ChemComm

Initiating ribosomal peptide synthesis with exotic building blocks

Journal:	*ChemComm*
Manuscript ID	CC-HIG-02-2020-001291.R2
Article Type:	Highlight

Initiating ribosomal peptide synthesis with exotic building blocks

Dr. Christos Tsiamantas[+], Dr. Joseph M. Rogers[+], Prof. Hiroaki Suga*

[+] These authors are equally contributed

Department of Chemistry, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, 113-0033 (Japan)

Abstract

Ribosomal peptide synthesis begins almost exclusively with the amino acid methionine, across all domains of life. The ubiquity of methionine initiation raises the question; to what extent could polypeptide synthesis be realized with other amino acids, proteinogenic or otherwise? This highlight describes the breadth of building blocks now known to be accepted by the ribosome initiation machinery, from subtle methionine analogues to large exotic non-proteinogenic structures. We outline the key methodological developments that have enabled these discoveries, including the exploitation of methionyl-tRNA synthetase promiscuity, synthetase and tRNA engineering, and the utilization of artificial tRNA-loading ribozymes, flexizymes. Using these methods, the number and diversity of validated initiation building blocks is rapidly expanding permitting the use of the ribosome to synthesize ever more artificial polymers in search of new functional molecules.

Introduction

Regardless of organism, methionine or one of its derivatives is the first amino acid used by the ribosome in the assembly of peptide chains. However, there are good reasons to initiate peptide synthesis with an alternative, non-proteinogenic amino acid. Non-proteinogenic amino acids can offer new chemical or photo-reactivity, spectroscopically useful moieties, radio- or isotopic-labels, convenient affinity handles, or desirable structural or physicochemical features.[1] Using these amino acids during initiation offers certain advantages over their incorporation during peptide elongation. Firstly, as termini of proteins tend to be solvent exposed, the N-terminal position is likely to tolerate substitutions without significantly affecting protein structure or function. Secondly, it ensures a single instance of an unnatural building block in the peptide chain, even when using randomized codon libraries. Lastly, as the amino acid does not require an amine, or amine-like nucleophile, to be incorporated in the growing peptide chain, a greater variety of substrates can be envisioned, at least in principle. It is interesting therefore to note that the early work on "site-specific incorporation" of non-proteinogenic or exotic amino acids focused on elongation process,[2-4] whereas replacing Met during translation initiation appeared later. Perhaps this reflects the perceived stricter governance of the initiation process, as hinted at by the universality of Met as the key initiator amino acid. However, in the last few years there has been much progress in engineering ribosomal initiation to use non-Met substrates. This highlight describes this progress, including the exploitation of MetRS promiscuity, tRNA and aminoacyl-tRNA synthetase (aaRS) engineering, and the development and utilization of artificial tRNA-loading ribozymes, flexizymes, that most aggressively expand the scope of translation initiation with exotic building blocks in peptide synthesis. We describe how these advances have redefined the limitations of, and expanded our understanding of, nature's translational machinery.

Mechanism of initiation

E. coli is the most used organism for ribosome and translation engineering. As with other prokaryotes, polypeptide synthesis starts exclusively with N^{α}-formyl-methionine (fMet).[5] Methionyl-tRNA synthetase (MetRS) aminoacylates a tRNA dedicated to translation initiation ($tRNA^{fMet}_{CAU}$), yielding Met-$tRNA^{fMet}_{CAU}$ (Fig. 1).[6] The same MetRS also aminoacylates a separate tRNA to supply Met during elongation. However, the presence of recognition elements on the body of the initiator tRNA render the Met-$tRNA^{fMet}_{CAU}$ complex a substrate for methionyl-$tRNA^{fMet}$ transformylase (MTF), which catalyzes a formylation reaction, yielding fMet-$tRNA^{fMet}_{CAU}$ (Fig. 1). fMet-$tRNA^{fMet}_{CAU}$ is then bound to the 30S subunit by IF2 (initiation factor 2), and together with the IF1 and IF3, is ready to accept the 50S subunit and begin translation. Following the aminoacylation step, the amino acid attached to an initiator tRNA is inspected at least twice prior to its use in initiation of protein synthesis, by MTF followed by IF2. It is this sequential recognition that guarantees the universality of Met as the initiating amino acid during ribosomal polypeptide synthesis. Another important feature is that fMet-$tRNA^{fMet}_{CAU}$ is bound directly to the P-site, in contrast to all aminoacylated tRNA used during elongation, which must enter first through the A-site (Fig. 1).

