ACCOUNTS
of chemical research

Flexizymes: Their Evolutionary History and the Origin of Catalytic Function

JUMPEI MORIMOTO,[†,‡] YUUKI HAYASHI,[†,§]
KAZUHIRO IWASAKI,[†,‡] AND HIROAKI SUGA[*,†,‡,§]

[†]Department of Chemistry, School of Science, [‡]Department of Chemistry and Biotechnology, School of Engineering, and [§]Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

RECEIVED ON MARCH 28, 2011

CONSPECTUS

Transfer RNA (tRNA) is an essential component of the cell's translation apparatus. These RNA strands contain the anticodon for a given amino acid, and when "charged" with that amino acid are termed aminoacyl-tRNA. Aminoacylation, which occurs exclusively at one of the 3′-terminal hydroxyl groups of tRNA, is catalyzed by a family of enzymes called aminoacyl-tRNA synthetases (ARSs). In a primitive translation system, before the advent of sophisticated protein-based enzymes, this chemical event could conceivably have been catalyzed solely by RNA enzymes. Given the evolutionary implications, our group attempted in vitro selection of artificial ARS-like ribozymes, successfully uncovering a functional ribozyme (r24) from an RNA pool of random sequences attached to the 5′-leader region of tRNA. This ribozyme preferentially charges aromatic amino acids (such as phenylalanine) activated with cyanomethyl ester (CME) onto specific kinds of tRNA.

During the course of our studies, we became interested in developing a versatile, rather than a specific, aminoacylation catalyst. Such a ribozyme could facilitate the preparation of intentionally misacylated tRNAs and thus serve a convenient tool for manipulating the genetic code. On the basis of biochemical studies of r24, we constructed a truncated version of r24 (r24mini) that was 57 nucleotides long. This r24mini was then further shortened to 45 nucleotides. This ribozyme could charge various tRNAs through very simple three-base-pair interactions between the ribozyme's 3′-end and the tRNA's 3′-end. We termed this ribozyme a "flexizyme" (Fx3 for this particular construct) owing to its flexibility in addressing tRNAs.

To devise an even more flexible tool for tRNA acylation, we attempted to eliminate the amino acid specificity from Fx3. This attempt yielded an Fx3 variant, termed dFx, which accepts amino acid substrates having 3,5-dinitrobenzyl ester instead of CME as a leaving group. Similar selection attempts with the original phenylalanine-CME and a substrate activated by (2-aminoethyl)amidocarboxybenzyl thioester yielded the variants eFx and aFx (e and a denote enhanced and amino, respectively). In this Account, we describe the history and development of these flexizymes and their appropriate substrates, which provide a versatile and easy-to-use tRNA acylation system. Their use permits the synthesis of a wide array of acyl-tRNAs charged with artificial amino and hydroxy acids.

In parallel to these efforts, we initiated a crystallization study of Fx3 covalently conjugated to a microhelix RNA, which is an analogue of tRNA. The X-ray crystal structure, solved as a co-complex with phenylalanine ethyl ester and U1A-binding protein, revealed the structural basis of this enzyme. Most importantly, many biochemical observations were consistent with the crystal structure. Along with the predicted three regular-helix regions, however, the flexizyme has a unique irregular helix that was unexpected. This irregular helix constitutes a recognition pocket for the aromatic ring of the amino acid side chain and precisely brings the carbonyl group to the 3′-hydroxyl group of the tRNA 3′-end. This study has clearly defined the molecular interactions between Fx3, tRNA, and the amino acid substrate, revealing the fundamental basis of this unique catalytic system.

History of Aminoacylation Ribozymes

In the "modern world" translation system, aminoacylation of the 3′-terminus of tRNA is catalyzed by a family of protein enzymes, aminoacyl-tRNA synthetases (ARSs). On the other hand, the present evidence from the crystal structural as well as biochemical studies of ribosome have revealed that

Published on the Web 06/28/2011 www.pubs.acs.org/accounts
10.1021/ar2000953

ribosome's catalytic center consists of only RNA, that is, ribosome is a ribozyme. This suggests that a primitive translation catalytic system, including not only ribosome but also ARSs, could have consisted of entirely RNA molecules. However, naturally occurring ARS ribozymes are yet unknown and thus the above hypothesis is not fully supported by the available knowledge from nature. Because sophisticated protein enzymes such as ARSs could not be evolved before the advent of the translation system, it is critical to see if RNA molecules are able to catalyze aminoacylation.

In vitro selection (or SELEX) is a powerful technique that aims at isolating functional RNAs from a pool of random sequences of RNA. In fact, in the last nearly two decades, we have witnessed the discovery of many artificial ribozymes capable of catalyzing various chemical reactions in vitro even though such naturally occurring ribozymes are unknown in the modern life. Among them, a few research teams have successfully isolated ribozymes capable of charging certain amino acids onto RNA. One of the earlier pioneer works was reported by Yarus et al. where they isolated artificial ribozymes that catalyze self-aminoacylation of their own CCG 3′-terminal 2′/3′-OH with Phe-AMP or Tyr-AMP[1,2] as an aminoacylation donor. In 1998, Famulok and Jenne also reported an artificial rybozyme that catalyzes self-aminoacylation with *N*-biotin-Phe-AMP, but the site of aminoacylation is not 3′-terminal 2′/3′-OH but an internal 2′-OH.[3] Both classes of ribozymes were unable to aminoacylate onto the 3′-end of tRNA or tRNA-like CCA, and therefore, a major question remained unanswered whether ribozymes are able to charge amino acids onto the specific site of 3′-terminal CCA end of tRNA or tRNA molecules.

