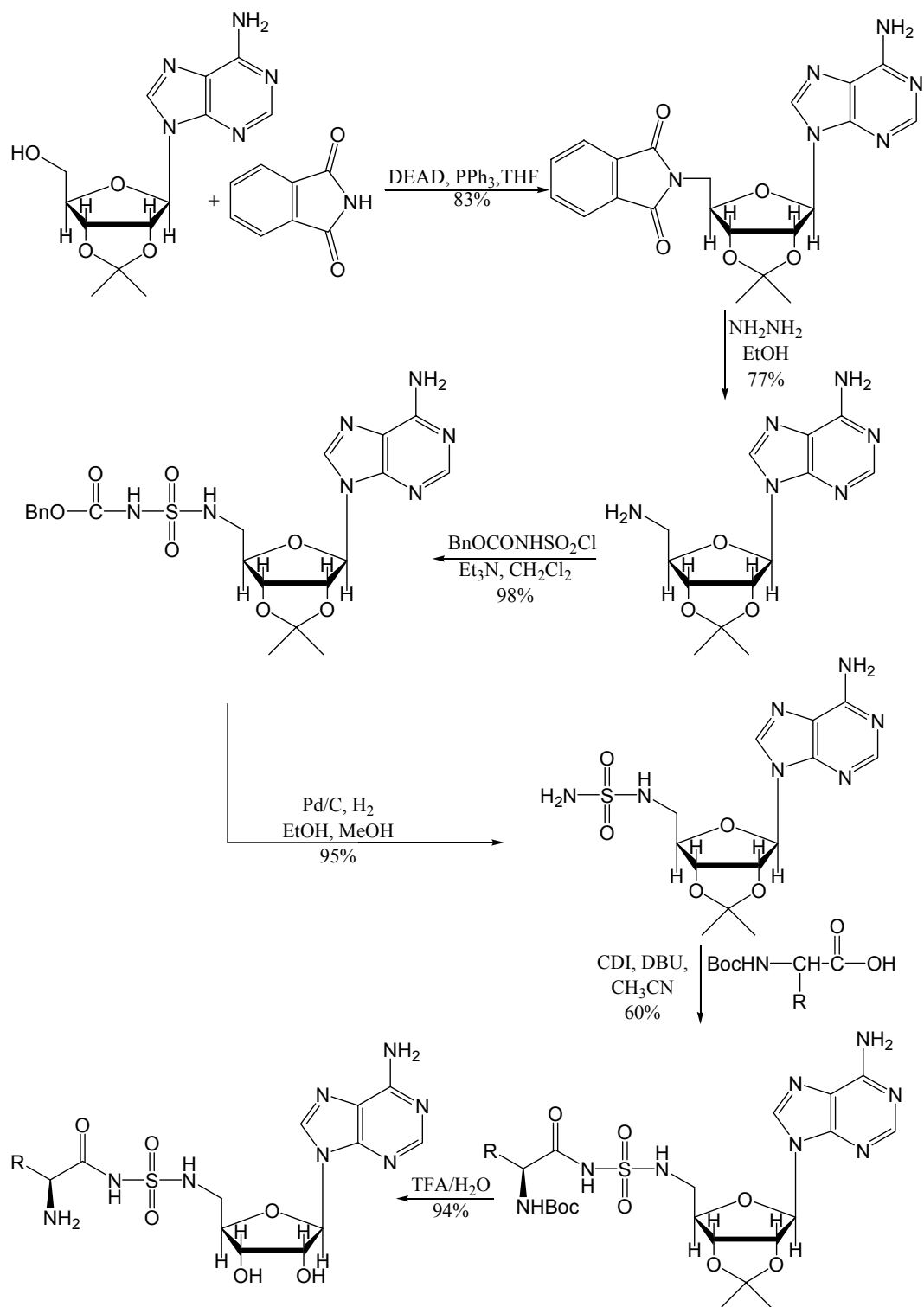
Scheme 2. Synthesis of aminoacyl sulfamides.

Addition of micromolar concentrations of synthetically constructed sulfamides allows us to specifically block the action of one or more synthetases as we choose. Inhibition is general to a variety of templates we tested, such as globin, Ubiquitin, and Ras mRNA templates (Figure 5). Clearly the potency of different aminoacyl sulfamides to inhibit protein translation varies according to the amino acid bearing and the mRNA template tested.
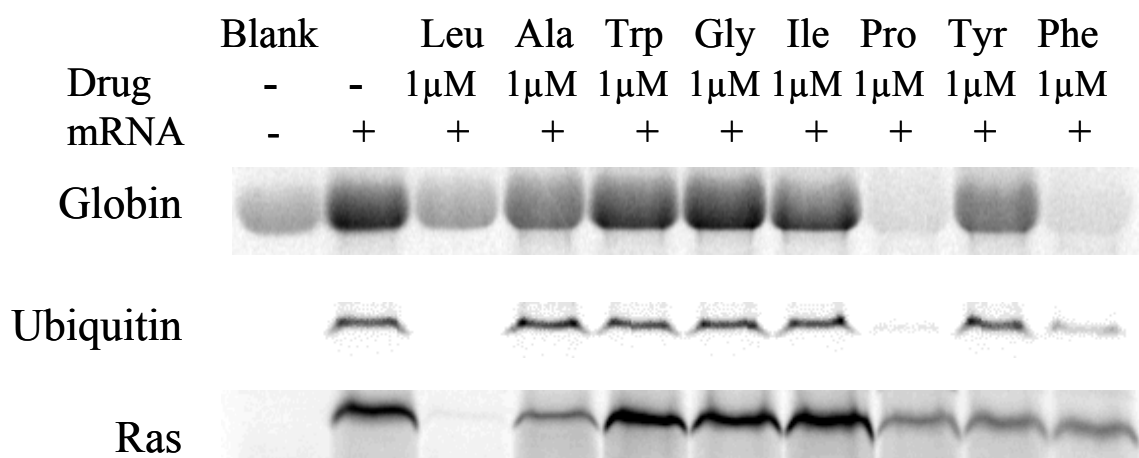
Figure 5. Inhibition of protein translation by aminoacyl-sulfamide. Tricine-SDS-PAGE analysis of [^{35}S]Met-protein translation reactions in the presence of 1 μM of aminoacyl-sulfamides.

We then measured the activity of each compound in a high dynamic-range IC_{50} potency assay using the rabbit reticulocyte protein synthesis system.[24] The various aminoacyl sulfamides inhibit globin translation with IC_{50} in the range of 22 nM to 28 μM (Figure 6). There appears to be a general correlation between the potency of the drugs and the amino acid abundance for a particular template. For example, Leu-sulfamide is a very potent inhibitor (IC_{50} 22 nM), whereas Ile-sulfamide is a poor inhibitor (IC_{50} 6.7 μM) for globin translation, and there are more positions coded for Leu (17/18) than Ile (3/1) in globin template. Presumably, the larger abundance of an amino acid in the template, the easier to inhibit the translation by shutting down the aminoacyl tRNA synthetase activity. All inhibitors are specific to corresponding aminoacyl tRNA synthetases due to the structural similarity to their corresponding aminoacyl adenylates.

A key issue is whether we can restore translation of globin by supplying aminoacyl tRNAs orthogonal to the codon blank created by the inhibitor. We tested the restoration of globin translation by biocytin-tRNAs coding for valine (Figure 7). We used an orthogonal tRNA charged with biocytin to eliminate it as a substrate for the cognate aminoacyl-tRNA synthetase and provide an easy handle for probing.

A

	1	2	3	4	5	6	7	8	9	10
Leu-Sulfamide	-	-	0.1 nM	1 nM	10 nM	100 nM	250 nM	500 nM	1 μM	10 μM
Globin mRNA	-	+	+	+	+	+	+	+	+	+

B

C

Aminoacyl Sulfamide	IC50 (μM)	Abundance of amino acids (α/β)
Met	0.221	1/2
Val	0.215	10/18
Ala	0.355	13/15
Leu	0.022	17/18
Ile	6.7	3/1
Tyr	1.0	3/3
Phe	0.056	8/8
Trp	28	1/2
Pro	0.036	7/4
Gly	15	9/11
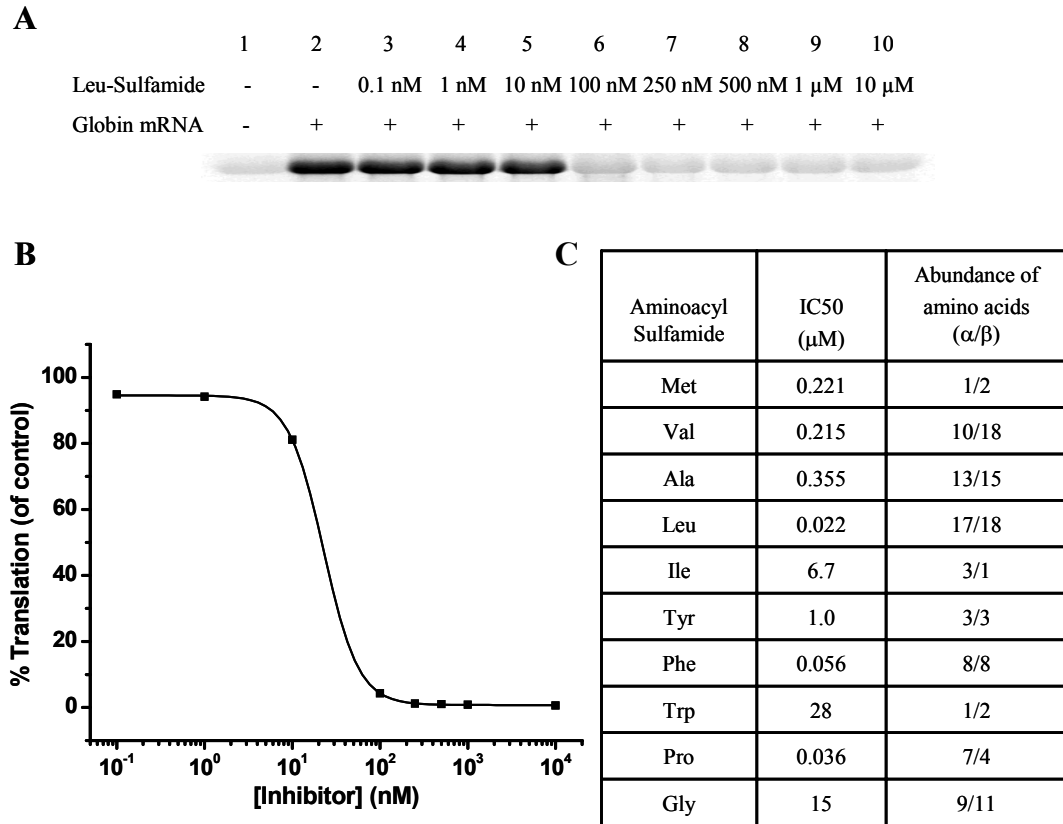
Figure 6. Inhibition of globin translation by aminoacyl sulfamide. (A) Tricine-SDS-PAGE analysis of [^{35}S]Met-globin translation reactions in the presence of Leu-Sulfamide: Lane 1, no template; lane 2, globin alone; lanes 3-10, concentrations of aminoacyl sulfamide from 0.1 nM to 10 μM. (B) Percent globin translation relative to the no drug control for Leu-Sulfamide from gel analysis of (A). (C) IC$_{50}$ values for various aminoacyl sulfamides.

The total globin translation level with the Val-sulfamide inhibitor and biocytin-tRNAs coding for valine is shown in Figure 7. At 1 μM of val-sulfamide, we were able to reduce the translation level to ~10% (lane 3). In the presence of biocytin-tRNAs, globin translation was restored to about 40% relative to no drug translation, depending on which biocytin-tRNAs were added (lane 4-6).

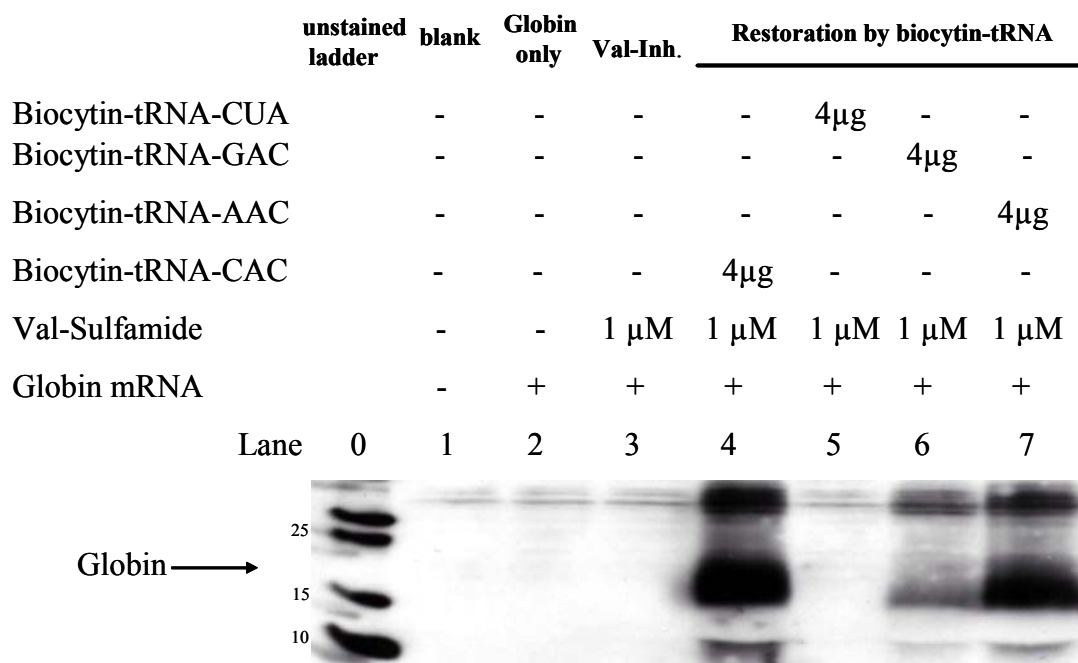|					
Biocytin-tRNA-CAC	-	-	-	4 μg	-	-
Biocytin-tRNA-GAC	-	-	-	-	4 μg	-
Biocytin-tRNA-AAC	-	-	-	-	-	4 μg
Val-Sulfamide	-	-	1 μM	1 μM	1 μM	1 μM
Globin mRNA	-	+	+	+	+	+
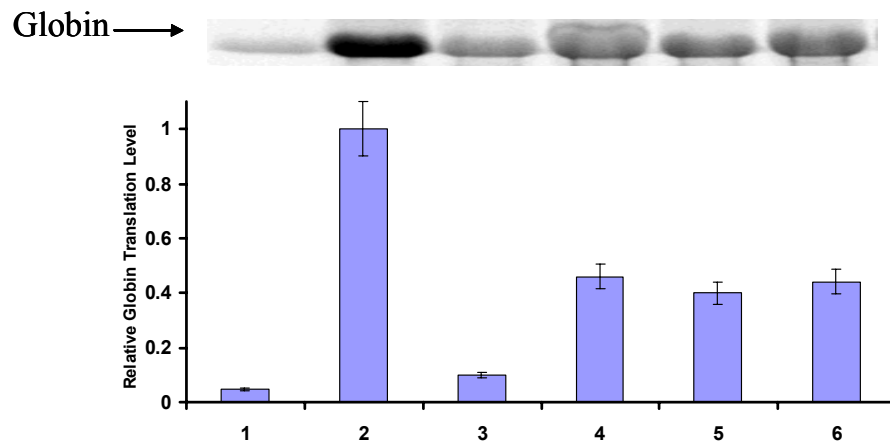
Figure 7. Globin translation restoration by biocytin-tRNA THG73 variant containing anticodon CAC, AAC or GAC in the presence of Val-sulfamide inhibitor: Lane 1, no template; lane 2, globin alone; lane 3, 1μM of val-sulfamide; lanes 4 - 6, biocytin-tRNA-CAC, GAC or AAC was added individually.