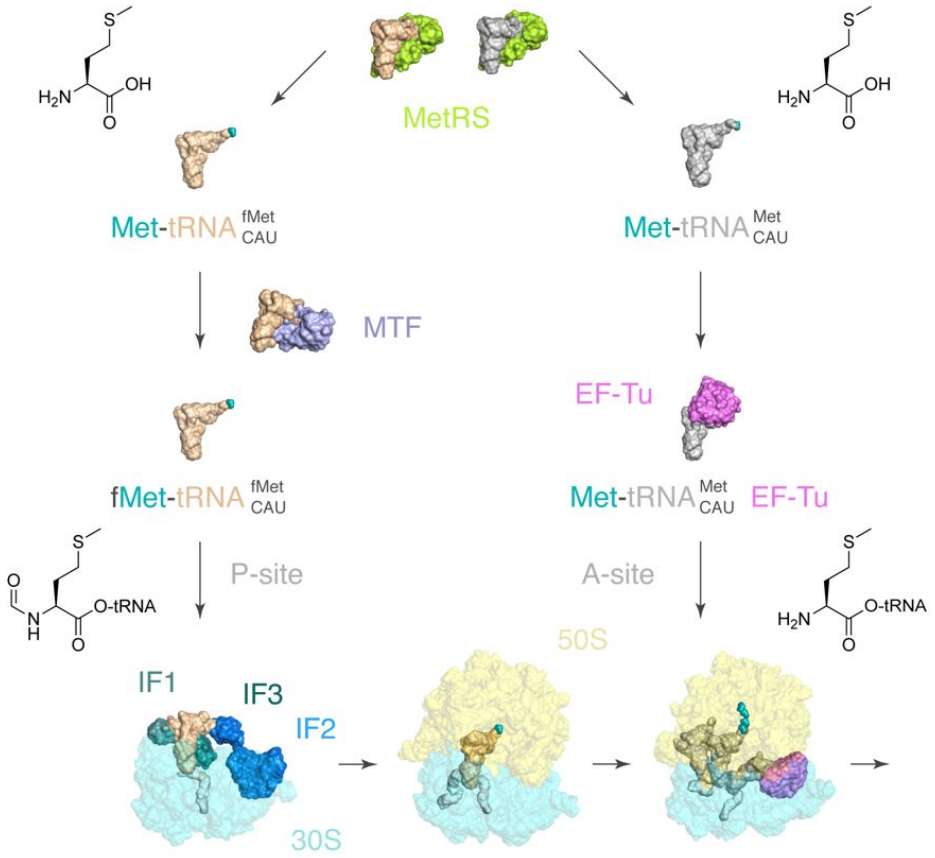
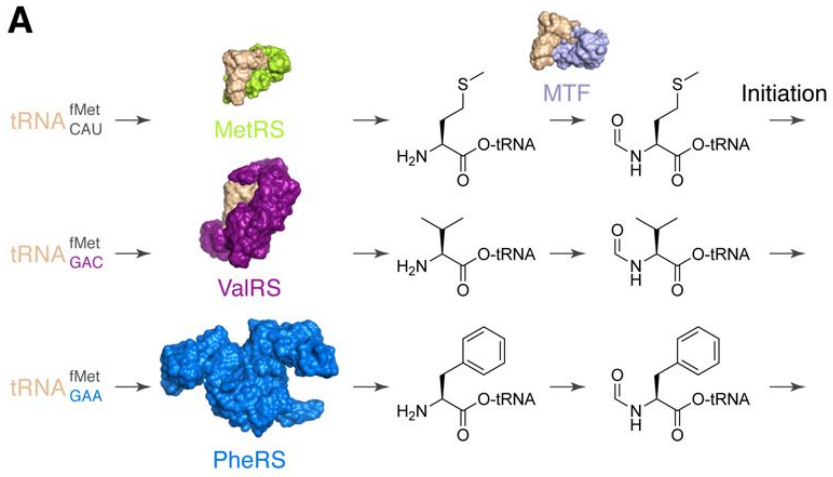


Figure 1. Mechanism of peptide synthesis initiation at the ribosome. Methionine is loaded by MetRS onto both initiation $tRNA^{fMet}_{CAU}$ and elongation $tRNA^{Met}_{CAU}$. Only Met-$tRNA^{fMet}_{CAU}$ is formylated by MTF, brought to the 30S ribosome subunit together with initiation factors, and engaged by the 50S subunit to start

peptide elongation. During the elongation cycle, EF-Tu delivers of loaded tRNA, like Met-tRNA$^{Met}_{CAU}$, to the ribosome.