Meanwhile, Szostak and Lohse reported an acyl-transferase ribozyme (ATRib) capable of transferring *N*-biotin-Phe (biotin is a selectable tag using streptavidin resin) from the 3′-end of a short RNA to its own 5′-hydroxyl group.[4] Because this ribozyme was originally aimed at mimicking the function of ribosome as a peptidyl transfer catalyst, the acyl-donor RNA was designed to be 5′-CAACCA-3′ as a mimic of tRNA's 3′-end and the 3′-end of ATRib has an internal guide sequence complementary to this RNA sequence. Using the microreversibility of acyl-transfer reaction, Suga et al. turned ATRib into two new ribozymes that are able to charge amino acids onto the 3′-terminus of tRNAs. A ribozyme, reported in 2000 and referred to as AD02,[5] has a 70 nucleotide (nt) accessory domain at the 3′-terminus of ATRib and catalyzes two steps of reactions; in the first step, AD02's accessory domain catalyzes self-aminoacylation of *N*-biotin-Gln assisted by cyanomethyl ester (CME) onto the 5′-OH group, and in the second step the 5′-acyl group is transferred to the 3′-end of tRNA upon binding to the ATRib internal guide sequence. Although the catalytic ability of AD02 is very modest where only 4% of tRNA is aminoacylated, this ribozyme showed remarkable specificity toward Gln side chain over Met (0.047), Leu (0.004), Phe (<0.001), and Val (<0.001) (the values in parenthese are relative specificity constants).[6] This work represents the first example of a tRNA aminoacylation catalyst consisting of only a RNA scaffold. The second ribozyme, reported in 2002 and referred to as BC28,[7] also has an accessory domain at the 3′-terminus of ATRib and a loop, referred to as anti-anticodon (AC) loop, embedded in this domain, and recognizes a specific AC loop sequence of tRNA by forming six base pairs. BC28 accepts various kinds of amino acids from the 3′-end of 5′-AACCA-3′ (note that one complementary base less than the original donor substrate) to its own 5′-terminal OH group such as ATRib, and also transfers back to the 3′-end of a specific tRNA designated by the interaction between the anti-AC loop and tRNA's AC loop, yielding up to 17% aminoacyl-tRNA. Interestingly, a base mutation(s) in the anti-AC loop is able to reprogram the specificity of BC28 that aminoacylates a desired tRNA with a complete match of six base pairs over other tRNAs containing a mispair(s). Although these ribozymes consisting of the ATRib scaffold endorse some of unique characteristic functions as ARSs, their modest efficiencies of tRNA aminoacylation due to their complex mechanisms involving the equilibrium shift of acyl-transfer chemistry dismissed further development of these ribozymes.

In Vitro Evolution of Flexizyme

In prokaryotes, precursor tRNAs, in which 5′-leader accessory sequences are attached to the 5′-end of the body of tRNA, are transcribed and processed by ribonucleoprotein enzymes, RNase P, to yield mature tRNAs onto which amino acids are charged by cognate ARSs (Figure 1A). The catalytic component of RNase P is known to be RNA (referred to as M1 RNA), representing one of the molecular fossils of the RNA world. One can hypothesize that even though the 5′-leader sequences have no function in the "modern world", they could have catalytic ability for self-aminoacylation of the 3′-terminus of tRNA sequences and later be removed by a M1 RNA-like ribozyme in the "RNA world" (Figure 1B). This hypothesis is also able to give a simple explanation on how such 5′-leader ribozymes could specify cognate tRNAs, that is, they are covalently linked to cognate tRNAs for self-aminoacylation, and thus the specificity is set in the respective precursor tRNAs. We thus decided to set an experiment

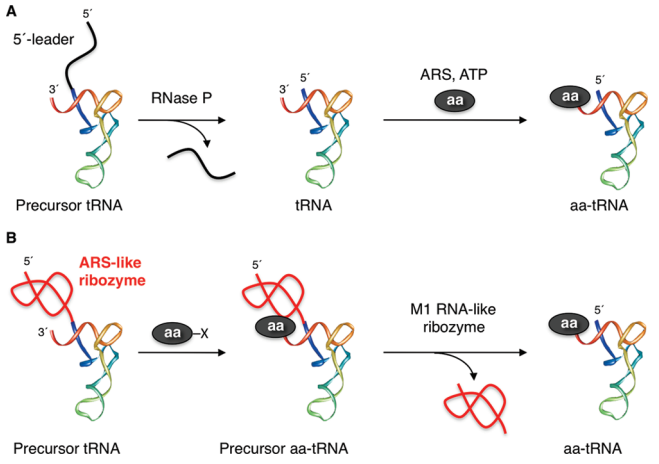

FIGURE 1. Comparison of (A) the current aa-tRNA maturation pathway and (B) a possible early aa-tRNA maturation pathway. aa denotes amino acid. (A) 5′-leader sequences of precursor tRNAs are removed by a RNA–protein complex, called RNase P, and then the 3′-terminal OH group is aminoacylated by protein ARSs. (B) ARS-like ribozymes could have existed in 5′-leader sequences of precursor tRNAs. They might have catalyzed aminoacylation of their own 3′-terminal OH and then be removed by M1-like ribozyme to yield mature aminoacyl-tRNAs. Note that all the components other than amino acid are composed of RNA molecules in this possible primitive pathway.

to discover such a catalytic precursor tRNA by means of in vitro selection.

Our previous selection outcome of AD02, in which the 3′-region of accessory domain was able to charge Gln onto the 5′-OH group of ATRib using the corresponding CME donor, particularly encouraged us to design a new selection scheme for the discovery of catalytic precursor tRNAs. We thus constructed a precursor tRNA library bearing random RNA sequences (20-nt 5′-primer region followed by 70-nt random) at the 5′-leader region of a chosen tRNA (Figure 2A) and performed in vitro selection to isolate RNA species capable of self-aminoacylating in the presence of *N*-biotin-L-phenylalanine cyanomethyl ester (*N*-biotin-Phe-CME) (Figure 3A). This selection campaign successfully yielded a single family of ribozyme, referred to as pre-24, that specifically charged *N*-biotin-Phe onto the 3′-terminal OH group(s) of the tRNA region.[8] Moreover, its 90-nt catalytic domain, referred to as r24, could be truncated down to 57-nt r24mini (Figure 2B) without any loss of catalytic activity.[9]

Biochemical studies on r24 revealed several intriguing features: r24 and r24mini (1) have an internal guide sequence, GGU, in L3 that forms complementary base pairs with the $\underline{A}^{73}\underline{CCA}^{76}$ (pairing bases are underlined);[9] (2) could be disconnected from tRNA, becoming a trans-acting ribozyme;[8] (3) could use not only *N*-biotin-Phe-CME but also Phe-CME (Figure 4A), Phe-AMP, and Phe-thioester;[8] (4) could selectively charge Phe onto the 3′-OH group, not 2′-OH, at the tRNA 3′-terminus;[10] and (5) of which critical bases in unpaired joining regions (J1/2 and J2/3) for Phe binding were defined by chemical mappings.[9]