More important is the relative ratio of biocytin incorporation to total translation level. Here, specific biocytin incorporation was probed by Western blot (Figure 8). In α-globin,
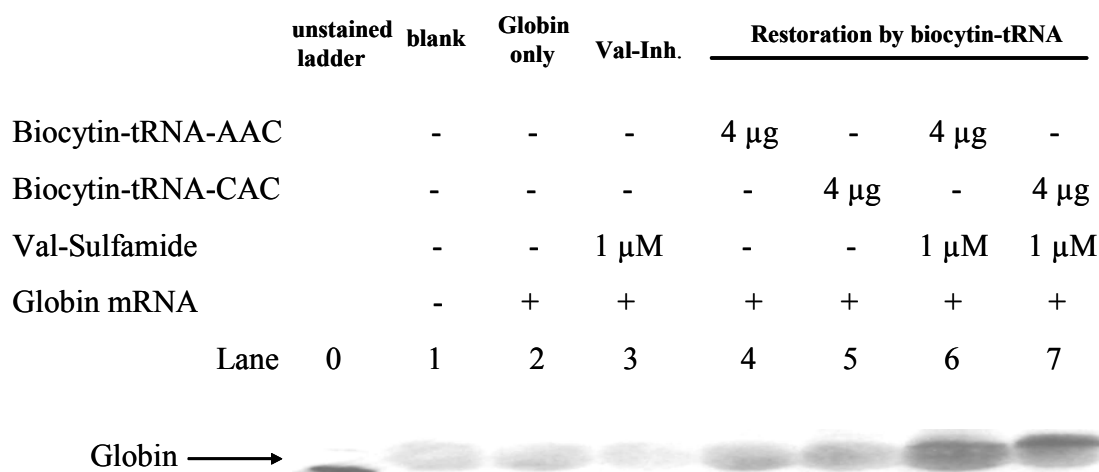
there are 10 valine codons, all are GUG (anticodon CAC); in β-globin, there are 18 valine codons, 12 of them are GUG, and 2 GUC (anticodon GAC) and 4 GUU (anticodon AAC), no GUA. The globin bands may represent heterogeneous populations of biocytin incorporation and the more abundant codon sites (lane 4, CAC) are over-represented than less abundant codons (lane 6 GAC, lane 7 AAC). Also when we use biocytin–tRNA with stop codon UAG (lane 5, anticodon CUA), there is no significant biocytin incorporation, and this demonstrates that the Val sulfamide inhibitor shows specificity toward the corresponding Val synthetase due to the structural similarity.
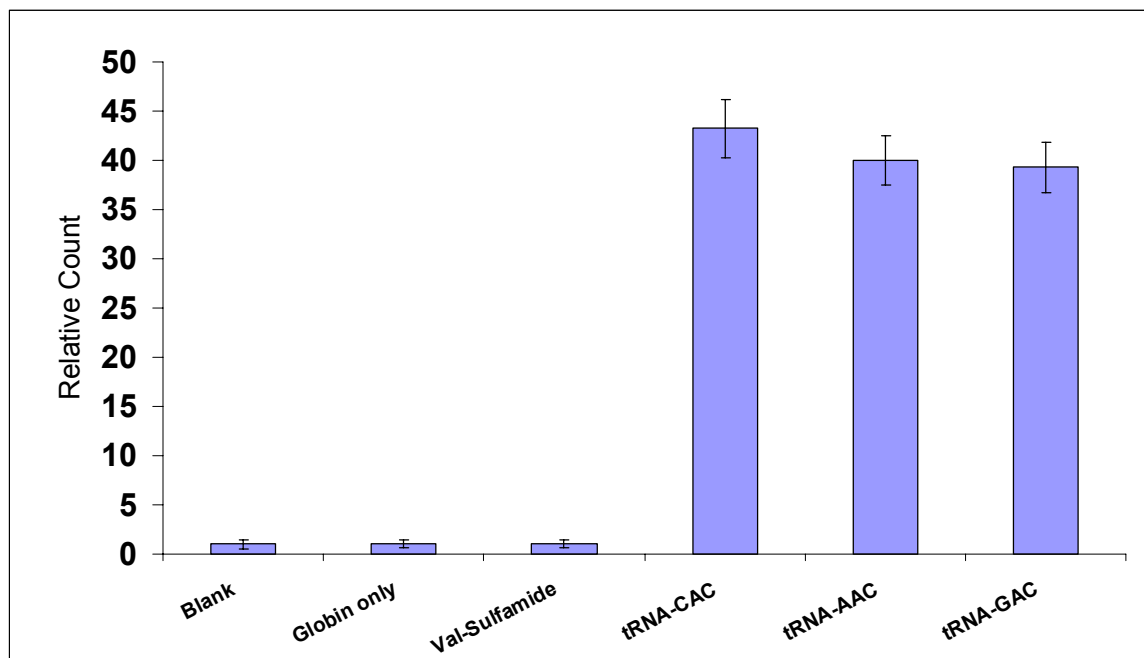
	unstained ladder	blank	Globin only	Val-Inh.	Restoration by biocytin-tRNA		
Biocytin-tRNA-CUA	-	-	-	-	4µg	-	-
Biocytin-tRNA-GAC	-	-	-	-	-	4µg	-
Biocytin-tRNA-AAC	-	-	-	-	-	-	4µg
Biocytin-tRNA-CAC	-	-	-	4µg	-	-	-
Val-Sulfamide	-	-	1 µM	1 µM	1 µM	1 µM	1 µM
Globin mRNA	-	+	+	+	+	+	+
Lane	0	1	2	3	4	5	6	7

Figure 8. Globin Translation Restoration probed with Western blot. Lane 0, precision protein standard unstained; lane 1, no template; lane 2, globin alone; lane 3, 1µM of val-sulfamide; lane 4, biocytin-tRNA-CAC (Valine codon GUG); lane 5, biocytin-tRNA-CUA (stop codon UAG); lane 6, biocytin-tRNA-GAC (Valine codon GUC); lane 7, biocytin-tRNA- AAC (Valine codon GUU) was added individually.

We further compared biocytin incorporation level in the presence and absence of Val-sulfamide. As shown in Figure 9, biocytin incorporation level in the presence of the inhibitor (lane 6 and 7) is higher than that in the absence of the inhibitor (lane 4 and 5). This suggests that the presence of the inhibitor provides enhancement of unnatural amino acids incorporation.

	unstained ladder	blank	Globin only	Val-Inh.	Restoration by biocytin-tRNA			
Biocytin-tRNA-AAC		-	-	-	4 µg	-	4 µg	-
Biocytin-tRNA-CAC		-	-	-	-	4 µg	-	4 µg
Val-Sulfamide		-	-	1 µM	-	-	1 µM	1 µM
Globin mRNA		-	+	+	+	+	+	+
Lane	0	1	2	3	4	5	6	7

Globin ⟶

Figure 9. Globin Translation Restoration probed with Western blot in the presence of inhibitor and in the absence of inhibitor. Lane 0, precision protein standard unstained; lane 1, no template; lane 2, globin alone; lane 3, 1 µM of Val-sulfamide; lane 4, biocytin-tRNA-AAC was added without the inhibitor; lane 5, biocytin-tRNA-CAC was added without the inhibitor: lane 6, biocytin-tRNA-AAC in the presence of 1 µM of Val-sulfamide; lane 7, biocytin-tRNA-CAC in the presence of 1 µM of Val-sulfamide.

We also tested the restoration of globin translation by biocytin-tRNAs coding for valine probed with Streptavidin-capture of biocytin-containing globin. [^{35}S]Met-biocytin-containing-globin was immobilized to streptavidin-agarose beads and then eluted by adding tricine loading buffer directly to the streptavidin-agarose beads with heating at 90 ºC for 10 min. The amount of immobilized [^{35}S]Met-biocytin containing globin was determined by scintillation counting of the eluent (Figure 10). This also demonstrated that in the presence of biocytin-tRNAs, globin translation was restored to about 35-40% relative to no drug translation.


Figure 10. Restoration of translation probed with streptavidin-capture of biocytin-containing globin.

Discussions

tRNA synthetase inhibitors enhance In Vitro suppression

In this work, we prepared a series of aminoacyl-adenylate analogues and evaluated their potency for inhibition of protein translation. Further, we characterized the ability of aminoacyl sulfamides to facilitate incorporation of an unnatural amino acid, biocytin, into globin protein. This approach provides a facile means to rewrite the universal genetic code at will, without extensively re-engineering cells. Therefore, this work provides an experimental means to pursue fundamental questions relating to the construction, functional diversity, and organization of the genetic code (Figure 11).

Figure 11.   Enhance *In Vitro* suppression by tRNA synthetase inhibitors.

Uniting mRNA display, In Vitro Suppression and tRNA synthetase inhibitors

The incorporation of unnatural amino acid into selectable, amplifiable peptide and protein libraries expands the chemical diversity of such libraries, thus considerably facilitating the process of obtaining ligands with improved properties (affinity, specificity, and function), particularly against therapeutically interesting targets. Here we also demonstrated that sense codon suppression was enhanced by tRNA synthetase inhibitors. To apply this unnatural strategy to selection cycle via mRNA display library, we can facilitate the enrichment of desired unnatural amino acid residue and speed up the selection process (Figure 12).
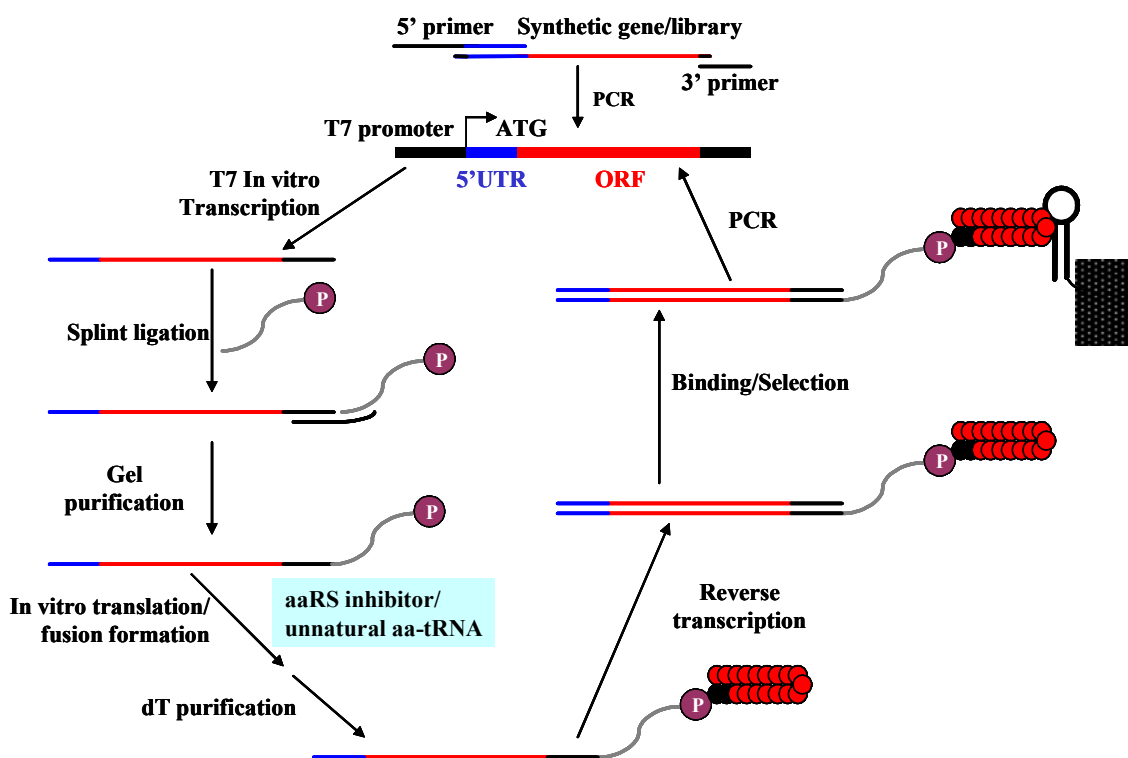
Figure 12. Selection cycle via mRNA display enhanced by tRNA synthetase inhibitor.

Conclusions

We generated various high-affinity inhibitors of the tRNA synthetases—aminoacyl sulfamides to create synthetic blanks in our translation extracts. Aminoacyl tRNA synthetase inhibitors were combined with translation extracts and protein translation was shown to be inhibited. We then measured the activity of each compound in a high dynamic-range IC_{50} potency assay using the rabbit reticulocyte protein synthesis system. The various aminoacyl sulfamides inhibit globin translation with IC_{50} in the range of 22 nM to 28 μM.

We further demonstrated the restoration of globin translation by biocytin-tRNAs coding for valine in the presence of Val-sulfamide inhibitor. Here, specific biocytin incorporation was probed by Western blot, and biocytin incorporation level in the presence of inhibitor is higher than that in the absence of inhibitor. This suggests that the presence of inhibitor provides enhancement of unnatural amino acids incorporation.

In this work, we demonstrated that we can create and reprogram synthetic "blanks" in the genetic code. Most importantly, we can facilitate the enrichment of desired unnatural amino acid residue and speed up the selection process when we unit mRNA display, *in vitro* Suppression and tRNA synthetase inhibitors.

References

1. Roesser, J. R., Chorghade, M. S. & Hecht, S. M. Ribosome-catalyzed formation of an abnormal peptide analogue. *Biochemistry* **25**, 6361-5 (1986).

2. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182-8 (1989).

3. Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C. J. & Schultz, P. G. Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol* **202**, 301-36 (1991).

4. Bain, J. D., Glabe, C. G., Dix, T. A., Chamberlin, A. R. & Diala, E. S. Biosynthetic Site-Specific Incorporation of a Non-Natural Amino-Acid into a Polypeptide. *J Am Chem Soc* **111**, 8013-8014 (1989).

5. Bain, J. D., Wacker, D. A., Kuo, E. E. & Chamberlin, A. R. Site-Specific Incorporation of Nonnatural Residues into Peptides - Effect of Residue Structure on Suppression and Translation Efficiencies. *Tetrahedron* **47**, 2389-2400 (1991).

6. Bain, J. D. et al. Site-Specific Incorporation of Nonnatural Residues During Invitro Protein-Biosynthesis with Semisynthetic Aminoacyl-Transfer Rnas. *Biochemistry* **30**, 5411-5421 (1991).

7. Karginov, V. A., Mamaev, S. V. & Hecht, S. M. In vitro suppression as a tool for the investigation of translation initiation. *Nucleic Acids Res* **25**, 3912-6 (1997).

8. Nowak, M. W. et al. In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Methods Enzymol* **293**, 504-29 (1998).

9. Nowak, M. W. et al. Nicotinic Receptor-Binding Site Probed with Unnatural Amino-Acid-Incorporation in Intact-Cells. *Science* **268**, 439-442 (1995).

10. Saks, M. E. et al. An engineered Tetrahymena tRNA(Gln) for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression. *J Biol Chem* **271**, 23169-23175 (1996).

11. Tao, J. S. & Schimmel, P. Inhibitors of aminoacyl-tRNA synthetases as novel anti-infectives. *Expert Opinion on Investigational Drugs* **9**, 1767-1775 (2000).