Initiation with non-Met proteinogenic amino acids

It is possible to direct non-Met proteinogenic amino acids to start translation *in vivo*. Chattapadhyay *et al.*, demonstrated that altering the identity of the anticodon sequence of tRNA^{fMet} can alter the amino acid that will initiate the translation event in *E. coli*.[7] Briefly, alteration of the anticodon to GAC or GAA is sufficient to drive misaminoacylation of the initiator tRNA with Val or Phe, by ValRS and PheRS, respectively. tRNA^{fMet}, bearing these amino acids, initiated protein synthesis *in vivo* when provided with initiation codons complementary to the modified anticodons (Fig. 2a). This process resulted in fairly low incorporation efficiency which was mainly attributed to three factors: 1) the less efficient aminoacylation of the tRNA mutants by valyl- and phenylalanylRS, and the preference of 2) MTF and 3) IF2 for Met, as opposed to other amino acids.[8, 9] In a subsequent investigation by Mayer *et al.*, the effect of overproduction of aaRS, MTF, and IF2 on the translation efficiency of reporter proteins was investigated, leading to the identification of the limiting factors for initiating translation with 4 different amino acids, namely: Val, Phe, Ile and Gln.[10] Their results, which are summarized in Figure 2b, highlight the absence of apparent restrictions for initiating translation with other amino acids or codons, providing sufficient aminoacylation, formylation and delivery to the ribosome occurs.



B

Anticodon	Amino acid	Initiation yield	Yield increased most by overexpression of
CAU	Met	100 (set)	-
CUA	Gln	80	IF2
GAC	Val	160	MTF
GAU	Ile	120	MTF
GAA	Phe	20	aaRS
CUG	Gln	40	IF2

Figure 2. Non-Met amino acids can initiation translation *in vivo*. (A) Changing the initiator tRNA^{fMet} anticodon is sufficient to drive misacylation of tRNA by other aaRS, resulting in some degree of formylation and translation initiation. (B) Depending on the amino acid, potential bottlenecks

(aminoacylation, formylation or initiation complex formation) are relieved by overexpression of different associated proteins (aaRS, MTF, IF2).

Initiation with non-proteinogenic amino acids *in vivo*

More than half a century ago Trupin *et al.* demonstrated that a non-proteinogenic amino acid could replace Met in initiation.[11] They showed that norleucine, an analogue of both Leu and Met, is a good substrate both for MetRS and MTF, leading to the formation of fNorleucine-tRNA$^{fMet}_{CAU}$. Subsequent studies demonstrated that fNorleucine-tRNA$^{fMet}_{CAU}$ could replace fMet-tRNA$^{fMet}_{CAU}$ in the formation of the initiation complex and initiate protein expression *in vivo* (Fig. 3a).[12] Separately, the groups of Hoffman and Habtman demonstrated that *E. coli* MetRS was unable to discriminate Met from selenomethionine and ethionine, respectively, resulting in polypeptide synthesis using these amino acids, provided Met auxotrophs were used.[13-15] The use of the former is particularly valuable in heavy atom replacement for X-ray crystallography. These pioneering investigations illustrate some flexibility in the side-chain recognition by MetRS and MTF. In the following years, a cascade of investigations was triggered, in which the incorporation of an increasing number of Met analogs was realized, deviating more and more from the native substrate.[16, 17] Examples included 2-aminohexanoic acid,[16] homoallylglycine, homopropargylglycine[18] and azidohomoalanine (Fig. 3b).[19] It is noteworthy that overproduction of MetRS, in *E. coli* Met-depleted cultures, enabled the utilization of modest substrates for MetRS, including *cis/trans*-crotylglycine, 2-aminoheptanoic acid, norvaline, 2-butynylglycine and allylglycine (Fig. 3b).[20] Exploitation of the promiscuity of MetRS had thus become an invaluable tool for the incorporation of non-proteinogenic amino acids into proteins, an approach that is still applicable up to today, both *in vivo* and *in vitro*. However, the significant drawback of this methodology is that the substrate must be compatible with MetRS, essentially limiting the available registry to amino acids that are structurally and electronically similar to Met.