Although the specificity toward tRNA and Phe achieved in r24mini are essential features as an ARS-like ribozyme in the evolutionary point of view, during the course of our studies we had become more interested in developing a versatile aminoacylation catalyst rather than a specific one. We envisioned that such a ribozyme should allow us to prepare misacylated tRNAs at our will and thus serve as a convenient tool for manipulating the genetic code in translation. With such an application in mind, we decided to turn our effort to the evolution of more versatile catalysts. Based on r24mini, we constructed a doped RNA pool where bases in P3 and L3 were randomized (Figure 2C). Bases composing J2/3 and $U^{40}U^{41}$ in L3 (underlined bases in Figure 2B, C) and 5′-GGU-3′ in L3 were kept as the original sequences because their importance in the recognition of amino acid substrate and tRNA, respectively, was supported by biochemical studies.[9] With this RNA pool, we performed in vitro selection of active species. The outcome of active sequences showed that the bases in the top strand of P3 were completely conserved with the original bases whereas those beyond GGU in L3 had no sequence similarity in active species. These results suggested that the nonconserved bases are unnecessary for catalytic activity. To this end, r24mini was further truncated to a 45-nt construct (Figure 2D). This new shorter construct exhibited better activity in terms of the yield of Phe-tRNA product in trans, exhibiting k_{cat} of 0.15 min^{-1} and K_M to tRNA of 5 μM with an ability of multiple turnovers up to 16 times. Moreover, it was able to charge a variety of aromatic amino acids, including Phe analogues, onto tRNAs with A, G, and U at position 73. We thus referred to this construct of ribozyme as flexizyme3[11] (Fx3; 3 stands for the number of base pair interactions with the tRNA 3′-end).

Although this prototype Fx3 acquired versatility toward tRNAs, it was not flexible toward amino acid substrates; it was capable of charging only aromatic amino acids such as Phe and their derivatives onto tRNA. It should be noted that Fx3 retained all properties described for r24mini. As a result, we hypothesized that the critical recognition element could be embedded in the aromatic side chain (benzyl group) of the amino acid substrate. To overcome the limitation of substrate recognition of Fx3, we decided to perform in vitro evolution of Fx3 variants that are able to accept more diverse kinds of side chains in the substrate.[12] To increase the chance of success, we set two designs in the selection

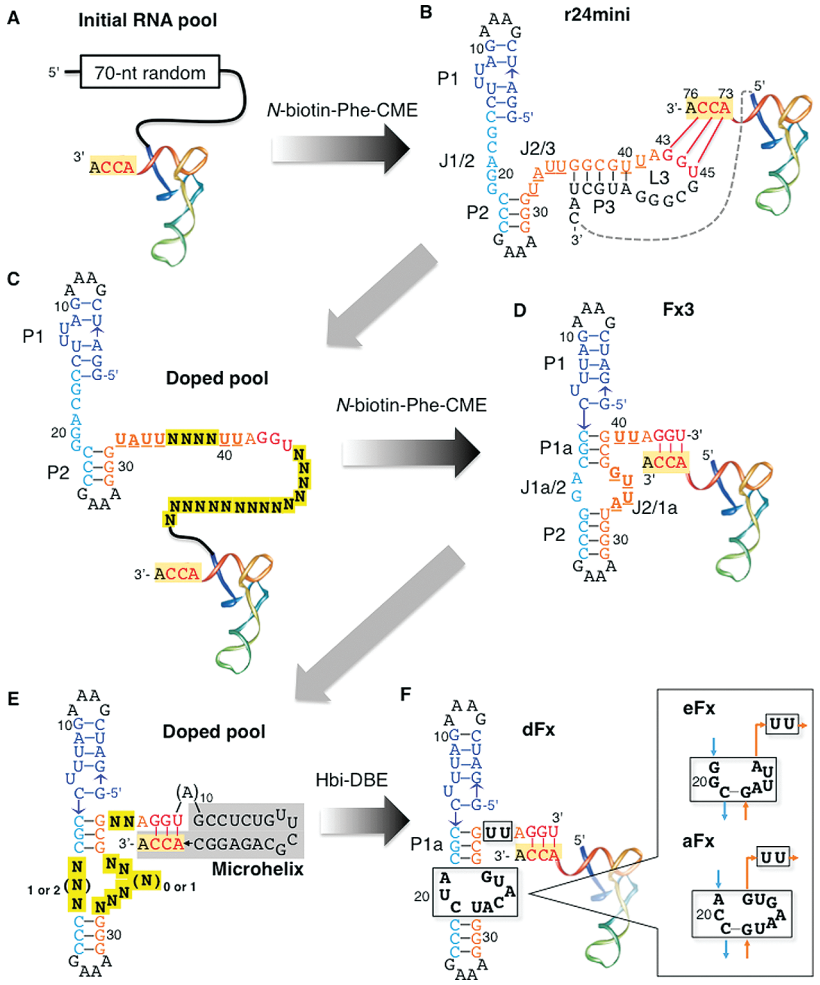

FIGURE 2. Diagram summarizing the evolutionary history of flexizyme. Black arrows indicate in vitro selection and sequence optimization steps. Amino acid substrate used for selection is written above the arrows. Gray arrows indicate the construction of doped pool (partially randomized RNA sequences). The secondary structures of (A), (B), and (C) are determined by biochemical studies. Those of (D), (E), and (F) are described based on the crystal structural analysis. In (B), (C), and (D), underlined bases are critical for binding sites of amino acid substrate determined by biochemical data.[9] In (C) and (E), randomized bases are highlighted as bold characters on a yellow background. In (F), sequences derived from randomized positions are highlighted as bold characters in a rectangle box. (A) Initial RNA pool for in vitro selection of ARS-ribozymes. (B) First generation of the artificial ARS-ribozyme. (C) Doped RNA pools constructed based on the sequence of r24mini. (D) Second generation ARS-ribozyme, that is flexible toward tRNA, referred to as Fx3. (E) Doped RNA pool constructed based on the sequence of Fx3. Bases of microhelix RNA are highlighted by a gray background. (F) Third generation ARS-ribozymes. The entire sequences of dFx is shown and sequences derived from randomized positions of eFx and aFx are in boxes on the right.