12. Li, S., Millward, S. & Roberts, R. In vitro selection of mRNA display libraries containing an unnatural amino Acid. *J Am Chem Soc* **124**, 9972-3 (2002).

13. Frankel, A. & Roberts, R. W. In vitro selection for sense codon suppression. *RNA* **9**, 780-786 (2003).

14. Smith, G. P. & Petrenko, V. A. Phage display. *Chem Rev* **97**, 391-410 (1997).

15. Zwick, M. B., Shen, J. Q. & Scott, J. Phage-displayed peptide libraries. *Current Opinion in Biotechnology* **9**, 427-436 (1998).

16. Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Applications of Chimeric Genes and Hybrid Proteins, Pt C* **328**, 430-444 (2000).

17. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A* **94**, 12297-302 (1997).

18. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett* **414**, 405-8 (1997).

19. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. The structural basis of the activation of Ras by Sos. *Nature* **394**, 337-343 (1998).

20. Milligan, J. F. & Uhlenbeck, O. C. Synthesis of Small Rnas Using T7 Rna-Polymerase. *Methods Enzymol* **180**, 51-62 (1989).

21. Hill, J. M. & Kluge, A. F. Cubist Pharmaceuticals, Inc., USA, patent 5,824,657 (1998).

22. Kolb, M., Danzin, C., Barth, J. & Claverie, N. Synthesis and Biochemical-Properties of Chemically Stable Product Analogs of the Reaction Catalyzed by S-Adenosyl-L-Methionine Decarboxylase. *J Med Chem* **25**, 550-556 (1982).

23. Schagger, H. & von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368-379 (1987).

24. Starck, S. R., Qi, X., Olsen, B. N. & Roberts, R. W. The puromycin route to assess stereo- and regiochemical constraints on Peptide bond formation in eukaryotic ribosomes. *J Am Chem Soc* **125**, 8090-1 (2003).

Chapter 3

Probing Flexibility of Protein Synthesis *In Vitro* by Puromycin Analogs

Abstract

Puromycin is a small-molecule mimic of aminoacyl-tRNA (aa-tRNA), and can be mistakenly inserted in place of aminoacyl-tRNA by ribosome and acts as a universal protein translation inhibitor. We have constructed a series of puromycin analogs with natural and unnatural amino acids side chains in order to systematically test the effects of side chain characteristics of amino acid moiety on the activity of puromycin. We then measured the activity of each compound in a high dynamic-range IC_{50} potency assay using the rabbit reticulocyte protein synthesis system. Both IC_{50} values and carboxypeptidase Y (CPY)-resistant results lead us to conclude that all of the compounds we synthesized both inhibit translation and participate in ribosome-mediated peptide bond formation.

Introduction

Puromycin has played an important role in our understanding of the ribosome and protein synthesis. Puromycin is a small-molecule mimic of aminoacyl-tRNA (aa-tRNA), and can be mistakenly inserted in place of aminoacyl-tRNA by ribosome and acts as a universal protein translation inhibitor (Figure 1). It enters the ribosomal A site and participates in peptide bond formation with the nascent peptidyl chain,[1] resulting in truncated proteins containing the drug at their C-terminus. Puromycin and puromycin analogs have been extremely useful in exploring the activity of various nucleophiles (-OH vs. -NH$_2$ vs. -SH) in the ribosome active site.[2-5]
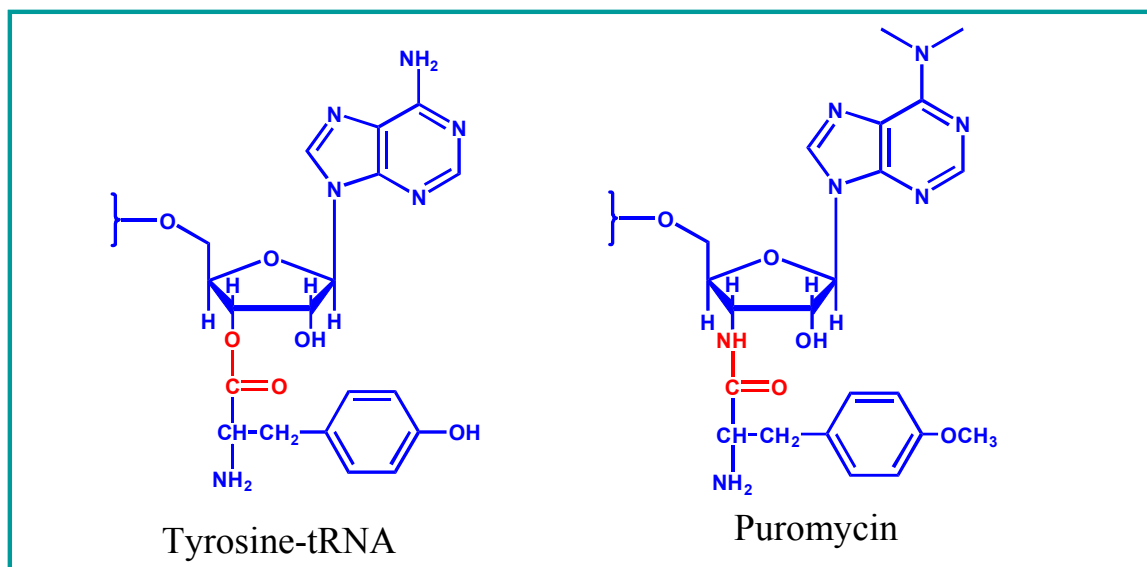
Figure 1. Structural comparison between Tyrosine-tRNA and puromycin.

Unlike aa-tRNA, puromycin is able to enter the ribosome independently (Figure 2), does not induce EF-Tu•GTPase activiy,[6] and does not appear to require soluble

translation factors for function.[7] Puromycin and related analogs thus provide a direct means to address ribosome-mediated peptide bond formation even in an intact translation extract.
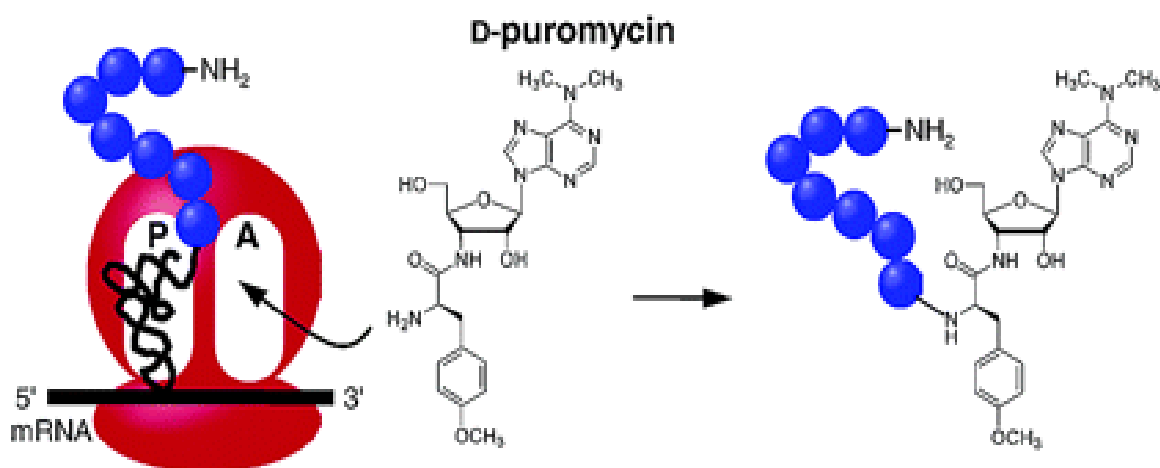
D-puromycin

Figure 2. D-puromycin incorporated in the peptidyltransferase center of the ribosome.

The protein synthesis machinery can be used to incorporate unnatural amino acids into peptides,[8-10] proteins,[11,12] and molecular libraries.[13-18] These studies indicate that the ribosome displays a broad ability to utilize residues beyond the 20 naturally occurring amino acids. Chemically misacylated tRNA fragments and tRNAs provide one route to probe the stereo- and regiospecificity of isolated ribosomes[19-22] and intact translation systems.[23] This approach has expanded our understanding of the range of residues that can be incorporated by the ribosome.[24-27] However, entry of both β- and D-amino acids has proved challenging.[9,11] Analysis of incorporation of these residues would be valuable and deepen our understanding of the stereo- and regiochemical constraints of ribosome-mediated peptide bond formation.

Materials and Methods

General Information.

^{1}H and ^{13}C NMR spectra were recorded on a Varian, Inc. UNITY INOVA instrument operating at 500 MHz using D_2O or DMSO-*d6* as the solvent. ^{1}H NMR data are reported as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br s, broad singlet; dd, doublet of doublets. High-resolution mass spectra (FAB) were recorded on a JMS-600H double-focusing, high-resolution, magnetic sector mass spectrometer at the Mass Spectrometry Laboratory, Division of Chemistry and Chemical Engineering, California Institute of Technology. Column chromatography was carried out on silica gel (40-63 μm, EM Science). Analytical HPLC was performed using a Vydac C18 column (5 mm, 4.5 x 250 mm) with buffer A (5 mM NH_4OAc, pH 5.5 with 10% acetonitrile) and buffer B (5 mM NH_4OAc, pH 5.5 with 90% acetonitrile); a linear gradient of 100% buffer B in 50 min was used with a flow rate of 1 mL/min. All reagents were of highest available commercial quality and were used without further purification. Puromycin aminonucleoside (3'-amino-3'-deoxy-*N, N'*-dimethyl-adenosine) (PANS) was purchased from Sigma Chemical Co. FMOC-(4-methoxy-D-phenylalanine) and FMOC-(D-alanine) were purchased from Bachem. FMOC-(4-methyl-L-phenylalanine), FMOC-(L-alanine), and FMOC-(L-β-homoalanine) were purchased from Fluka. FMOC-(4-methyl-D-phenylalanine) and FMOC-(4-methyl-L-β-phenylalanine) were from Peptech. Puromycin and puromycin analog concentrations were determined with the following extinction

coefficients ($M^{-1}cm^{-1}$) at 260 nm: L- and D-puromycin (**1a** and **1b**) [$\varepsilon = 11{,}790$] in H_2O; L-(4-Me)-Phe-PANS, D-(4-Me)-Phe-PANS, and L-β-(4-Me)-Phe-PANS (**2a** – **2c**) [$\varepsilon = 10{,}500$] in H_2O; and L-Ala-PANS, D-Ala-PANS, and L-β-Ala-PANS (**3a** – **3c**) [$\varepsilon = 11{,}000$] in phosphate buffered saline (pH 7.3).

Rabbit reticulocyte lysate was purchased from Novagen. Rabbit globin mRNA was obtained from Life Technologies Gibco BRL. L-Puromycin (**1a**) was purchased from Sigma Chemical Co. Ras mRNA was prepared by using two DNA primers complementary to the 5'- and 3'-ends of the coding region for H-Ras (pProEX HTb vector, a kind gift from Dafna Bar-Sagi)[1] to amplify the gene using PCR. mRNA was produced by T7 runoff transcription[2] of the H-Ras DNA in the presence of RNAsecure (Ambion) followed by gel purification via denaturing urea-PAGE and 'crush and soak' RNA isolation. L-[^{35}S]methionine (1,175 Ci/mmol) was purchased from NEN Life Science Products. Carboxypeptidase Y was obtained from Pierce. GF/A glass microfiber filters were from Whatman. Scintillation counting was carried out using a Beckman LS-6500 liquid scintillation counter.

General Procedure for Preparation of Puromycin Analogs.

N, N'-dicyclohexylcarbodiimide (DCC) (0.0539 mmol) was added to a cold (0 °C) solution of PANS (0.0520 mmol), FMOC-protected amino acid (0.0541 mmol), and *N*-hydroxysuccinimide (NHS) (0.0556 mmol) in dried *N,N'*-dimethylformamide (DMF) (0.900 mL). The solution was stirred for 30 min in an ice-water bath and then for 25 h at

ambient temperature. *N,N'*-dicyclohexylurea was filtered and washed (EtOAc, 4 mL), and the filtrate was concentrated *in vacuo*. For **1b**, the residue was resuspended in EtOAc, sonicated, and the mixture was filtered and then dried. The material was purified by gradient flash chromatography using $CHCl_3$ → MeOH/$CHCl_3$ (4:96) for **1b** or MeOH/$CHCl_3$ (7:93) for **2a-2c** and **3a-3c**. Homogenous product fractions were dried *in vacuo* to yield the FMOC-protected product. FMOC-deprotection was carried out in 20% (v/v) piperidine in DMF (5mL) with stirring for 30 min at ambient temperature. The solvent was removed *in vacuo* and the residue was subjected to gradient flash chromatography using $CHCl_3$ → MeOH/$CHCl_3$ (8:92) for **1b**, **2a**, and **2b** and TEA/MeOH/$CHCl_3$ (2:10:88) for **2c** and **3a-3c** to afford the titled products. Confirmation of purity was assessed using analytical HPLC.

9-{3'-Deoxy-3'-[(4-methoxy-D-phenylalanyl)amino]-β-D-ribofuranosyl}-6-(*N,N'*-dimethylamino)purine (D-puromycin) (1b).[3] White solid (31.5 mg, 87.3%): [1]H (DMSO-*d6*) δ 1.85 (br s, 2H), 2.58-2.63 (m, 1H), 2.93 (dd, *J* = 4.5, 14 Hz, 1H), 3.42 (dd, *J* = 4.5, 8.5 Hz, 2H), 3.51-3.56 (m, 2H), 3.72 (s, 6H), 3.93-3.96 (m, 1H), 4.47-4.51 (m, 4H), 5.17 (t, *J* = 5.5 Hz, 1H), 5.97 (d, *J* = 2.0 Hz, 1H), 6.17 (d, *J* = 5.0, 1H), 6.85 (d, *J* = 9.0 Hz, 2H), 7.15 (d, *J* = 8.5 Hz, 2H), 8.08 (br s, 1H), 8.24 (s, 1H), 8.45 (s, 1H); [13]C (DMSO-*d6*) δ 25.4, 50.6, 56.0, 56.8, 61.7, 73.8, 84.4, 90.2, 114.3, 131.0, 138.7, 150.4, 152.6, 158.0, 175.5; HRMS (FAB), *m/z* calculated for $C_{22}H_{30}N_7O_5$ (M+H)[+] 472.2311, found 472.2307.

9-{3'-Deoxy-3'-[(4-methyl-L-phenylalanyl)amino]-β-D-ribofuranosyl}-6-

(*N*,*N*'-dimethylamino)purine [L-(4-Me)-Phe-PANS] (2a). Pale white solid (18.8 mg,

80.8%): ^{1}H NMR (DMSO-*d6*) δ 1.84 (br s, 2H), 2.26 (s, 6H), 2.52-2.57 (m, 1H), 2.94

(dd, *J* = 4.5, 14 Hz, 1H), 3.44-3.52 (m, 2H), 3.67-3.70 (m, 2H), 3.92-3.95 (m, 1H), 4.44-

4.50 (m, 4H), 5.14 (t, *J* = 5.5 Hz, 1H), 5.98 (d, *J* = 3.0 Hz, 1H), 6.14 (d, *J* = 4.0 Hz, 1H),

7.10 (dd, *J* = 8.0, 18 Hz, 4H), 8.07 (d, *J* = 5.5 Hz, 1H), 8.24 (s, 1H), 8.45 (s, 1H); ^{13}C

(DMSO-*d6*) δ 21.3, 41.2, 50.7, 56.9, 61.7, 73.9, 84.3, 90.2, 120.3, 129.4, 129.8, 135.7,

136.3, 138.7, 150.4, 152.6, 155.0, 175.5; HRMS (FAB), *m/z* calculated for $C_{22}H_{30}N_7O_4$

(M+H)$^+$ 456.2362, found 456.2367.