There is a large effort to engineer aaRSs to accept non-cognate amino acids.[21-24] Only limited protein engineering studies have been carried out on MetRS, and thus translation initiation. However, these have shown that a single mutation, L13G, in the Met binding pocket of the MetRS, enables the aminoacylation reaction to be carried out with azidonorleucine (Anl) instead of Methionine (Figure 3c).[25] Subsequent studies showed that protein expression *in vivo* can be carried out using Anl instead of Met with the Anl incorporation going as high as 90% of the total Met content in Met auxotrophs.[26, 27] Despite this methodology being among the most efficient for the *in vivo* N-terminal incorporation of non-proteinogenic amino acids into proteins, its scope is still limited to amino acids with structural or electronic similarities to Met. However, Tharp *et al.* provided an alternative approach.[28] They transplanted identity elements from *Methanocaldococcus jannaschii* (Mj) tRNA^{Tyr} to *E. coli* tRNA^{fMet} yielding a novel chimera orthogonal to *E. coli* MetRS, a good substrate for MjTyrRS and accepted by the initiation machinery. This allowed for initiation by a range of aromatic non-proteinogenic amino acids (Figure 4c), with the removal of redundant copies of tRNA^{fMet} to overcome the low efficiencies initially observed. This work demonstrated for the first time that deviating from Met derivatives during the initiation of translation is feasible *in vivo*.