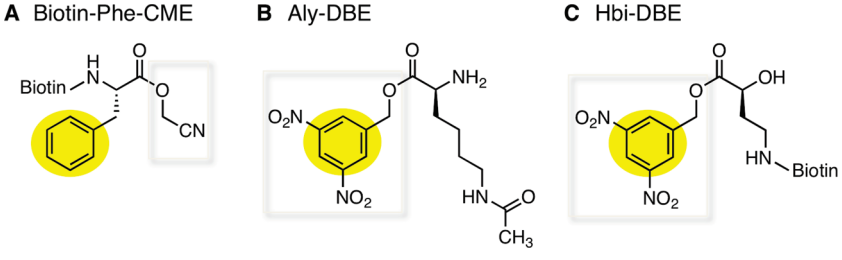

FIGURE 3. Amino acid substrates used in the evolutional attempts described in this Account. Aromatic rings accommodated in flexizymes are highlighted with yellow circles. Leaving groups are highlighted by gray boxes. (A) *N*-Biotin-phehylalanine cyanomethyl ester (Biotin-Phe-CME). (B) ε-(*N*-acetyl)lysine 3,5-dinitrobenzyl ester (Aly-DBE). (C) δ-(*N*-Biotinyl)Amino-α-(*S*)-hydroxybutanoic acid 3,5-dinitrobenzyl ester (Hbi-DBE).

strategy: (1) In order to alter Fx3's recognition element from the side chain to a generic group of substrate, the benzyl group was embedded in the leaving group; and (2) in order to alter the amino acid substrate recognition of Fx3, random

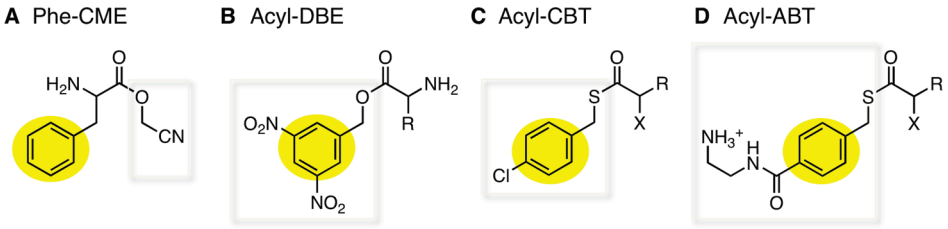

FIGURE 4. Generic structure of amino acid substrate for flexizymes. Aromatic ring accommodated in flexizymes is highlighted with a yellow circle. Leaving group is shown in a gray rectangle. (A) Phenylalanine cyanomethyl ester (Phe-CME), suitable substrate for r24mini, Fx3, and eFx. (B) Amino acid dinitrobenzyl ester (Acyl-DBE), suitable substrate for dFx. (C) Amino acid 4-chlorobenzyl thioester (Acyl-CBT), suitable substrate for eFx. (D) Amino acid (2-aminoethyl)amidocarboxybenzyl thioester (Acyl-ABT), suitable substrate for aFx.

sequences were introduced to putative bases interacting with the aromatic side chain. To see if the strategy of (1) could be valuable, we synthesized ε-*N*-acetyl-lysine 3,5-dinitrobenzyl ester (Aly-DBE) (Figure 3B) where the cyanomethyl ester group was replaced with 3,5-dinitrobenzyl ester and tested it using Fx3. Although the yield of aminoacyl-tRNA was poor (~3%), Aly-DBE was accepted as an amino acid substrate of Fx3 in contrast to Aly-CME that was an inert substrate of Fx3.

This result motivated us to perform in vitro evolution of flexizyme mutants capable of more efficiently using acyl-DBE. We thus constructed a library of Fx3–microhelix RNA conjugate in which the bases involved in amino acid recognition of Fx3 were randomized (Figure 2E) and performed in vitro selection of flexizyme mutants capable of charging microhelix with δ-(*N*-biotinyl)amino-α-(*S*)-hydroxybutanoic acid DBE (Hbi-DBE) (Figure 3C). This experiment turned Fx3 to a new family of ribozymes capable of catalyzing aminoacylation onto not only microhelix RNA but also tRNAs in cis as well as trans with much higher efficiencies. One of such a 46-nt ribozyme, referred to as dinitro-flexizyme (dFx, Figure 2F), was found to be tolerant with alteration of a variety of side chains, chirality, *N*-acyl/alkyl substitutions, and non-α-amino acid structures as far as they are esterified with DBE, for example, artificial α-L-amino acids, α-D-amino acids,[13] α-*N*-acyl amino acids,[14] α-*N*-alkyl amino acids,[15,16] β-amino acids (unpublished data), and α-hydroxy acids[17] (Figure 5A–F). In parallel to the above selection, we also performed a selection to improve Fx3 catalytic activity using *N*-biotin-Phe-CME. This selection yielded enhanced flexizyme (eFx) (Figure 2F) that shows higher catalytic activity toward various aromatic amino acid CMEs. Coincidentally, we later found that eFx surprisingly accepts substrates activated by chlorobenzyl thioester (CBT) (Figure 4C). Because the benzyl group is embedded in the leaving group of the substrate, eFx turns out to be also a versatile catalyst that accepts those with nonproteinogenic side chains.

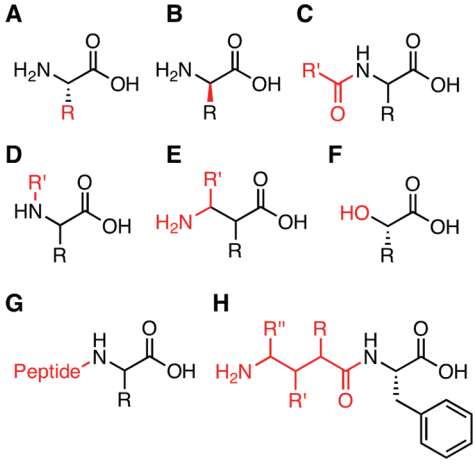

FIGURE 5. Representative examples of nonproteinogenic and artificial amino acids that have been tested for flexizymes. Nonstandard structures are highlighted in red. (A) L-Amino acid with nonproteinogenic side chain. (B) D-Amino acid. (C) α-*N*-acyl-amino acid. (D) α-*N*-Alkyl-amino acid. (E) β-Amino acid. (F) α-Hydroxy acid. (G) α-*N*-Peptidyl-amino acid. (H) γ-Aminoacyl-amino acid.