9-{3'-Deoxy-3'-[(4-methyl-D-phenylalanyl)amino]-β-D-ribofuranosyl}-6-

(*N*,*N*'-dimethylamino)purine [D-(4-Me)-Phe-PANS] (2b). Pale white solid (20.7 mg,

88.8%): ^{1}H NMR (DMSO-*d6*) δ 1.85 (br s, 2H), 2.26 (s, 6H), 2.61 (dd, *J* = 8.0, 13 Hz,

1H) 2.96 (dd, *J* = 4.5, 14 Hz, 1H), 3.42-3.45 (m, 1H), 3.51-3.56 (m, 1H), 3.71-3.73 (m,

2H), 3.94-3.96 (m, 1H), 4.40-4.49 (m, 4H), 4.48 (d, *J* = 12 Hz, 1H), 5.17 (t, *J* = 5.5 Hz,

1H), 5.97 (d, *J* = 2.5 Hz, 1H), 6.19 (br s, 1H), 7.11 (dd, *J* = 8.0, 16 Hz, 4H), 8.10 (br s,

1H), 8.23 (s, 1H), 8.45 (s, 1H); ^{13}C (DMSO-*d6*) δ 21.4, 41.0, 50.6, 56.8, 61.7, 73.8, 84.4,

90.2, 120.3, 129.4, 129.9, 135.8, 136.1, 138.7, 150.4, 152.6, 155.0, 175.5; HRMS (FAB),

m/z calculated for $C_{22}H_{30}N_7O_4$ (M+H)$^+$ 456.2362, found 456.2360.

9-{3'-Deoxy-3'-[(4-methyl-L-β-phenylalanyl)amino]-β-D-ribofuranosyl}-6-(N,N'-dimethylamino)purine [L-β-(4-Me)-Phe-PANS] (2c). Pale white solid (17.8 mg, 73.0%): ^{1}H NMR (D$_2$0) δ 2.06 (s, 6H), 2.45-2.49 (m, 1H), 2.67 (d, J = 6.5 Hz, 1H), 3.18 (t, J = 6.0 Hz, 1H), 3.28 (br s, 3H), 3.37-3.39 (m, 1H), 3.49-3.50 (m, 1H), 3.59-3.62 (m, 1H), 3.80 (dd, J = 2.0, 13 Hz, 1H), 4.08-4.10 (m, 1H), 4.35 (dd, J = 6.0, 8.5 Hz, 1H), 4.46 (dd, J = 3.0, 5.5 Hz, 1H), 5.94 (d, J = 2.5 Hz, 1H), 7.03 (s, 4H), 8.03 (s, 1H), 8.15 (s, 1H); ^{13}C NMR (D$_2$0) δ 19.1, 26.0, 40.5, 49.8, 50.5, 54.5, 60.8, 73.6, 82.7, 89.7, 111.0, 120.0, 129.4, 129.6, 134.0, 137.2, 138.0, 148.8, 152.3, 173.6; HRMS (FAB), m/z calculated for C$_{23}$H$_{32}$N$_7$O$_4$ (M+H)$^+$ 470.2519, found 470.2508.

9-{3'-Deoxy-3'-[(L-alanine)amino]-β-D-ribofuranosyl}-6-(N,N'-dimethylamino)purine (L-Ala-PANS) (3a). Pale yellow solid (5.7 mg, 30.2%): ^{1}H NMR (D$_2$O) δ 1.41 (d, J = 7.0 Hz, 3H), 3.28 (br s, 6H), 3.63 (dd, J = 3.5, 13 Hz, 1H), 3.82 (dd, J = 2.5, 13 Hz, 1H), 3.97 (q, J = 7.0 Hz, 1H), 4.17-4.18 (m, 1H), 4.55-4.58 (m, 2H), 4.62-4.64 (m, 1H), 5.98 (d, J = 3.0 Hz, 1H), 8.03 (s, 1H), 8.17 (s, 1H); ^{13}C NMR (D$_2$O) δ 17.3, 39.0, 49.4, 51.0, 60.7, 73.5, 82.7, 89.6, 119.5, 138.0, 148.8, 152.2, 154.6, 172.4; HRMS (FAB), m/z calculated for C$_{15}$H$_{24}$N$_7$O$_4$ (M+H)$^+$ 366.1892, found 366.1889.

9-{3'-Deoxy-3'-[(D-alanine)amino]-β-D-ribofuranosyl}-6-(N,N'-dimethylamino)purine (D-Ala-PANS) (3b). Pale yellow solid (8.6 mg, 45.4%): ^{1}H NMR (D$_2$0) δ 1.43 (d, J = 7.5 Hz, 3H), 3.28 (br s, 6H), 3.65 (dd, J = 4.0, 13 Hz, 1H), 3.83 (dd, J = 2.5, 13 Hz, 1H), 4.02 (q, J = 7.5 Hz, 1H), 4.14-4.17 (m, 1H), 4.55-4.58 (m, 2H), 4.64-4.66 (m, 1H), 5.98 (d, J = 3.0 Hz, 1H), 8.04 (s, 1H), 8.17 (s, 1H); ^{13}C NMR (D$_2$0)

δ 17.0, 39.0, 49.3, 50.9, 60.8, 73.3, 82.8, 89.6, 106.0, 119.6, 138.1, 152.3, 154.8, 172.0;

HRMS (FAB), *m/z* calculated for $C_{15}H_{24}N_7O_4$ (M+H)$^+$ 366.1892, found 366.1898.

9-{3'-Deoxy-3'-[(L-β-homoalanine)amino]-β-D-ribofuranosyl}-6-(*N,N*'-dimethylamino)purine (L-β-Ala-PANS) (3c). Pale yellow solid (4.6 mg, 25.6%): ^{1}H NMR (D$_2$0) δ 1.16 (d, J = 6.5 Hz, 3H), 1.74 (s, 1H), 2.52 (d, J = 3.5 Hz, 2H), 3.24 (br s, 6H), 3.52-3.61 (m, 2H), 3.78 (d, J = 13 Hz, 1H), 4.11 (d, J = 5.5 Hz, 1H), 4.62-4.64 (m, 2H), 5.93 (s, 1H), 7.99 (s, 1H), 8.13 (s, 1H); ^{13}C NMR (D$_2$0) δ 18.2, 39.0, 39.5, 45.0, 50.7, 60.7, 73.5, 82.7, 89.6, 119.5, 138.0, 148.8, 152.2, 172.7; HRMS (FAB), *m/z* calculated for $C_{16}H_{26}N_7O_4$ (M+H)$^+$ 380.2049, found 380.2054.

IC$_{50}$ Determination.

Translation reactions containing [^{35}S]Met were made up in batch on ice and added in aliquots to microcentrifuge tubes containing an appropriate amount puromycin or puromycin analog dried *in vacuo*. Typically, a 20 μl translation mixture consisted of 0.8 μL of 2.5 M KCl, 0.4 μL of 25 mM MgOAc, 1.6 μL of 12.5X Translation Mixture without methionine (25 mM dithiothreitol (DTT), 250 mM HEPES (pH 7.6), 100 mM creatine phosphate, and 312.5 μM of 19 amino acids, except methionine), 3.6 μL of nuclease-free water, 0.6 μL (6.1 μCi) of [^{35}S]Met (1175 Ci/mmol), 8 μL of Red Nova$^®$ nuclease-treated lysate, and 5 μL of 0.05 μg/μL globin mRNA. Inhibitor, lysate preparation (include all components except template), and globin mRNA were mixed

simultaneously and incubated at 30 ºC for 60 min. Then 2 µL of each reaction was combined with 8 µL of tricine loading buffer (80 mM Tris-Cl (pH 6.8), 200 mM DTT, 24% (v/v) glycerol, 8% sodium dodecyl sulfate (SDS), and 0.02 % (w/v) Coomassie blue G-250), heated to 90 ºC for 5 min, and applied entirely to a 4% stacking portion of a 16% tricine SDS-polyacrylamide gel containing 20% (v/v) glycerol[4] (30 mA for 1.5h). Gels were fixed in 10% acetic acid (v/v) and 50% (v/v) methanol, dried, exposed overnight on a PhosphorImager screen, and analyzed using a Storm PhosphorImager (Molecular Dynamics). Analysis in Figure S1 was carried out as described above except 6 µL of each reaction and 24 µL of tricine loading buffer were loaded (1.5-fold increase in stacking and resolving portion of gel; 30mA for 7 h).

Carboxypeptidase Assay.

Translation reactions were prepared as described for IC_{50} determination except reactions (50 µL) contained 2 µL of 2.5 M KOAc, 1 µL of 25 mM MgOAc, 4 µL of 12.5X Translation Mixture without methionine (25 mM dithiothreitol (DTT), 250 mM HEPES (pH 7.6), 100 mM creatine phosphate, and 312.5 µM of 19 amino acids, except methionine), 16 µL (163 µCi) of [^{35}S]Met (1175 Ci/mmol), 20 µL of Red Nova® nuclease-treated lysate, and 6.96 µL of 230 µg/mL Ras mRNA.[3] Inhibitor, lysate components, and Ras mRNA were mixed simultaneously and incubated at 30 ºC for 60 min. Then 2 µL of reaction was combined with 150 µL of 0.1 M sodium acetate (pH 5.0)
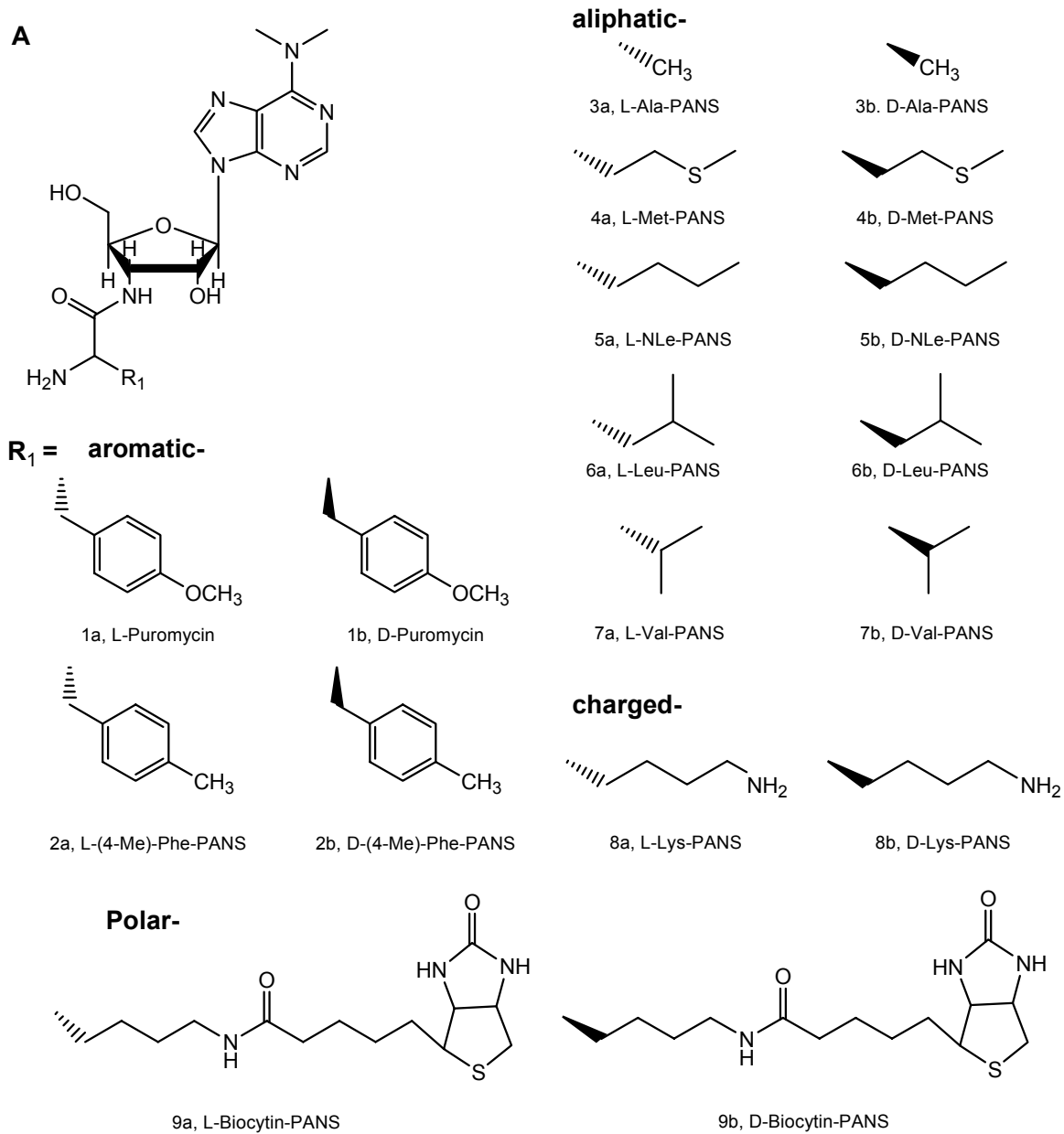
and 17 µL carboxypeptidase Y (CPY) (1 mg/mL in 0.05 M sodium citrate (pH 5.3) Pierce), and incubated at 37 ºC for 18 h. After incubation, reactions were mixed with 100 µL of 1 N NaOH/2% H_2O_2 (hydrolyzes charged tRNAs and removes the red color that may quench scintillation counting) and incubated at 37 ºC for 10 min to hydrolyze the charged tRNAs. Then 0.9 mL of 25% trichloroacetic acid (TCA)/2% casamino acids was added to the samples, vortexed, and put on ice for 10 min. The samples were filtered on GF/A filters (pre-soaked in 5% TCA), washed 3 times with 3-mL portions of cold 5% TCA, and scintillation counted to determine the amount of [^{35}S]Met-Ras. For the no CPY-treated samples, [^{35}S]Met-Ras (2 µL of reaction) was TCA precipitated without CPY treatment as described above.
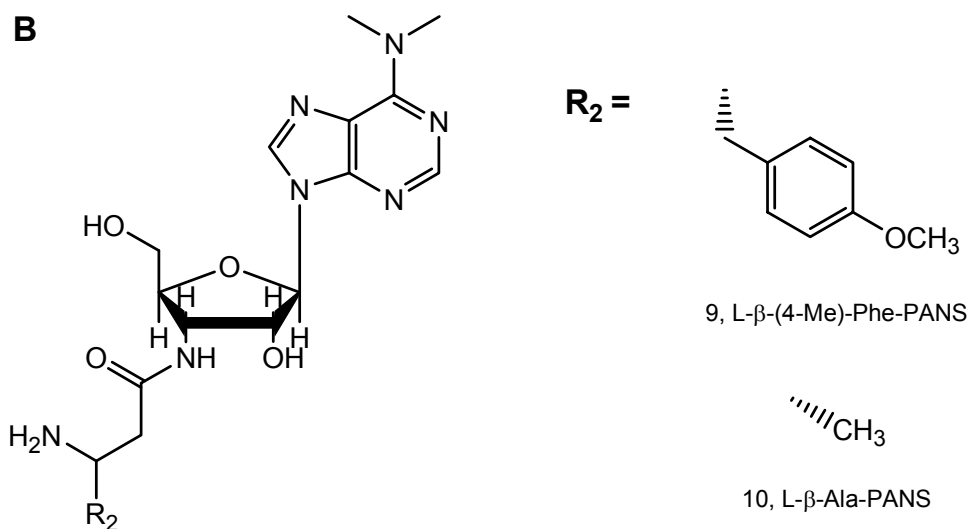
Results

Here, we have used a series of synthetic puromycin analogs with a variety of α-amino acid moieties in both L- and D-configurations, and β-amino acids, to measure their activity in an intact eukaryotic translation system. We have synthesized a series of puromycin derivatives (Figure 3) that differ in the 1) amino acid moiety, 2) amino acid stereochemistry, and 3) number of carbon units in the amino acid backbone. The synthesis route of these compounds is outlined in Scheme 1.