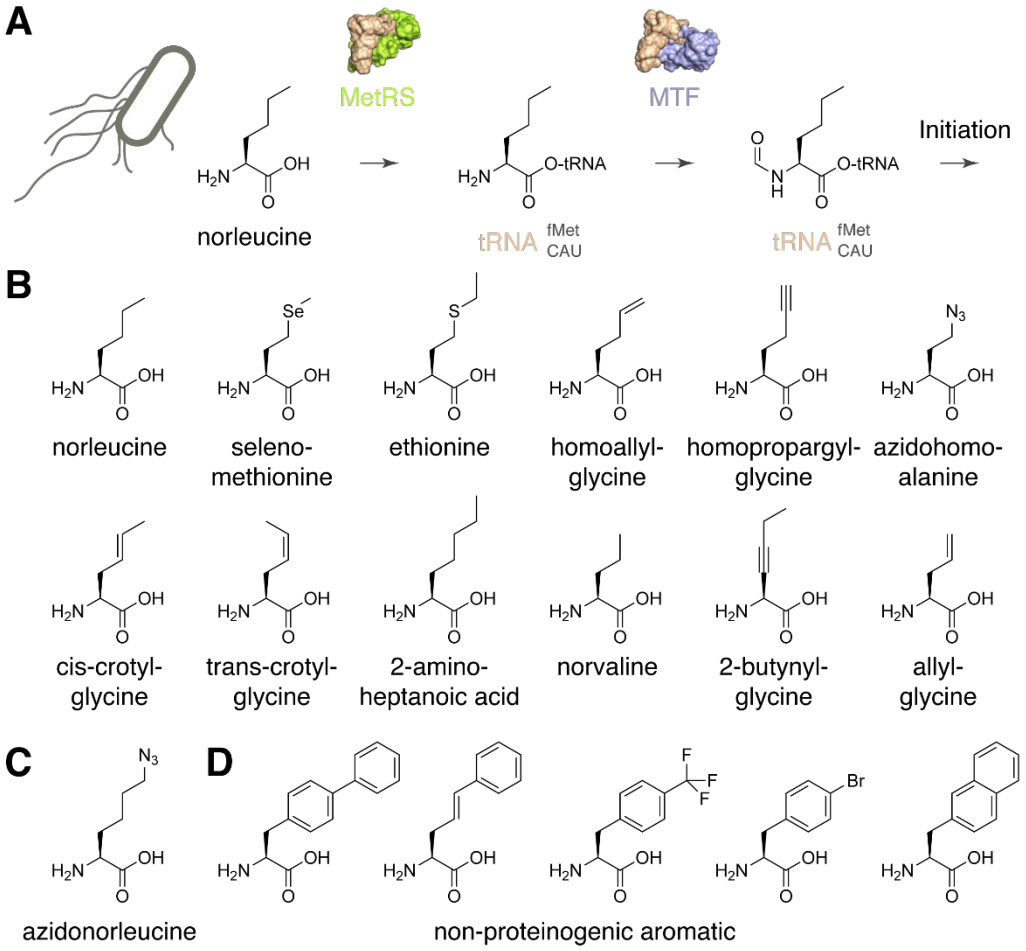


Figure 3. Non-proteinogenic Met substitution *in vivo*. ((A) In *E. coli*, Met can be substituted for norleucine and undergo the same initiation process. (B) Other Met analogues accepted by the same mechanism. (C) Mutant MetRS can accept azidonorleucine. (D) tRNA engineering allows aromatic non-proteinogenic amino acids to be used as the initiator amino acid.

Reprogramming of *in vitro* translation initiation

In order to include a broader range of non-proteinogenic amino acids, scientists have attempted to circumvent the natural aaRS step by employing a plethora of alternative approaches to forming amino acylated tRNA, including chemoenzymatic tRNA acylation.[29] Chemoenzymatic tRNA acylation involves the synthesis of a 5'-phospho-2'-deoxyribocytidylylriboadenosine (pdCpA) dinucleotide, which is then acylated with the desired amino acid. T4-mediated ligation of the resulting 2'(3')-O-acylated pdCpA derivative to a tRNA missing cytidine-75 and adenosine-76, affords a chemoenzymatically acylated tRNA (Fig. 4a). Characteristic examples of chemoenzymatically acylated $tRNA^{fMet}$ can be found in the works of McIntosh *et al.* and Gite *et al.* where chemical aminoacylation of initiator tRNA with fluorophore-amino acid conjugates and subsequent supplementation into cell-free *E. coli* translation systems, yielded N-terminally, fluorescently labeled proteins.[30, 31] Initial attempts were carried out using AUG initiation codon, which inevitably resulted in low incorporation efficiency due to the competition with Met available in the *in vitro* translation system. Subsequently, utilization of the amber stop codon led to a slight increase in the

translation efficiency, which remained as low as 2%.[32] The very low yield in these cases is attributed to aminoacylation of the deacylated $tRNA^{fMet}_{CUA}$ by cognate endogenous aaRSs, resulting in contamination with polypeptides initiated with other amino acids. Additionally, only N^{α}-fluorophore labeled Met, Val and Lys were used in the context of these studies. It is noteworthy though that attachment of the bulky residue at the amine did not cause a significant drop in the affinity for IF2.

Non-Met initiation in the PURE system

The assembly of a cell-free translation system in the beginning of the millennium by Ueda and coworkers,[33] consisting of purified translation factors, namely the "protein synthesis using recombinant elements" system (PURE), led to the realization that nature's translational apparatus can be utilized as a catalyst for template encoded polymer synthesis.[34] The PURE system allows for the withdrawal of amino acids and/or aaRSs at will, creating vacant codons. In this way, and as far as translation initiation is concerned, depletion of Met from the system renders the initiation codon vacant, allowing an alternative amino acid to be reassigned to initiation.

In the pioneering work of Josephson *et al.*, several Met analogues were placed within the context Met-deficient PURE to, using the natural promiscuity of MetRS (Fig. 4a), generate peptides equipped with non-proteinogenic amino acid initiators. In particular, incorporation of N-formyl-2-aminohex-5-ynoic acid allowed the N-terminal posttranslational derivatization of peptides with various azides.[35] Hacker *et al.*, used 2-amino-4-azidobutanoic acid in a Met depleted system, which upon combination with a *p*-ethynyl phenylalanine at elongation position enabled the generation of vast libraries of peptide macrocycles by using copper-catalyzed azide-alkyne cycloaddition.[36]