Having these two flexizymes, dFx and eFx, enables us to prepare a wide variety of acyl-tRNAs. The dFx/acyl-DBE system is our generic system for tRNA aminoacylation, since (1) synthesis of acyl-DBEs is simple, and (2) dFx is active against most acyl-DBEs dissolved in a generic reaction buffer containing some portions of DMSO if necessary. On the other hand, some acyl-DBEs could be poor or insufficient substrates of dFx, for instance, due to their steric hindrance of side chain in the combination of *N*-alkyl modification. Then, the eFx/acyl-CBT system is a preferable choice to increase the substrate reactivity, resulting in higher yields of acyl-tRNAs.

More recently, we have added another variant of flexizymes, referred to as aFx ("a" stands for amino) (Figure 2F), that uses amino acids that are activated by (2-aminoethyl)-amidocarboxybenzyl thioester (ABT) (Figure 4D). This leaving group has a primary amino group that is protonated in the generic reaction buffer at pH 7.5–8.5 to give the

ammonium salt.[18] Even though dFx and eFx function with many substrates activated with DBE and CBT, due to intrinsic hydrophobicity of the leaving groups we have encountered cases where the substrate activated by these leaving groups are poorly soluble in such a generic reaction buffer. In such cases, we first attempt to use DMSO as a cosolvent to dissolve the substrate since dFx and eFx still function up to 40% and 30% DMSO, respectively. However, DMSO occasionally fails to rescue the solubility of the substrate. In this case, the aFx/aminoacyl-ABT system often overcomes such an insolubility problem since the ABT group is able to increase water-solubility of substrate.

The choice of flexizymes in the combination of cognate leaving groups on the substrate gives us opportunities to test many different acyl-substrates; thus far, more than 300 kinds of nonstandard structures of amino and hydroxy acids (Figure 5A−F) have been charged onto tRNAs and tested for the incorporation into polypeptide chain using in vitro translation systems in our laboratory. Moreover, we have found that the flexizyme system also charges polypeptides up to pentapeptides, including those containing α-D-amino acids, β-amino acids, and γ-amino acids (Figure 5G and H), onto tRNA. These peptidyl-tRNAs derived from initiator tRNA were used to initiate the translation to yield peptides containing *N*-terminal unique peptides.[19,20]

A virtue of the flexizyme system is the simplicity of procedures. All we need are just a few steps of chemical synthesis of acyl-donor substrate and a few hours of its incubation with a tRNA in the presence of an appropriate flexizyme. Virtually any kind of amino and hydroxy acid could be used in the combination with any tRNAs with a variety of anticodon and body sequences. This flexibility has facilitated the genetic code reprogramming by integration with custom-made cell-free translation systems, recently referred to as flexible in-vitro translation (FIT) systems.[21] Because the primary focus in this Account is on flexizymes, we discuss this system elsewhere.

An Overview of Flexizyme Structure Determined by Crystal Structural Studies

The most intriguing feature of flexizymes is their substrate flexibilities toward amino acids bearing the benzyl (or aromatic) groups residing in the side chain or leaving group and tRNA albeit their small size consisting of only 45 or 46 nt in total length. Earlier biochemical data performed on r24 were fully accountable by the three-dimensional structure of Fx3 generated by the X-ray crystal diffraction data. Therefore, we here concentrate our discussion based on the crystal structure of Fx3.

After nearly 5 years of collaborations with Ferré-D'Amaré's group, we were able to successfully solve the crystal structure of Fx3 tethered to a microhelix RNA, a tRNA acceptor stem analogue.[22] A cocrystallization technique pioneered by Ferré-D'Amaré[23] was utilized in this study, where a U1A-binding loop sequence inserted into the Fx3's functionally dispensable L1 loop was cocrystallized with U1A protein. To define the active site of Fx3, the complex was also cocrystallized with an inhibitor, L-phenylalanine ethyl ester (Phe-EE), instead of active L-Phe-CME substrate (Figure 6B). In its crystal structure, bases are numbered differently from those in the original Fx3 construct, and the following discussion utilizes the numbers indicated in Figure 6A. In this section and the following two sections, A, U, G, or C represents a base in flexizyme and tA, tU, tG, or tC represents a base in tRNA.

Earlier biochemical studies[8] have indicated that flexizyme is composed of two helixes (P1 and P2) and its 3′-terminal three bases form an additional helix (P3) with tRNA. This prediction has turned out to be mostly true in the crystal structure of Fx3 albeit base pairs in P1 stem predicted by biochemical study are slightly different from those observed in the crystal structure. Significantly, two additional "unexpected" short helices, P1a and J1a/2-J2/1a, are newly found in the crystal structure (Figure 6A), showing that four of these helixes (P1, P1a, J1a-J2/1a, and P2) form a large main helical stack. Nucleotides of 52−54 near the 3′-end of flexizyme form a unique hairpin-shaped turn (J1a/3), helping the neighboring P3 stem to direct away perpendicularly from the main helical stack. There are three magnesium ions binding to flexizyme in the crystal structure. This is consistent with the conclusion from the biochemical data showing that there are at least two magnesium ion binding sites.[24] One of the magnesium ions exists in the irregular helix, likely participating in the formation of a catalytic core; and the other two exist in P1 stem, likely stabilizing its structure.

Structural Basis of Amino Acid Recognition

As described above, each flexizyme acquired its versatility through the independent courses of evolution experiments. Although they have different properties in terms of acyl-donor substrate specificities, all flexizymes have the common structural scaffold. Thus, we here focus on discussions of the catalytic domains on the basis of an available crystal structure of Fx3.