Fmoc-L-Lys(Fmoc)-PANS
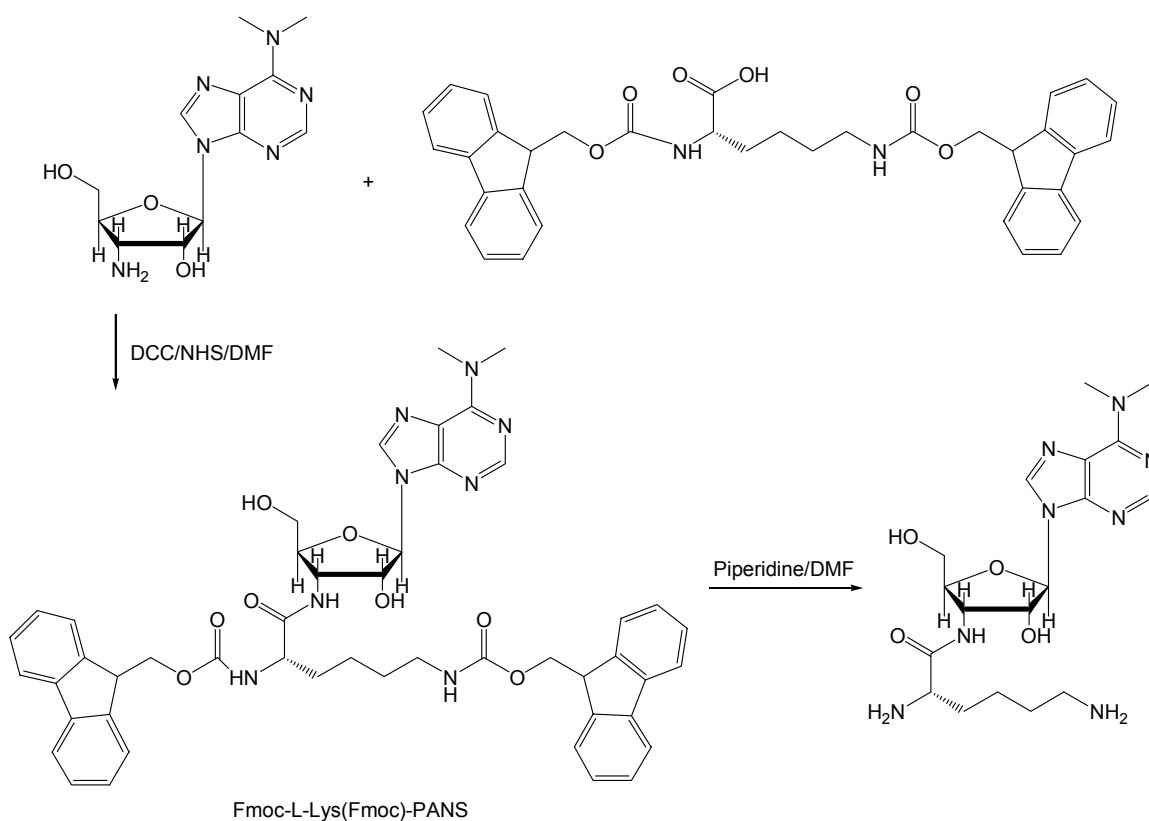
**Scheme 1.** Synthesis of puromycin analogs (Lys- derivative is used here as an example)

A

R₁ =   **aromatic-**

1a, L-Puromycin

1b, D-Puromycin

2a, L-(4-Me)-Phe-PANS

2b, D-(4-Me)-Phe-PANS

aliphatic-

3a, L-Ala-PANS

3b. D-Ala-PANS

4a, L-Met-PANS

4b, D-Met-PANS

5a, L-NLe-PANS

5b, D-NLe-PANS

6a, L-Leu-PANS

6b, D-Leu-PANS

7a, L-Val-PANS

7b, D-Val-PANS

charged-

8a, L-Lys-PANS

8b, D-Lys-PANS

Polar-

9a, L-Biocytin-PANS

9b, D-Biocytin-PANS

Figure 3A. Puromycin analogs with L- and D- α-amino acid side chains.

B

$R_2 =$

9, L-β-(4-Me)-Phe-PANS

10, L-β-Ala-PANS

Figure 3B. Puromycin analogs with L-β- amino acid sidechains.

We then measured the activity of each compound in a high dynamic-range IC_{50} potency assay using the rabbit reticulocyte protein synthesis system.[7] The naturally occurring compound, L-puromycin (**1a**), inhibits globin mRNA translation with an IC_{50} of 1.8 μM (Table 1). Surprisingly, D-puromycin (**1b**) also inhibited translation giving an IC_{50} of 280 μM, a difference of 150-fold (Figure 4).
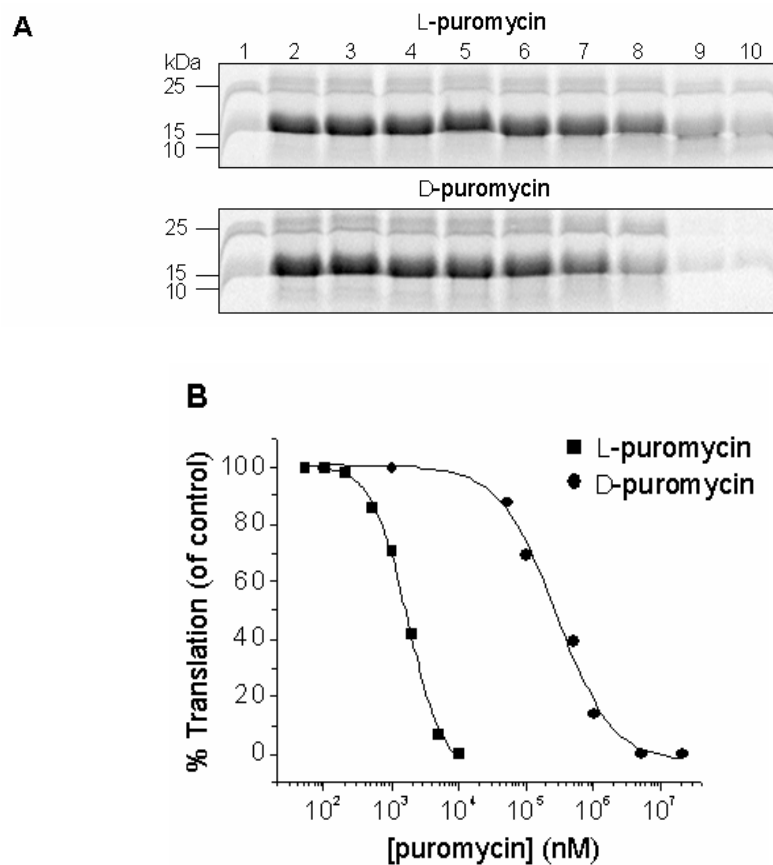
Figure 4. IC$_{50}$ determination for puromycin. (A) Tricine-SDS-PAGE analysis of [^{35}S]Met-globin translation reactions in the presence of L-puromycin (**1a**) and D-puromycin (**1b**): Lane 1, no template; lane 2, globin alone; lanes 3-10, concentrations from 50 nM to 10 mM for **1a** and from 100 nM to 20 mM for **1b**. (B) Percent globin translation relative to the no drug control for L-puromycin (**1a**) and D-puromycin (**1b**).

We reasoned that stereoselectivity should be a function of the side chain size and geometry. To test this, we constructed compounds where the puromycin side chain was altered to bear either a bulky aromatic side chain (L- and D-4-methyl-phenylalanine; **2a** and **2b**), aliphatic side chains of different length (L- and D-alanine, **3a** and **3b**; L- and D-methionine, **4a** and **4b**; L- and D-norleucine, **5a** and **5b**; L- and D-leucine, **6a** and **6b**; L- and D-valine, **7a** and **7b**), a charged side chains (L- and D-lysine, **8a** and **8b**), or a polar side chains (L- and D-Biocytin, **9a** and **9b**).

**Table 1.** IC_{50} values for puromycin analogs.

| | Puromycin Analogs | IC50 (μM) | | |
		L- (a)	D- (b)	D/L
1	Puromycin	1.8	280	156
2	(4-Me)-Phe-PANS	1.0	2400	2400
3	Ala-PANS	730	1900	2.6
4	Met-PANS	15	765	51
5	Nle-PANS	61	937	15
6	Leu-PANS	105	776	7.4
7	Val-PANS	3500	5300	1.5
8	Lys-PANS	289	347	1.2
9	Biocytin-PANS	2900	2900	1.0
10	β-(4-Me)-Phe-PANS	600	-	-
11	β-Ala-PANS	1700	-	-

Compound **2a** inhibits translation better than puromycin itself (IC_{50} = 1.0 μM) and is the most potent compound we constructed. The D-amino acid variant (**2b**) shows significantly reduced activity (IC_{50} = 2400 μM) with ~9-fold reduced activity compared to D-puromycin (**1b**), and is 2400-fold less potent than the L-isomer.

Compared to the natural occurring L-puromycin (1.8 μM), the other analogs we constructed, including Methionine (Met), Leucine (Leu), NorLeucine (Nle), Lysine (Lys), Biocytin, and Valine (Val), Alanine (Ala), are all less potent as inhibitor of protein synthesis (Table 1), but they show some interesting trend. For example, the IC_{50} values for L-Lys and D-Lys analogs are very similar, 289 μM and 347 μM, respectively. This shows that the ribosome perhaps does not discriminate L- and D-configurations for charged amino acid analogs. To be consistent with this observation, the puromycin analogs with biocytin side chain (biotinylated lysine) also have the same IC_{50} (2900 μM) between the L- and D-configuration, though much less potent.

The IC_{50} values for the puromycin analog with unnatural amino acid L-Nle and D-Nle are 61 μM and 937 μM, respectively. The Nle has the same aliphatic side chain as Lys but lacks the terminal amino group therefore lacks the positive charge. In contrast to Lys derivatives, L-Nle is 15-fold more potent than its D- counterpart. This allows us to conclude that the more hydrophobic the amino acid moiety, the more different is the potency of the D- and L- puromycin analog, consistent with our previous observation that larger hydrophobic aromatic side chain is a key element of ribosome discrimination.

In line with this observation, the Leu analogs have IC$_{50}$ values of 105 μM and 776 μM for L- and D-configuration, a 7-fold difference (Figure 5). These are very similar to the Nle analogs because Leu and Nle have similar aliphatic side chains. For Met, L-configuration (15 μM) is about 51-fold different than the D-configuration (765 μM). On the contrary, Val analogs are much less potent, having IC$_{50}$ values of 3500 μM and 5300 μM, with little discrimination, similar to the case of Ala. The alanine analogs show only a 3-fold selectivity for the L- versus D- isomers (**3a** vs. **3b**).
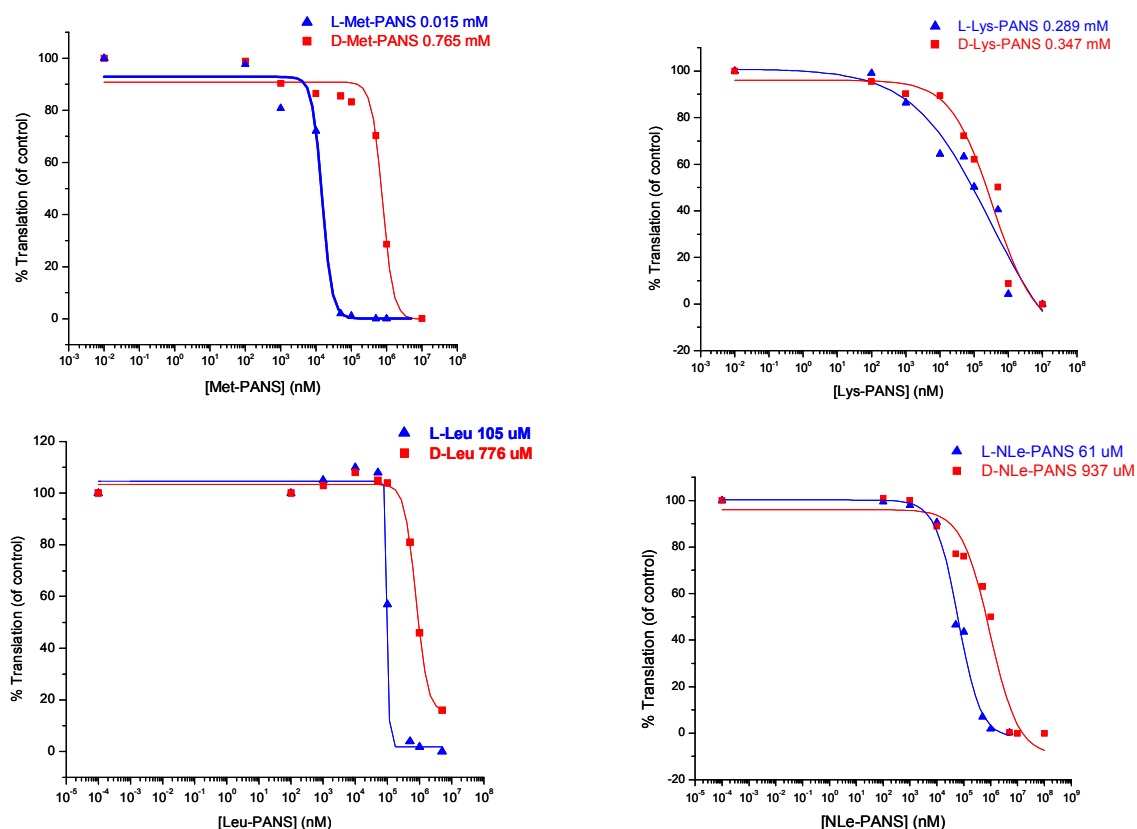
Figure 5. IC$_{50}$ comparison for L- and D- puromycin.

We then next examined puromycin derivatives bearing β-amino acids. β-amino acids have been previously incorporated at low levels using nonsense suppression techniques.[10,11] In our experiments, both L-β-(4-Me)-Phe-PANS (**10**) and L-β-Ala-PANS (**11**) gave measurable inhibition constants (IC_{50} = 600 and 1700 μM, respectively) (Table 1).