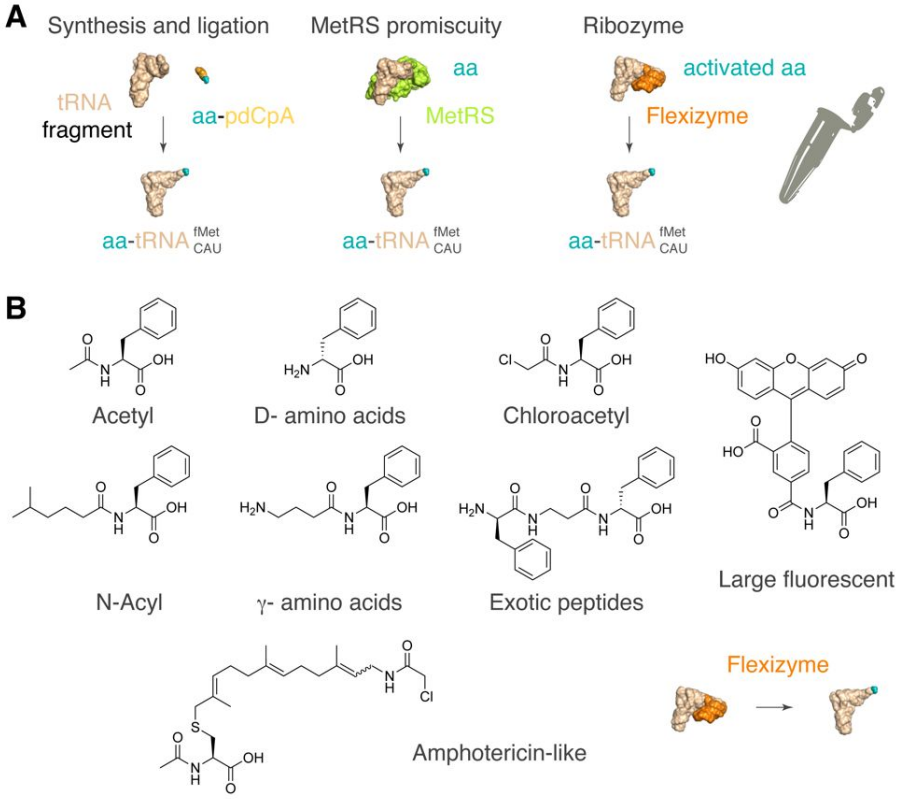


Figure 4. Reprogramming of *in vitro* translation initiation is particularly flexible. (A) Alternative methods of generating loaded $tRNA^{fMet}_{CAU}$. (B) Example non-proteinogenic substrates used in initiation, enabled by flexizyme (for comprehensive list, see Rogers OBC 2015[37]).

Flexizyme reprogramming of initiation

The most notable development in the field of initiation engineering is the discovery and utilization of flexizymes - artificial ribozymes capable of aminoacylating tRNA. However, unlike aaRS which are specific for their cognate amino acids, flexizymes are general catalysts which, when provided suitable activated amino acids, can aminoacylate tRNA of any sequence or anti-codon (Fig. 4a). Loaded tRNA can then be added to amino acid or aaRS deficient PURE systems, allowing for extensive reprogramming of the genetic code. This system was referred to as FIT (Flexible *In vitro* Translation) system, where flexizymes are utilized as the key genetic code reprogramming tool. The use of FIT system to reprogram elongation has been extensively reviewed elsewhere,[37] here we focus on the reprogramming of translation initiation.

Goto *et al.* in a pioneering study in 2008, used flexizymes to load proteinogenic amino acids onto $tRNA^{fMet}_{CAU}$ and showed that the majority could initiate translation.[38] Interestingly, they discovered that MTF formylated most of these amino acids, with the exception of proline that gave solely the non-formylated product. Based on the structure of the complex between fMet-$tRNA^{fMet}_{CAU}$ and MTF, in which the side chain of Met is accommodated into a binding pocket composed of hydrophobic amino acids, they explained the inefficient incorporation of charged amino acids (Lys, Asp, Arg, and Glu).[39] In the same study the incorporation of N-acylated amino acids was attempted (Fig. 4b). In particular, it was shown that two moderately incorporating amino acids, Lys and Arg, demonstrated a 2- and 4-fold increase in the initiation efficiency with pre-acylation, presumably because acylated amino acids better mimic the formylated natural fMet substrate during the process of initiation.