Fx3 accepts various phenylalanine derivatives, indicating its critical recognition element for amino acids is embedded in the benzyl (aromatic) side chain. The crystal structure

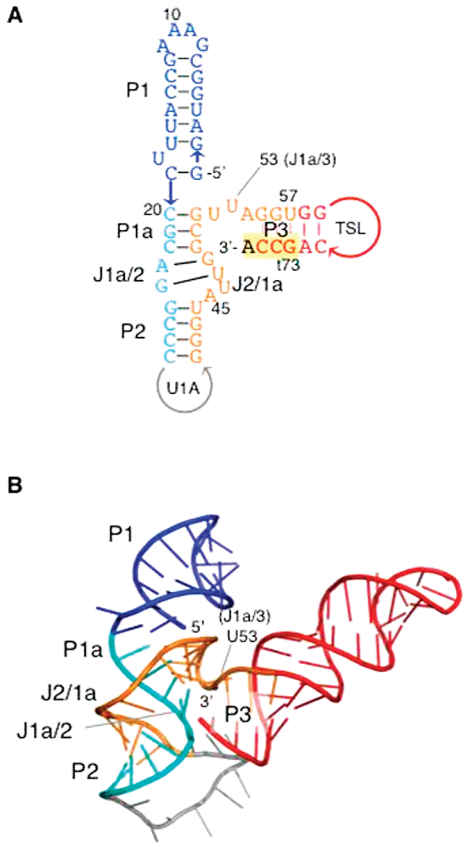

FIGURE 6. Secondary and tertiary structures of Fx3. G1–C7 and C19–G12 compose P1 stem. C20–C22 and G51–G49 compose P1a stem. G26–C28 and U43–G41 compose P2 stem. G55-U57 and C75–G73 compose P3 stem. A23–G25, U44–G48, and U52–A54 compose junctions J1a/2, J2/1a, and J1a/3, respectively. Sequences of flexizyme are drawn using three different colors, and sequences of tRNA are colored in red to facilitate comparison of the secondary structure and the tertiary structure. The U1A loop was inserted to the original Fx3 construct to enhance crystallization of the complex. The microhelix sequences other than 5′- and 3′-sequences were represented simply as TSL (T-stem loop). (A) The secondary structure determined based on the crystal structural analysis. (B) The 3D structure of Fx3 complexed with a microhelix RNA. This figure was generated by the data of accession code 3CUL from the Protein Data Bank using The PyMol Molecular Graphics System.

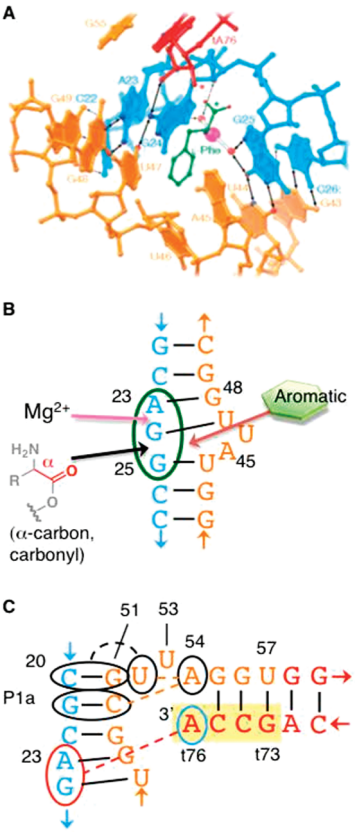

FIGURE 7. Amino acid and tRNA recognition elements of Fx3. Colors of bases are the same as those in Figure 6. Base-pairing interactions are indicated by solid lines. Bases and/or base pairs interacting with each other by noncanonical interactions via hydrogen bonding or stacking are circled by solid lines and indicated by dashed lines. (A) The 3D structure of flexizyme's active site. Adapted with permission from ref 22. Copyright 2008 Nature. (B) Amino acid recognition elements. Binding sites of the aromatic ring, α-carbon, and carbonyl group of Phe-EE are indicated by arrows. (C) tRNA recognition elements.

bound to Phe-EE shows that the aforementioned irregular helix composed of J1a/2 (A23–G25) and J2/1a (U44–G48) provides the site for the benzyl side chain (Figure 7A, B). In more details, three non-Watson–Crick base pairs (A23•G48, G24•U47, and G25•U44) and intervening two bases (A45 and U46) constitute the irregular helix, where A23•G48 expands the minor groove, making the neighboring G24•U47 pair more stretched than the ordinary G•U. This stretched G24•U47 and G25•U44 pairs together with the adjoining two unpaired bases, A45 and U46, forming the amino acid binding pocket. A hydrated magnesium binds to the major groove of this irregular helix and stabilizes the binding pocket. The phenyl group of Phe-EE is accommodated in the pocket formed by G24•U47 and U46 and interacting with O6 of G24 by its partial positive charge at the center of the phenyl ring. The α-carbon and carbonyl group of Phe-EE are accommodated in the cleft between G24 and G25. On the other hand, no electron density was observed for the ethyl group of ester moiety, but there is a space opening to the outer space and directing away from the flexizyme. This observation agrees with the fact that Fx3 accepts a relatively bulky leaving group, such as AMP, suggesting that the leaving group of the activated amino acid does not interact with the active pocket. Moreover, the α-amino group points away from the active pocket of flexizyme and thus is not precisely recognized by the flexizyme, either.

From these observations, we can speculate the mechanism how other advanced flexizymes such as dFx accept acyl-donor substrates irrespective of side chain or α-position

heteroatom. dFx very likely recognizes the 3,5-dinitrobenzyl leaving group of acyl-DBE; that is, the substrate presumably binds to the active site of dFx in inverted orientation respect to Phe-EE (Phe-CME) to Fx3. Therefore, it is reasonably assumed that side chain of acyl-DBE is directed away from the active site as similar to the orientation of the ethyl group of ester moiety seen in the crystal structure of Fx3. This results in granting dFx the ability to catalyze aminoacylation toward acyl-DBE with various side chains. The reason why Fx3 does not accept acyl-DBE is that the 3,5-dinitrobenzyl group is too large to be accommodated in the pocket of Fx3. In fact, dFx acquires a single base insertion in J2/1a. A reasonable assumption is that this base insertion along with base mutations enlarges the pocket to accommodate the 3,5-dinitrobenzyl group, though we currently have no clear vision how this alteration of bases could constitute the active cavity.

Structural Basis of Recognition of tRNA Acceptor End

All flexizymes accept virtually any kind of tRNA irrespective of their anticodon and body sequences, since tRNA binding requires the formation of only three base pairs between flexizyme's G55−U57 and tRNA's tC75−tG73 (Figure 7C). These three base pairs, consisting of two Watson−Crick and one G•U wobble pairs, initially predicted by biochemical experiments are also observed in the crystal structure. All tRNAs have the CCA-3′ end without exception, but the base at position 73, namely, discriminator base, potentially has all possible four bases. It has been shown that U57 in Fx3 tolerates pairing with tA73 (Watson−Crick) as well as tU73 (U•U wobble).[9] tRNA with tC73 is the worst substrate for Fx3, but prolonging the incubation time with Fx3 still yields the desired aminoacyl-tRNAs. Alternatively, U57G mutation in Fx3 pairs with tC73 would increase the efficiency like that of the normal Fx3-tRNA pair (unpublished data).