Finally, we wished to confirm that our puromycin analogs participated in peptide bond formation within the ribosome. Incorporation of puromycin blocks the C-terminus, rendering the protein carboxypeptidase resistant.[1] Previous work in our laboratory indicated that puromycin incorporation is measureable when the analog concentration is near the IC_{50}.[7] Protein synthesis performed in the presence of our puromycin derivatives resulted in a 12- to 16-fold increase in carboxypeptidase Y (CPY)-resistant protein compared with the no puromycin, Ras only control (Figures 6A & B). The production of truncated protein fragments in the presence of each puromycin derivative, previously shown to result from puromycin attachment to globin,[7] along with the CPY-resistance results lead us to conclude that all of the compounds we synthesized both inhibit translation and participate in ribosome-mediated peptide bond formation.
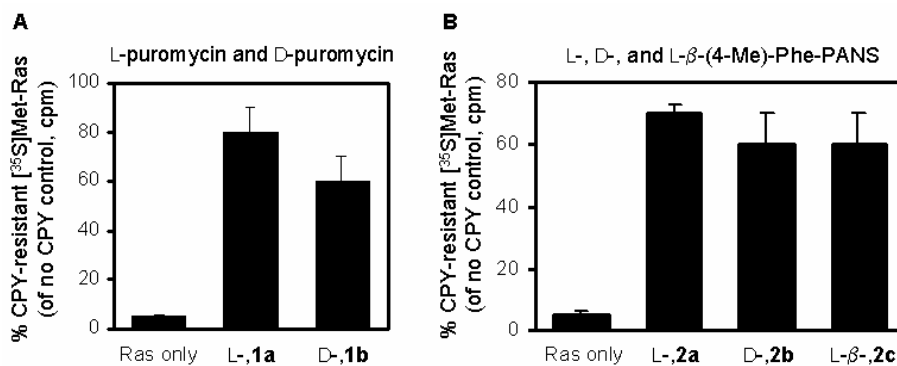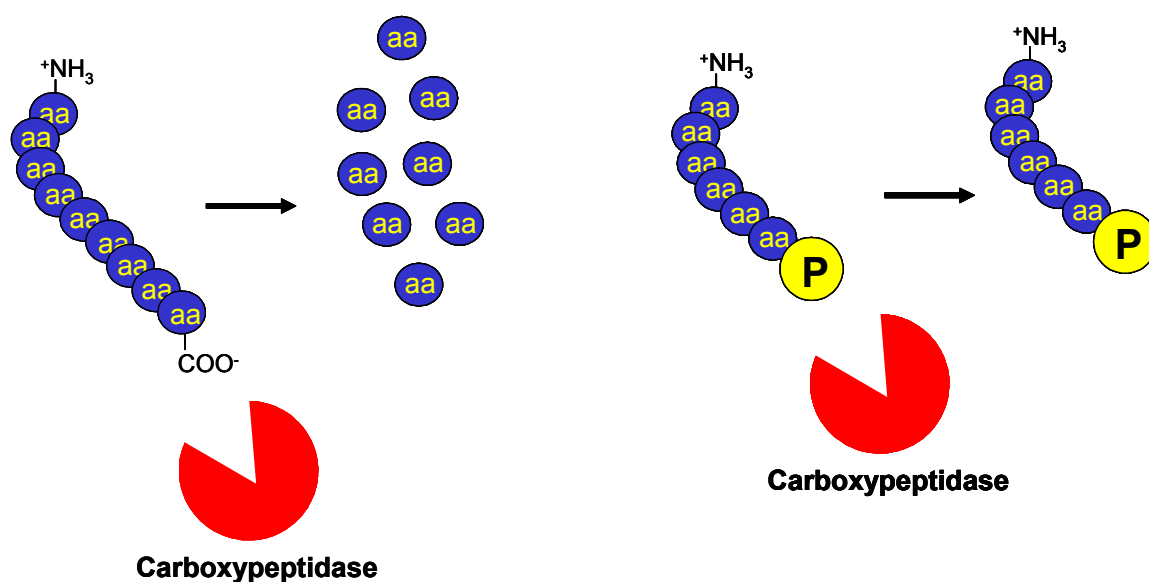
Figure 6. Carboxypeptidase Y (CPY) analysis of protein-puromycin products.[11] TCA precipitation of [^{35}S]Met-protein (Ras) from translation reactions after CPY treatment containing (A) Ras only, L-puromycin (**1a**) at 2 μM and D-puromycin (**1b**) at 500 μM and (B) Ras only, L-(4-Me)-Phe-PANS (**2a**) at 1 μM, D-(4-Me)-Phe-PANS (**2b**) at 1500 μM, and L-β-(4-Me)-Phe-PANS (**2c**) at 1000 μM. Data represent the mean ± standard error for at least three independent experiments.

Discussions

We vision that stereoselectivity should be a function of the side chain size and geometry. In order to systematically test the effects of size chain characteristics of amino acid moiety on the activity of puromycin, we have constructed a series of puromycin analogs with natural and unnatural amino acids side chains. These compounds include 4-methyl-Phenylalanine, Alanine(Ala), Methionine (Met), Leucine (Leu), NorLeucine (Nle), Lysine (Lys), Biocytin, and Valine (Val), in addition to naturally occurring puromycin, in both L- and D-configurations. They have either aromatic, aliphatic hydrophobic side chains of different length, or charged/polar side chains. This allows us to specifically probe different type of interactions between these analogs with the ribosome. We have used the same high dynamics IC_{50} assay to measure the activities of these puromycin analogs in inhibition of globin translation.

These observations argue that ribosomal stereoselectivity falls over a broad range and is primarily dictated by the size and geometry of the pendant side chain. Within the L-amino acid series (**1a**, **2a**, and **3a**), marked variation is also seen based solely on side chain identity. Larger, hydrophobic side chains provide improved function. In the L-amino acid series, 4-methylphenylalanine and naturally occurring puromycin are the best, followed by long aliphatic amino acids (Met, Nle, Leu). This is consistent with previous observations.[4,5] In the D-amino acid series, the 4-O-methyltyrosine derivative (**1b**) functions the best overall, and shows even better activity than the natural L-Ala and L-Val variants.

Overall, an aromatic side chain provides the highest stereospecificity by ribosome between L- and D-configuration, followed by long hydrophobic aliphatic side chains. Charged and polar side chains provide little discrimination.

We proposed that steric hindrance is the main reason that L-puromycin analogs are in general more potent than D-analogs (Figure 7), and our observation that aromatic and hydrophobic amino acid side chains exhibit higher discrimination than either charged or smaller side chains is consistent with this hypothesis.
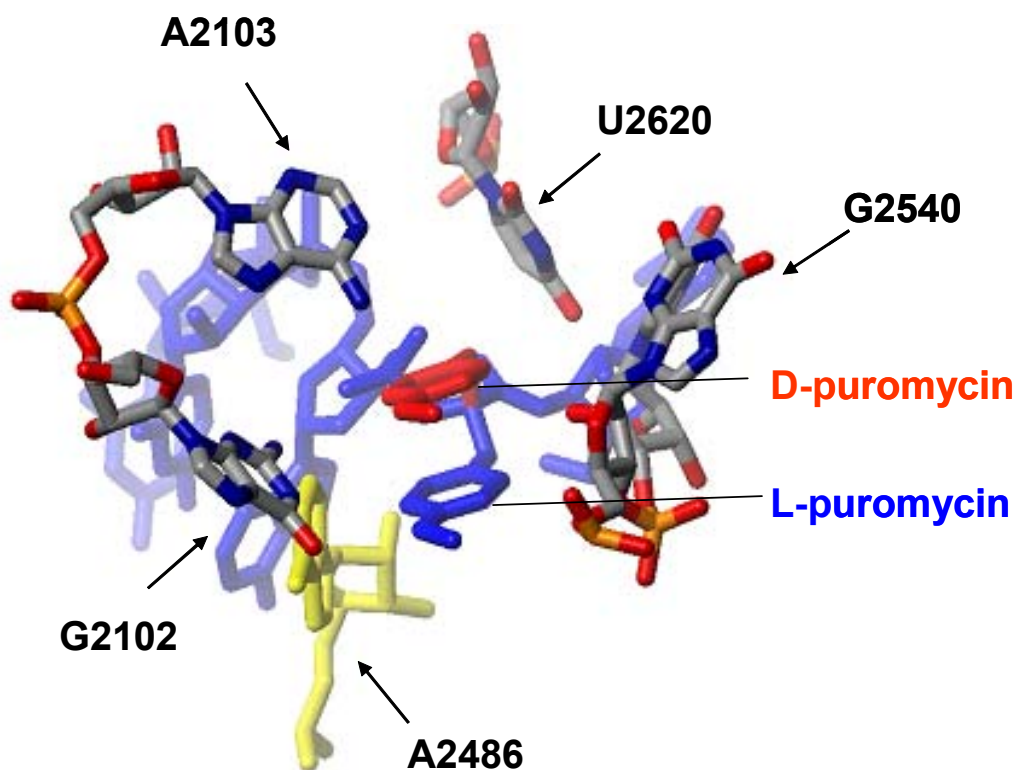
Figure 7. Model for D-puromycin (red) placement in the large 50S ribosome CCdA-p-L-puromycin (blue) complex. There is some potential hydrogen bond interactions of Met side chain with A2486.

The structural basis for stereoselectivity in rabbit ribosomes cannot be addressed presently, as there are no high-resolution structures available. However, modeling D-puromycin (**1b**) into the active site of the *Haloarcula marismortui* 50S subunit[28] is consistent with the idea that steric effects play a role in chiral discrimination. In the atomic resolution structure, U2620 (U2585 *E. coli*) is the closest nucleotide to the D-side chain. Also, while many of the ribosome active site nucleotides are highly conserved, the fact that critical residues can be mutated,[29] implies that construction of ribosomes with altered stereo- and regiospecificity may be possible. In the case of Met, there could potentially be hydrogen bonding interactions with the sulphur atom from the ribosome nucleotide for the L-derivative and this is consistent with the fact that there is about 51 fold difference between L- and D- Met-PANS.

Conclusions

Our work demonstrates that L-, D-, and β-amino acids can participate in ribosome-mediated peptide bond formation when constructed as analogs of puromycin. Our approach allows us to examine the activity of the ribosome directly in a physiologically complete protein-synthesizing system. Measurements using intact systems are critical as these can produce very different results from reconstituted or purified systems that are incapable of synthesizing proteins.[30,31] Our results provide one metric of the chiral and regiospecificity of mammalian ribosomes. We are hopeful that this data along with other information, such as the ability to optimize tRNA affinity for elongation factor Tu (EF-Tu)[32,33](EF1A in eukaryotes), will facilitate the incorporation of desirable but recalcitrant residues into peptides and proteins.

References

1. Nathans, D. Puromycin Inhibition of Protein Synthesis - Incorporation of Puromycin into Peptide Chains. *Proc Natl Acad Sci U S A* **51**, 585-& (1964).

2. Fahnesto.S, Neumann, H., Shashoua, V. & Rich, A. Ribosome-Catalyzed Ester Formation. *Biochemistry* **9**, 2477-& (1970).

3. Gooch, J. & Hawtrey, A. O. Synthesis of Thiol-Containing Analogues of Puromycin and a Study of Their Interaction with N-Acetylphenylalanyl-Transfer Ribonucleic-Acid on Ribosomes to Form Thioesters. *Biochemical Journal* **149**, 209-220 (1975).

4. Nathans, D. & Neidle, A. Structural Requirements for Puromycin Inhibition of Protein Synthesis. *Nature* **197**, 1076-& (1963).

5. Harris, R. J., Hanlon, J. E. & Symons, R. H. Peptide Bond Formation on Ribosome - Structural Requirements for Inhibition of Protein Synthesis and of Release of Peptides from Peptidyltransferrna on Bacterial and Mammalian Ribosomes by Aminoacyl and Nucleotidyl Analogues of Puromycin. *Biochimica Et Biophysica Acta* **240**, 244-& (1971).

6. Campuzano, S. & Modolell, J. Hydrolysis of Gtp on Elongation-Factor Tu.Ribosome Complexes Promoted by 2'(3')-O-L-Phenylalanyladenosine. *Proc Natl Acad Sci U S A* **77**, 905-909 (1980).

7.   Starck, S. R. & Roberts, R. W. Puromycin oligonucleotides reveal steric restrictions for ribosome entry and multiple modes of translation inhibition. *RNA* **8**, 890-903 (2002).

8.   Heckler, T. G., Zama, Y., Naka, T. & Hecht, S. M. Dipeptide Formation with Misacylated Transfer Rna-Phes. *Journal of Biological Chemistry* **258**, 4492-4495 (1983).

9.   Bain, J. D., Glabe, C. G., Dix, T. A., Chamberlin, A. R. & Diala, E. S. Biosynthetic Site-Specific Incorporation of a Non-Natural Amino-Acid into a Polypeptide. *J Am Chem Soc* **111**, 8013-8014 (1989).

10.   Bain, J. D., Wacker, D. A., Kuo, E. E. & Chamberlin, A. R. Site-Specific Incorporation of Nonnatural Residues into Peptides - Effect of Residue Structure on Suppression and Translation Efficiencies. *Tetrahedron* **47**, 2389-2400 (1991).

11.   Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182-8 (1989).

12.   Nowak, M. W. et al. Nicotinic Receptor-Binding Site Probed with Unnatural Amino-Acid-Incorporation in Intact-Cells. *Science* **268**, 439-442 (1995).

13.   Li, S., Millward, S. & Roberts, R. In vitro selection of mRNA display libraries containing an unnatural amino Acid. *J Am Chem Soc* **124**, 9972-3 (2002).

14.   Takahashi, T. T., Austin, R. J. & Roberts, R. W. mRNA display: ligand discovery, interaction analysis and beyond. *Trends Biochem Sci* **28**, 159-65 (2003).

15.   Ellman, J. A., Mendel, D. & Schultz, P. G. Site-Specific Incorporation of Novel Backbone Structures into Proteins. *Science* **255**, 197-200 (1992).

16. Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C. J. & Schultz, P. G. Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol* **202**, 301-36 (1991).

17. van Hest, J. C. M. & Tirrell, D. A. Protein-based materials, toward a new level of structural control. *Chem Commun*, 1897-1904 (2001).

18. Cornish, V. W., Mendel, D. & Schultz, P. G. Probing Protein-Structure and Function with an Expanded Genetic-Code. *Ange Chem-Int Ed* **34**, 621-633 (1995).

19. Yamane, T., Miller, D. L. & Hopfield, J. J. Discrimination between D-Tyrosyl and L-Tyrosyl Transfer Ribonucleic-Acids in Peptide-Chain Elongation. *Biochemistry* **20**, 7059-7064 (1981).

20. Chladek, S. & Sprinzl, M. The 3'-End of Transfer-Rna and Its Role in Protein-Biosynthesis. *Ange Chem-Int Ed* **24**, 371-391 (1985).

21. Heckler, T. G., Roesser, J. R., Xu, C., Chang, P. I. & Hecht, S. M. Ribosomal-Binding and Dipeptide Formation by Misacylated Transfer Rnaphes. *Biochemistry* **27**, 7254-7262 (1988).

22. Roesser, J. R., Xu, C., Payne, R. C., Surratt, C. K. & Hecht, S. M. Preparation of Misacylated Aminoacyl-Transfer Rna Phes Useful as Probes of the Ribosomal Acceptor Site. *Biochemistry* **28**, 5185-5195 (1989).

23. Mendel, D., Ellman, J. & Schultz, P. G. Protein-Biosynthesis with Conformationally Restricted Amino-Acids. *J Am Chem Soc* **115**, 4359-4360 (1993).

24. Killian, J. A., Van Cleve, M. D., Shayo, Y. F. & Hecht, S. M. Ribosome-mediated incorporation of hydrazinophenylalanine into modified peptide and protein analogues. *J Am Chem Soc* **120**, 3032-3042 (1998).

25. Koh, J. T., Cornish, V. W. & Schultz, P. G. An experimental approach to evaluating the role of backbone interactions in proteins using unnatural amino acid mutagenesis. *Biochemistry* **36**, 11314-11322 (1997).

26. Eisenhauer, B. M. & Hecht, S. M. Site-specific incorporation of (aminooxy)acetic acid into proteins. *Biochemistry* **41**, 11472-11478 (2002).

27. England, P. M., Zhang, Y. N., Dougherty, D. A. & Lester, H. A. Backbone mutations in transmembrane domains of a ligand-gated ion channel: Implications for the mechanism of gating. *Cell* **96**, 89-98 (1999).

28. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920-930 (2000).

29. Polacek, N., Gaynor, M., Yassin, A. & Mankin, A. S. Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature* **411**, 498-501 (2001).

30. Bhuta, P., Kumar, G. & Chladek, S. Aminoacyl Derivatives of Nucleosides, Nucleotides and Polynucleotides .36. The Peptidyltransferase Center of Escherichia-Coli Ribosomes - Binding-Sites for the Cytidine 3'-Phosphate Residues of the Aminoacyl-Trna 3'-Terminus and the Interrelationships between the Acceptor and Donor Sites. *Biochimica Et Biophysica Acta* **696**, 208-211 (1982).

31. Krayevsky, A. A. & Kukhanova, M. K. The peptidyltransferase center of ribosomes. *Prog Nucleic Acid Res Mol Biol* **23**, 1-51 (1979).

32. LaRiviere, F. J., Wolfson, A. D. & Uhlenbeck, O. C. Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294**, 165-168 (2001).

33. Asahara, H. & Uhlenbeck, O. C. The tRNA specificity of Thermus thermophilus EF-Tu. *Proc Natl Acad Sci U S A* **99**, 3499-504 (2002).

Chapter 4

Total Synthesis of Hydroxymethylacylfulvene

## Abstract

Hydroxymethylacylfulvene (**1**) (HMAF, also MGI 114) is a promising antitumor compound derived from the sesquiterpene illudin S (**2**). It is less cytotoxic than illudin S to normal cells and exhibits much greater selectivity in toxicity to malignant cells. A simple synthetic method is demonstrated to rapidly construct the skeleton of HMAF and its analogs starting from 2-(3-methoxyphenyl)ethanol. In particular this synthesis opens the door to enantioselective aromatic oxidation chemistry.

Introduction

Despite massive efforts put forward during the 20th century, human cancers continue to devastate the lives of millions of people worldwide. It has been estimated that over 100,000 Americans are diagnosed with cancer per year. The development of an antitumor drug that is highly efficient and less toxic to normal cells is needed. Hydroxymethylacylfulvene **1** is derived from the sesquiterpene illudin S by treatment with dilute sulfuric acid and excess paraformaldehyde. Illudin S was first isolated from the basidiomycete *Omphalotus illudens*. During the past 10 years, extensive preclinical studies on HMAF and related fulvenes have been conducted, leading to phase I human clinical trials which began in December 1995. Phase II clinical trials targeting several solid tumor types are now in progress under the sponsorship of MGI Pharma and the National Cancer Institute.[1,2]

HMAF
1
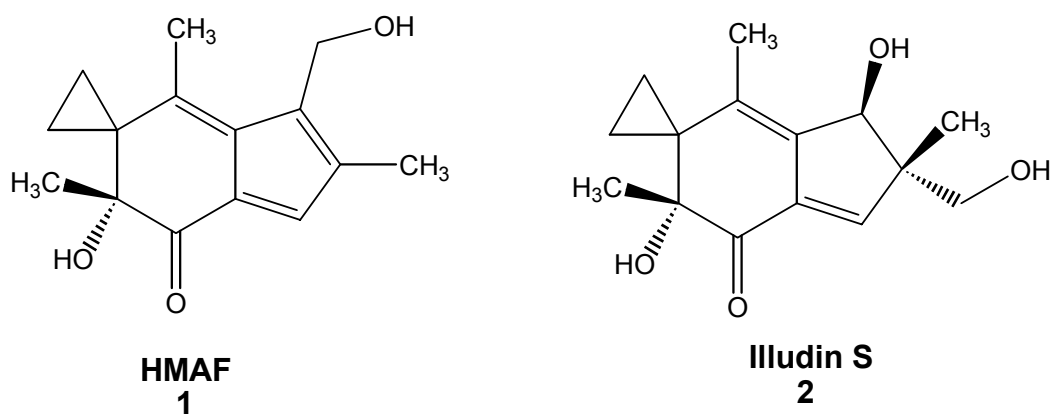
Illudin S
2

It is well known that many antitumor natural products behave as alkylating agents (Scheme 1). Among them are illudin compounds which react preferentially with thiols, and their cytotoxicity is attributed to the ability to react with vital thiol enzymes. Thiols

react readily at room temperature, adding to the α,β-unsaturated carbonyl and giving a cyclohexadiene intermediate which rapidly undergoes opening of the cyclopropane and loss of the tertiary hydroxyl. Reaction with thiols, e.g., methylthioglycolate, cysteine and glutathione is pH-dependent, with the optimum pH being 5.6-6.1. Not surprisingly, toxicity can be modulated by varying glutathione levels in cells. A third generation analog hydroxymethylacylfulvene (HMAF) caused complete tumor regression in all animals at the maximum tolerated dose of 10 mg/kg (iv) three times per week for 3 weeks. This resulted in increased life span of more than 150%. HMAF has also been found to exhibit outstanding activity against breast, colon, and skin cancer cell lines derived from human tumors.[1,3-9]

Scheme 1. Studies on the mechanism of action of illudins.
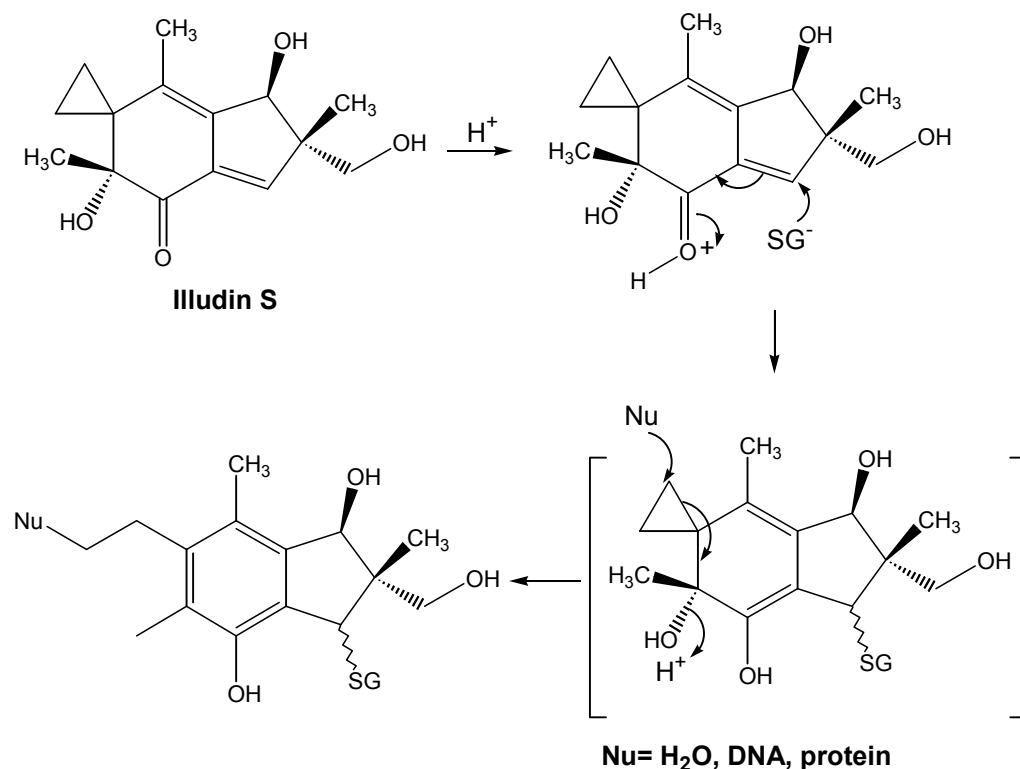
Biologically active natural products will have a profound effect on the field of medicine in the 21st century. However, these molecules are commonly isolated in small quantities and, therefore, cannot be subjected to advanced biological testing whereby the mechanism by which they act can be understood. Chemical synthesis can produce these compounds in the mass quantities necessary for intensive research and therapy.

Retrosynthesis

Several syntheses of structurally related compounds have appeared in the chemical literatures,[10-13] but these syntheses are not suitable for large scale production. A retrosynthetic analysis of the total synthesis of Hydroxymethylacylfulvene **1** (HMAF, also MGI 114) is illustrated in Scheme **2**. Starting from 2-(3-methoxyphenyl)ethanol, our strategy should lead to an efficient and versatile synthesis of HMAF and provide access to several derivatives of these products in sufficient amounts for further biological testing.

The key features of this synthesis are: (1)Nazarov reaction to form bicyclic system (**6**);[14] (2) Enantioselective aromatic oxidation chemistry to install the free hydroxy at the chiral center of intermediates (**11**); (3) Tandem oxidation/cyclopropanation at the final stage to afford HMAF efficiently.

Scheme 2. Retrosynthesis of Hydroxymethylacylfulvene

HMAF
1

tandem oxidation/cyclopropanation

enantioselective aromatic oxidation

11

Nazarov reaction

6
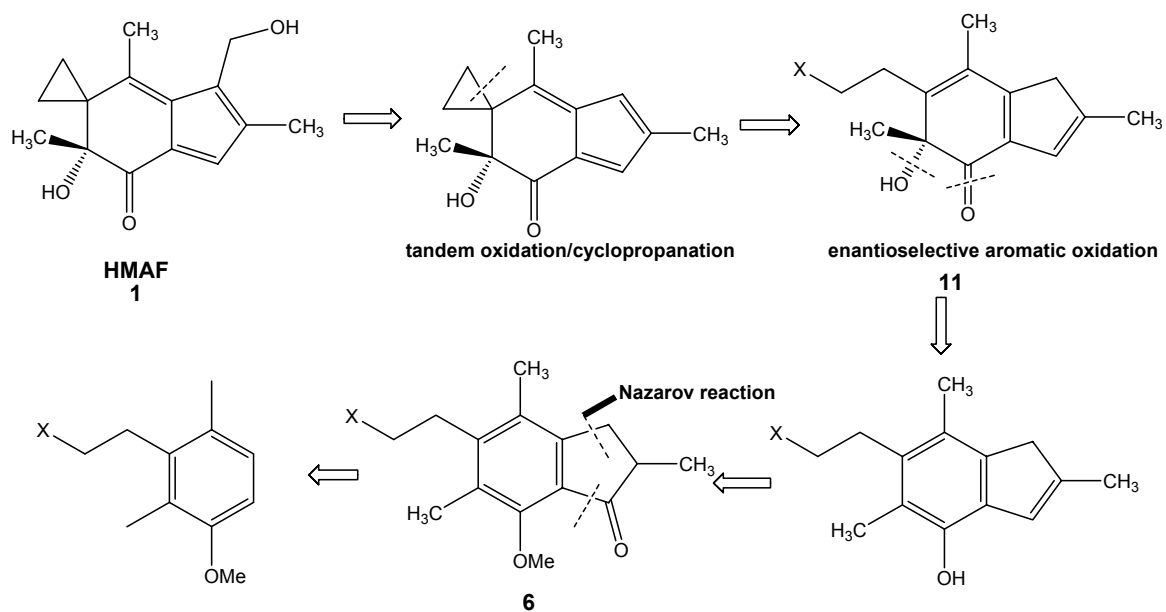
Results and Discussions

Starting from (3-methoxy)-2-phenylethanol, compound (**3**) was synthesized as described in the literature.[15] The two methyl groups were introduced to the aromatic ring by Friedel-Crafts alkylation, followed by LAH or Li/NH$_3$ reduction (Scheme 3).

Scheme 3. Preparation of the Aromatic Subunit of HMAF.[7]
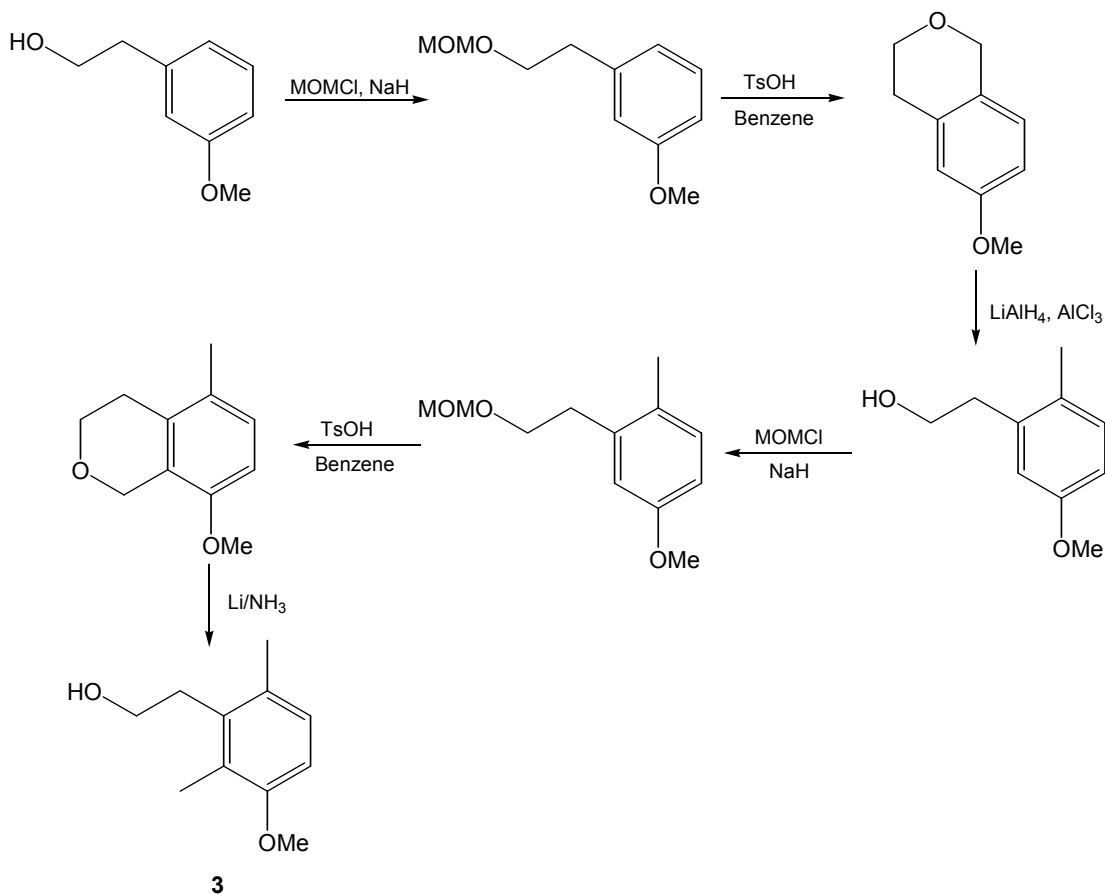
After benzyl protection of the hydroxyl group in 3-methoxy-2,6-dimethyl phenylethanol (**3**), selective α-bromization was achieved using NBS and silica gel

condition. Allylic alcohol (**4**) was formed by adding *t*-BuLi and methacrolein. Swern or

Dess Martin oxidation gave α,β-unsaturated ketone (**5**) which underwent Nazarov

reaction to form bicyclic system (**6**)(Scheme 4).