In a follow-up study, D-amino acids were loaded onto $tRNA^{fMet}_{CAU}$ using flexizymes and were found to yield D-amino acid-initiated peptides (Fig. 4b). Interestingly, the initiating residue in the vast majority of the cases corresponded to the non-formylated amino acid, suggesting that MTF poorly recognizes D-amino acids. However, upon pre-acylation of the amine they observed a significant increase in the formation of the desired product, circumvented recognition by MTF. In a similar manner, N-chloroacetyl amino acids were also shown to be efficient substrates for translation initiation. When incorporated into peptides, this N-terminal moiety was found to have reactivity poised for efficient intramolecular reaction with the nearest cysteine in the peptide chain, permitting the formation of non-reducible peptide macrocycles.

Having validated translation initiation with single L- and D-amino acids, chemically synthesized di- to penta-peptides, composed of different combinations of N-Methyl-, D-, β- and even δ-amino acids, were tested for their ability to initiate the translation event (Fig. 4b).[40] Having always their C-terminus appropriately activated, the oligopeptides were shown to be good substrates for Flexizymes, yielding the corresponding oligopeptide-$tRNA^{fMet}_{CAU}$ conjugates. All of them successfully initiated translation affording peptides bearing exotic peptides at their N-terminus. Interestingly, the ribosome initiation complex could still be assembled around these exotic peptides, despite their large size relative to proteinogenic Met. Presumably, this is because loaded $tRNA^{fMet}_{CAU}$ is delivered to the P-site, where the peptide can occupy the exit-tunnel within the assembled ribosome. Along the same lines, Takatsuji *et al.* incorporated a dipeptide consisting of (*R*)-thiazolidine-4-carboxylic acid (Thz) and (*S*)-2-amino-4-(2-chloroacetamido)butanoic acid

(Cab) which, combined with further reprogramming at elongation positions, led to the generation of libraries composed of backbone-cyclized peptides for the first time.[41]

The possibility to incorporate peptidic segments at the initiation position triggered another study in which the activated amino acid was coupled to an N-free γ-amino acid (Fig. 4b),[42] with the latter previously shown to be incompatible with aminoacylation due to intramolecular cleavage. Subsequently, head-to-tail macrocyclization yielded, for the first time, cyclic peptides having a γ-amino acid on their backbone, expanding even further the available amino acid registry.

The above mentioned studies paved the way for the incorporation of even larger and more complex substrates as demonstrated by the ribosomal synthesis of an amphotericin-B inspired macrocycle (Fig. 4b).[43] Amphotericin-B, a membrane-interacting antifungal natural product, consists of an amphiphilic polyene−polyol macrolide. Torikai *et al.* introduced a hydrophobic terpene at the N-terminus of a model peptide. An activated ^{D}Cys derivative of *N*-acetyl-S-12-(ClAc) farnesyl group was loaded on $tRNA^{fMet}_{CAU}$ which was then supplemented to a methionine deficient *in vitro* translation system, successfully yielding the desired product. Importantly, the presence of a downstream Cys enabled the generation of a macrocyclic analogue.

Further non-proteinogenic amino acids for the formation of macrocyclic peptides were tested by Kawakami *et al*.[44] First, they showed that a range of tertiary amine amino acids, N-alkyl plus N-chloroacetyl, are accepted by the ribosome during initiation (Fig. 5a). Then, they showed that certain chloroaryl substrates are also tolerated (Fig. 5b), with such moieties also being cysteine reactive to form various macrocyclic and bicyclic translated peptides. Most interestingly for the present discussion is that these substrates are entirely lacking amine groups, and yet are still recognized by the initiation factors and ribosome subunits during start-up of peptide synthesis. Similarly, a wide range of aryl substrates were tested for ribosome tolerance by Ad *et al.*[45] and Lee *et al.*[46] and many were successfully incorporated into peptides (Fig. 5c). The known tolerance of the initiation process was further expanded with the successful inclusion of additional non-amino acid substrates; polyketide-like moieties,[45] conjugated unsaturated,[46] and diverse alkyl chains (Fig. 5d).[46]

A highly exotic carborane-containing building block was recently shown to be accepted by the initiation process by Yin *et al.*[47] The side-chain of this amino acid contains a cluster of ten boron atoms and is structurally and electronically dissimilar to any of the proteogenic and non-proteinogenic amino acids described above (Fig. 5e). Moreover, when part of a specific binding cyclic peptide this appendage has the potential to be used for focused boron neutron capture therapy.