In addition to these three base pairs, A54 was fully conserved in all flexizymes. Although it was unclear if a noncanonical base pair of A54 with tA76 could be formed, our earlier biochemical data showed that tA76 was critical to maintain the full ribozyme activity where mutation of tA76 to C, G, or U decreases its aminoacylation activity by 20%, 15%, and 10%, respectively; and also A54 mutation diminishes aminoacylation activity.[9] The crystal structure of Fx3 has revealed that tA76 makes van der Waals contact with the ribose of G24 and its N1 makes a hydrogen bond with 2′-OH of A23.

Moreover, flexizyme utilizes the J1a/3 hairpin-shaped turn to precisely position the acceptor end of tRNA. Notably, U52−A54 form the J1a/3 turn by the following specific interactions: (1) A54 makes multiple hydrogen bonds with the C50•G21 pair of P1a from the minor groove of the helix, making the so-called type I A-minor motif,[25] (2) U52 stacks on A54 and also packs against G51•C20 of P1a, and (3) U53 is extruded from the turn. This turn structure helps to place G55−U57 perpendicular to the main helical stack in Fx3. The importance of this unique hairpin U-turn structure in activity was supported by the biochemical data where U52 was completely conserved in all flexizymes. Further, comparison of crystal structures of amino acid substrate bound and unbound conformers has revealed that this J1a/3 fully docks with P1a only when amino acid substrate (in this case Phe-EE) is bound. Together with the result from early chemical probing data on r24 where U52−53 were protected from Pb^{2+} cleavage upon binding of amino acid substrate,[9] the precise positioning of the acceptor end of tRNA by J1a/3 seems to be closely coupled with amino acid substrate binding.

An intriguing mechanistic question arisen from our earlier biochemical data that Fx3 exclusively aminoacylates on the 3′-OH group of tA76, not on the 2′-OH group.[10] The crystal structure has revealed that this selectivity is directed by two major interactions. First, the 3′-OH is positioned close to the carbonyl carbon of Phe-EE through the interactions between tA76 and J1a/2, where tA76 interacts with G24 and A23 as described above. Second, in the short helix consisting of G55−G56 and tC75−tC74, the 2′-OH is buried in the minor groove and less accessible than 3′-OH.

Conclusion

We have here comprehensively summarized the evolutionary history of a family of ARS ribozymes, flexizymes. Despite the fact that our attempt toward the generation of such ribozymes began from evolutionary interests of the origin of aminoacylation and translation chemistries, we have ended up developing a useful and practical catalytic system with almost no structural limitation of usable amino/hydroxy acids and tRNA, thus providing a versatile tool for the synthesis of acyl-tRNAs. This system offers us a facile method for the preparation of various mischarged tRNAs that allows for performing a new dimension of biochemical studies on the translation system, such as ribosome[26] and elongation factors. Moreover, acyl-tRNAs prepared by flexizymes could be integrated with a custom-made in vitro translation system, enabling us to reprogram the genetic

code assigning from proteinogenic amino acids to nonproteinogenic amino acids.[12−20,27−33] Thus, this methodology opens new means to express nonstandard peptides and their libraries for the discovery of therapeutic peptides.

Regardless of the potential applications of flexizymes to translation, the present studies on flexizymes imply a possible existence of similar ribozymes in the RNA world. We have shown that in a small ribozyme consisting of only 45 nt, a few mutations and insertions in the active site have successfully altered the specificity toward amino acid substrates with different aromatic leaving groups. This may imply that ribozymes set in the 5′-leader region of tRNAs could have evolved against amino acids with particular signatures of side chains or leaving groups (or their combinations). Because it has been shown that M1 RNA is able to cleave 5′-leader flexizyme away from tRNA after aminoacylation, in the RNA world such collaborations of catalytic precursor tRNAs with M1 RNA-like ribozyme could have established the genetic code currently used in all organisms. Yet, this is a still scientific imagination until we witness such a molecular fossil discovered in nature.

BIOGRAPHICAL INFORMATION

Jumpei Morimoto was born in Takasago City, Japan in 1983. He received his B. Eng. and M. Eng. from the University of Tokyo and entered the doctoral course in the Department of Chemistry and Biotechnology, Graduate School of Engineering in 2009 under the supervision of H. Suga. His current research interests are in the field of biotechnology related to translation and discovery of peptidomimetics against epigenetic targets.

Yuuki Hayashi was born in Tsukushino City, Japan in 1978. After receiving his B. Eng. and M. Eng. from Osaka University, he received his D. Eng. from the Department of Information Science, Graduate School of Engineering at the same University in 2006 under the supervision of Prof. T. Yomo. He has been working as a post-doctoral fellow under the supervision of H. Suga since 2006. His present research interest is to identify novel selective peptide inhibitors against kinases, which are responsible for human diseases such as pancreatic and ovarian carcinomas.

Kazuhiro Iwasaki was born in Osaka City, Japan in 1983. He received his B. Sci. from Tokyo University of Science and M. Eng. from the University of Tokyo. He is in the doctoral course of the Department of Chemistry and Biotechnology, Graduate School of Engineering under the supervision of H. Suga. His present research interest is in vitro selection of nonstandard peptides against cancer markers.

Hiroaki Suga was born in Okayama City, Japan in 1963. He received his B. Eng. and M. Eng. from Okayama University and Ph. D. in Chemistry from the Massachusetts Institute of Technology in 1994 under the supervision of S. Masamune. After three years of post-doctoral work in Massachusetts General Hospital working in the laboratory of J. W. Szostak, he was appointed as a tenure-track Assistant Professor in the Department of Chemistry in the State University of New York at Buffalo to carry his independent research group, and he became tenured Associate Professor in 2002. In 2003, he moved to the Research Center for Advanced Science and Technology at the University of Tokyo as Associate Professor, and soon after was promoted to Full Professor. In 2010, he changed his affiliation to the Department of Chemistry, Graduate School of Science. His research interests are in the field of chemical biology and biotechnology related to RNA, translation, and peptides. He also conducts drug discovery programs in collaboration with industrial partners, including his own biotechnology company, PeptiDream Inc. Tokyo.