Scheme 4. Advancing the Aromatic Subunit---Nazarov Reaction
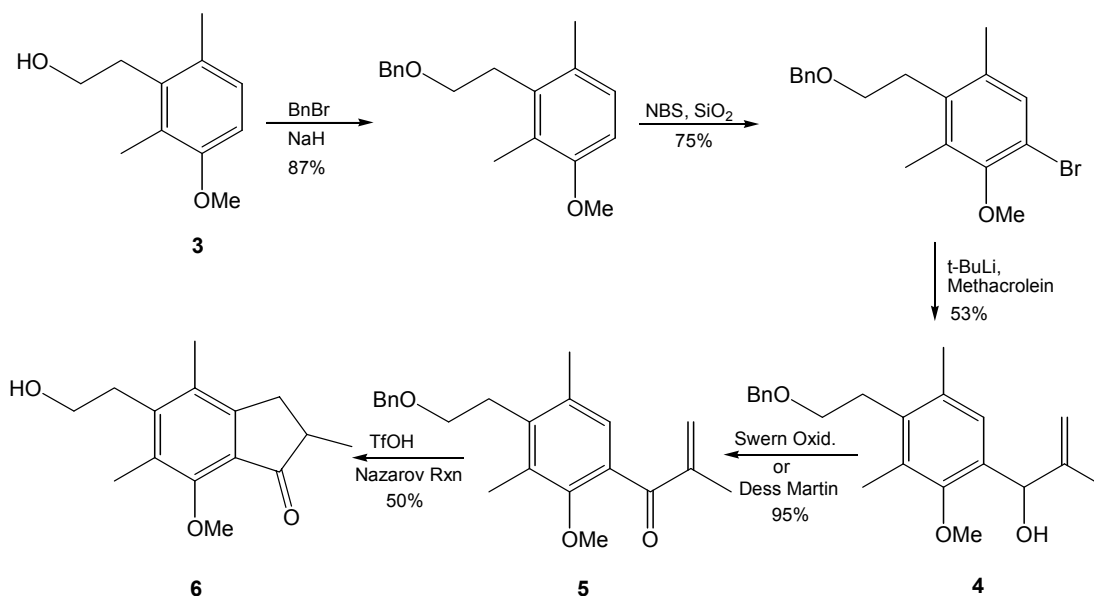
The free hydroxyl group of the compound (**6**) was protected by a silyl group.

Ketone-alcohol-alkene transformation was achieved by reduction using $NaBH_4$ followed

by treatment with Burgess' reagent, or MsCl and Et_3N. Phenol (**10**) was obtained by

demethylation using BBr_3. Free hydroxy at the chiral center of intermediate (**11**) was

designed to be installed by enantioselective aromatic oxidation chemistry, using

$Pb(OAc)_4$ and chiral ligand followed by deacetylation. Instead of screening chiral ligand,

enzymatic deacetylation will be a good way to achieve enantioselectivity (Scheme 5).

Scheme 5. Advancing the aromatic Subunit---Enantioselective aromatic oxidation
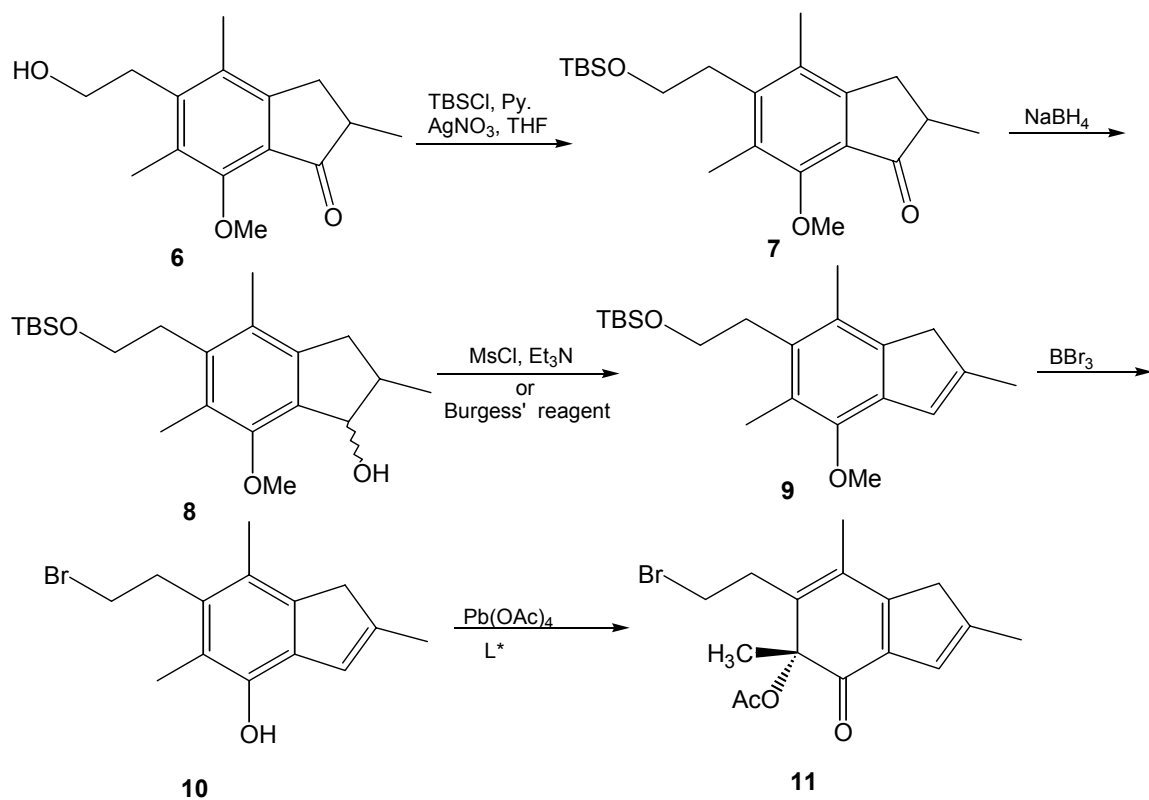
Final stage manipulation involves screening different base to achieve tandem oxidation/cyclopropanation to give known intermediate acylfulvene (**12**) which undergoes an ene reaction with formaldehyde to afford HMAF (**1**) (Scheme 6).

Scheme 6. Tandem oxidation/cyclopropanation

HMAF

1
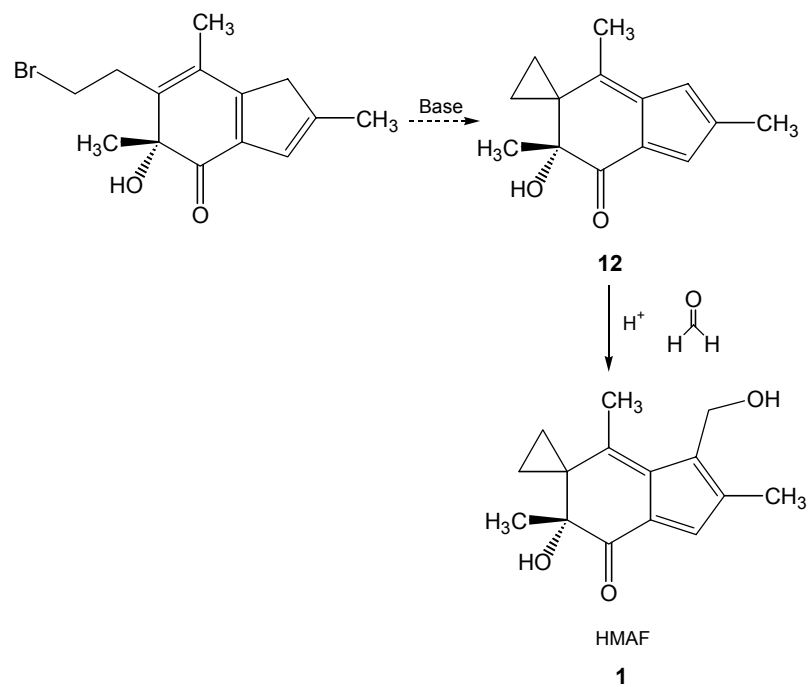
Conclusions

This research demonstrated a facile, efficient, and inexpensive synthesis for HMAF. In addition, a variety of acylfulvene analogues and their precursors will be synthesized and subsequently tested for biological activity. Although the HMAF and related molecules are known to have biological activity, the mechanisms by which they act remain to be further understood. The development of an efficient total synthesis of HMAF and its analogues would allow for the production of large quantities of each and thus render advanced biological testing of these molecules possible. The results of these experiments could establish precisely how these molecules function. Once the key molecular interactions are known, it should be possible to design molecules that have improved biological activity, bringing the world closer to uncovering the cure for cancer.

References

1. MacDonald, J. R., Muscoplay, C. C., Dexter, D. L., Mangold, G. L., Chen, S.-F., Kelner, M. J., McMorris, T. C., Von Hoff, D. D. Preclinical antitumor activity of 6-hydroxymethylacylfulvene, a semisynthetic derivative of the mushroom toxin illudin S. *Cancer Res* **57**, 279-283 (1997).

2. Kelner, M. J., McMorris, T. C., Estes, L. A., Wang, W., Samson, K. M., Taetle, R. Efficacy of HMAF (MGI-114) in the MV522 metastatic lung carcinoma xenograft model nonresponsive to traditional anticancer agents. *Investl New Drugs* **14**, 161-167 (1996).

3. McMorris, T. C., Yu, J., Estes, L. A. & Kelner, M. J. Reaction of antitumor hydroxymethylacylfulvene (HMAF) with thiols. *Tetrahedron* **53**, 14579-14590 (1997).

4. McMorris, T. C., Yu, J., Gantzel, P. K., Estes, L. A. & Kelner, M. J. An acetal derivative of illudin S with improved antitumor activity. *Tetrahedron Lett* **38**, 1697-1698 (1997).

5. Mcmorris, T. C., Kelner, M. J., Wang, W., Estes, L. A., Montoya, M. A., Taetle, R. Structure-Activity-Relationships of Illudins - Analogs with Improved Therapeutic Index. *J Org Chem* **57**, 6876-6883 (1992).

6. Kelner, M. J., Mcmorris, T. C. & Taetle, R. Preclinical Evaluation of Illudins as Anticancer Agents - Basis for Selective Cytotoxicity. *J Natl Cancer Inst* **82**, 1562-1565 (1990).

7.   Mcmorris, T. C. et al. Structure and Reactivity of Illudins. *Tetrahedron* **45**, 5433-5440 (1989).

8.   Kelner, M. J., Mcmorris, T. C., Beck, W. T., Zamora, J. M. & Taetle, R. Preclinical Evaluation of Illudins as Anticancer Agents. *Cancer Res* **47**, 3186-3189 (1987).

9.   Woynarowski, J. M. et al. Effects on DNA integrity and apoptosis induction by a novel antitumor sesquiterpene drug, 6-hydroxymethylacylfulvene (HMAF, MGI 114). *Biochem Pharma* **54**, 1181-1193 (1997).

10.   McMorris, T. C., Yu, J., Hu, Y., Estes, L. A. & Kelner, M. J. Design and synthesis of antitumor acylfulvenes. *J Org Chem* **62**, 3015-3018 (1997).

11.   McMorris, T. C., Hu, Y., Yu, J. & Kelner, M. J. Total synthesis of hydroxymethylacylfulvene, an antitumour derivative of illudin S. *Chem Comm*, 315-316 (1997).

12.   Brummond, K. M., Lu, J. L. & Petersen, J. A rapid synthesis of hydroxymethylacylfulvene (HMAF) using the allenic Pauson-Khand reaction. A synthetic approach to either enantiomer of this illudane structure. *J Am Chem Soc* **122**, 4915-4920 (2000).

13.   Brummond, K. M. & Lu, J. L. A short synthesis of the potent antitumor agent (+/-)-hydroxymethylacylfulvene using an allenic Pauson-Khand type cycloaddition. *J Am Chem Soc* **121**, 5087-5088 (1999).

14.   Suzuki, T., Ohwada, T. & Shudo, K. Superacid-catalyzed electrocyclization of 1-phenyl-2-propen-1-ones to 1-indanones. Kinetic and theoretical studies of

electrocyclization of oxonium-carbenium dications. *J Am Chem Soc* **119**, 6774-6780 (1997).

15.   Meyer, A. L. & Turner, R. B. Interesting Synthesis of 3-Methoxy-2,6-Dimethylphenethyl Alcohol. *Tetrahedron* **27**, 2609-& (1971).

Curriculum Vitae

Xin Qi
Division of Chemistry and Chemical Engineering
Mail Code 147-75 CH
California Institute of Technology
1200 E. California Blvd.
Pasadena, CA 91125

EDUCATION:
 Ph.D. Organic Chemistry/Biochemistry,   expected 5/05
 California Institute of Technology, Pasadena, CA
 M.S. / Ph.D. Candidate, Organic Chemistry,   *5/00*
 Wayne State University, Detroit, MI
 B.S. Polymer Chemistry,   *7/91*
 Peking University, Beijing, P. R. China

INDUSTRIAL EXPERIENCE:
 5 years pharmaceutical research experience.   *8/91-8/96*
 Beijing Pharmaceutical Company
 Engineer in pilot synthesis and quality control

RESEARCH EXPERIENCE:
 Graduate research, Division of Chemistry and Chemical Engineering, *9/00-present*
 California Institute of Technology
 Professor R. W. Roberts, principal investigator
- Synthesis and binding activity of peptide-acridine conjugates directed against RNA targets
- Synthesis of tRNA synthetase inhibitors and incorporation of non-natural amino acid residues using chemically aminoacylated tRNAs recognizing specific codons
- The Puromycin route to assess stereo- and regiochemical constraints on peptide bond formation in eukaryotic ribosomes
- Total synthesis of Hydroxymethylacylfulvene

 Graduate research, Department of Chemistry, Wayne State University   *8/96-5/00*
 Professor J. Montgomery, principal investigator
- Developed a nickel-catalyzed three-component synthesis of 1,3 dienes
- Synthesized indolizidine alkaloids by the nickel-catalyzed cyclization of ynals

Undergraduate research, Department of Chemistry, Peking University   *9/87-7/91*
Professor Shuojian Jiang, principal investigator
- Studied on synthesis and curing conditions of epoxy curing agents

AWARDS:
Thomas C. Rumble Fellowship — Wayne State University   *1997*
Achievement Award for the Sino-UK Technical Joint Venture Program   *1993*
— Beijing Pharmaceutical Company

PROFESSIONAL SOCIETIES:
Phi Lambda Upsilon   American Chemical Society

PRESENTATION:

218th ACS National Meeting, New Orleans, LA, August 22-26, 1999. New Three Component Synthesis of 1,3-Dienes Employing Nickel Catalysis.

PUBLICATIONS:
1.   **Xin Qi,** Richard W. Roberts, "Synthesis and binding activity of λ N peptide-acridine conjugates directed against *BoxB* RNA" *Biochemistry, manuscript in preparation.*
2.   **Xin Qi**, Shelley R. Starck, Richard W. Roberts, "Creating Synthetic "Blanks" in the Genetic Code Using Aminoacyl tRNA synthetase inhibitors" *J. Am. Chem. Soc., manuscript in preparation.*
3.   Rong-Gui Hu, Jun Sheng, **Xin Qi**, Zhenming Xu, Susanne M. Mumby, Terry T. Takahashi, Alexander Varshavsky, "Arginylation branch of the N-end rule pathway: a nitric oxide sensor controlling the levels of multiple regulators" *Nature* **2005**, *in revision.*
4.   Shelley R. Starck, **Xin Qi**, Jennifer B. Treweek, Binghai Ling, Richard W. Roberts, "Unexpected Flexibility of Protein Synthesis In Vitro and In Vivo" *manuscript in preparation.*
5.   Shelley R. Starck, **Xin Qi**, Brett N. Olsen, Richard W. Roberts, "The Puromycin Route to Assess Stereo- and Regiochemical Constraints on Peptide Bond Formation in Eukaryotic Ribosomes" *J. Am. Chem. Soc.* **2003**, *125,* 8090-8091.
6.   **Xin Qi**, John Montgomery, "New Three Component Synthesis of 1,3-Dienes Employing Nickel Catalysis" *J. Org. Chem.* **1999**, *64*, 9310-9313.