We thank Drs. H. Saito and H. Murakami for early contributions to the development of flexizymes, Drs. H. Xiao and A. R. Ferré-D'Amaré for contributions to solving the crystal structure, and all members of the Suga laboratory. This work was supported by grants from the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (S) (16101007), Specially Promoted Research (21000005), and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization to H.S.; and, in part, Grants-in-Aid for JSPS fellows to J.M. (21-9079) and K.I. (22-9695).

FOOTNOTES

*To whom correspondence should be addressed. Telephone/Fax: +81-3-5841-8372. E-mail: hsuga@chem.s.u-tokyo.ac.jp.

REFERENCES

1 Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **1995**, *267*, 643–647.

2 Illangasekare, M.; Yarus, M. Specific, rapid synthesis of Phe-RNA by RNA. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5470–5475.

3 Jenne, A.; Famulok, M. A novel ribozyme with ester transferase activity. *Chem. Biol.* **1998**, *5*, 23–34.

4 Lohse, P. A.; Szostak, J. W. Ribozyme-catalysed amino-acid transfer reactions. *Nature* **1996**, *381*, 442–444.

5 Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H. Ribozyme-catalyzed tRNA aminoacylation. *Nat. Struct. Biol.* **2000**, *7*, 28–33.

6 Lee, N.; Suga, H. A minihelix-loop RNA acts as a trans-aminoacylation catalyst. *RNA* **2001**, *7*, 1043–1051.

7 Bessho, Y.; Hodgson, D. R.W.; Suga, H. A tRNA aminoacylation system for non-natural amino acids based on a programmable ribozyme. *Nat. Biotechnol.* **2002**, *20*, 723–728.

8 Saito, H.; Kourouklis, D.; Suga, H. An in vitro evolved precursor tRNA with aminoacylation activity. *EMBO J.* **2001**, *20*, 1797–1806.

9 Saito, H.; Watanabe, K.; Suga, H. Concurrent molecular recognition of the amino acid and tRNA by a ribozyme. *RNA* **2001**, *7*, 1867–1878.

10 Saito, H.; Suga, H. A ribozyme exclusively aminoacylates the 3′-hydroxy group of the tRNA terminal adenosine. *J. Am. Chem. Soc.* **2001**, *123*, 7178–7179.

11 Murakami, H.; Saito, H.; Suga, H. A versatile tRNA aminoacylation catalyst based on RNA. *Chem. Biol.* **2003**, *10*, 655–662.

12 Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* **2006**, *3*, 357–359.

13 Goto, Y.; Murakami, H.; Suga, H. Initiating translation with D-amino acids. *RNA* **2008**, *14*, 1390–1398.

14 Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **2008**, *3*, 120–129.

15 Kawakami, T.; Murakami, H.; Suga, H. Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides. *Chem. Biol.* **2008**, *15*, 32–42.

16 Kawakami, T.; Murakami, H.; Suga, H. Ribosomal synthesis of polypeptoids and peptoid–peptide hybrids. *J. Am. Chem. Soc.* **2008**, *130*, 16861–16863.

17 Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. Synthesis of polyester by means of genetic code reprogramming. *Chem. Biol.* **2007**, *14*, 1315–1322.

18 Niwa, N.; Yamagishi, Y.; Murakami, H.; Suga, H. A flexizyme that selectively charges amino acids activated by a water-friendly leaving group. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3892–3894.

19 Goto, Y.; Suga, H. Translation initiation with initiator tRNA charged with exotic peptides. *J. Am. Chem. Soc.* **2009**, *131*, 5040–5041.

20 Ohshiro, Y.; Nakajima, E.; Goto, Y.; Fuse, S.; Takahashi, T.; Doi, T.; Suga, H. Ribosomal synthesis of backbone-macrocyclic peptides containing γ-amino acids. *ChemBioChem* **2011**, *12*, 1183–1187.

21 Goto, Y.; Katoh, T.; Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **2011**, *6*, 779–790.

22 Xiao, H.; Murakami, H.; Suga, H.; Ferré-D'Amaré, A. R. Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature* **2008**, *454*, 358–361.

23 Ferré-D'Amaré, A. R. Use of the spliceosomal protein U1A to facilitate crystallization and structure determination of complex RNAs. *Methods* **2010**, *52*, 159–167.

24 Saito, H.; Suga, H. Outersphere and innersphere coordinated metal ions in an aminoacyl-tRNA synthetase ribozyme. *Nucleic Acids Res.* **2002**, *30*, 5151–5159.

25 Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4899–4903.

26 Effraim, P. R.; Wang, J.; Englander, M. T.; Avins, J.; Leyh, T. S.; Gonzalez, R. L.; Cornish, V. W. Natural amino acids do not require their native tRNAs for efficient selection by the ribosome. *Nat. Chem. Biol.* **2009**, *5*, 947–953.

27 Nakajima, E.; Goto, Y.; Sako, Y.; Murakami, H.; Suga, H. Ribosomal synthesis of peptides with C-terminal lactams, thiolactones, and alkylamides. *ChemBioChem* **2009**, *10*, 1186–1192.

28 Sako, Y.; Goto, Y.; Murakami, H.; Suga, H. Ribosomal synthesis of peptidase-resistant peptides closed by a nonreducible inter-side-chain bond. *ACS Chem. Biol.* **2008**, *3*, 241–249.

29 Sako, Y.; Morimoto, J.; Murakami, H.; Suga, H. Ribosomal synthesis of bicyclic peptides via two orthogonal inter-side-chain reactions. *J. Am. Chem. Soc.* **2008**, *130*, 7232–7234.

30 Yamagishi, Y.; Ashigai, H.; Goto, Y.; Murakami, H.; Suga, H. Ribosomal synthesis of cyclic peptides with a fluorogenic oxidative coupling reaction. *ChemBioChem* **2009**, *10*, 1469–1472.

31 Kang, T. J.; Yuzawa, S.; Suga, H. Expression of histone H3 tails with combinatorial lysine modifications under the reprogrammed genetic code for the investigation on epigenetic markers. *Chem. Biol.* **2008**, *15*, 1166–1174.

32 Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H. Diverse backbone-cyclized peptides via codon reprogramming. *Nat. Chem. Biol.* **2009**, *5*, 888–890.

33 Goto, Y.; Iwasaki, K.; Torikai, K.; Murakami, H.; Suga, H. Ribosomall synthesis of dehydrobutyrine- and methyllanthionine-containg peptides. *Chem. Commun.* **2009**, *23*, 3419–3421.