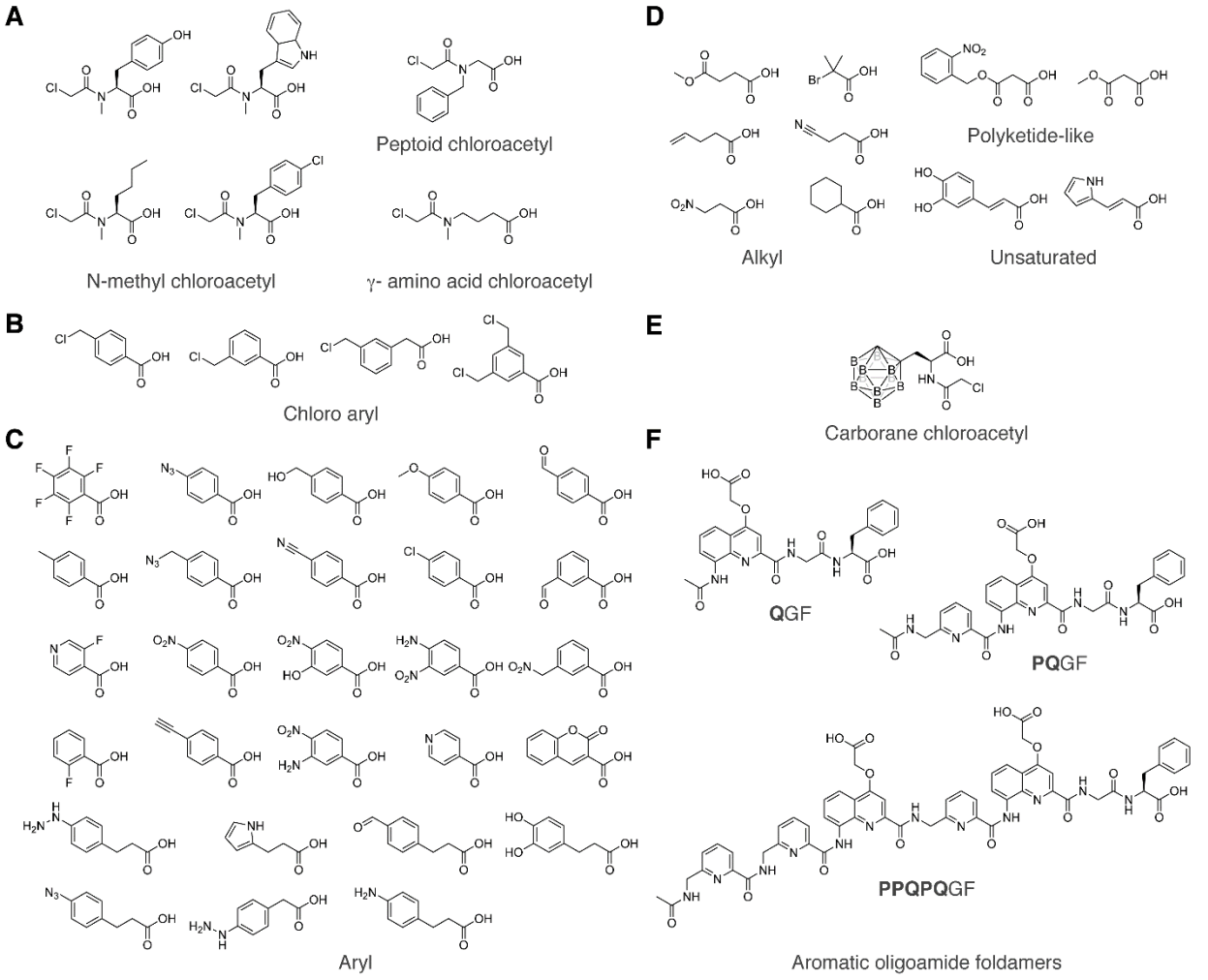


Figure 5. Highly non-proteinogenic substrates for peptide synthesis initiation at the ribosome. (A) N-chloroacetyl, (B) chloroaryl, (C) numerous aryl, (D) alkyl, polyketide-like and unsaturated, (E) carborane chloroacetyl and (F) foldamer initiators.

Lastly, the tolerance of the initiation process for large substrates has been explored using abiotic 'foldamer' building blocks. Foldamers are artificial folded molecular architectures inspired by the structures and functions of biopolymers[48], with building blocks ranging from β-, γ- and δ-amino acids to oligo-phenylene-ethynylenes and aromatic oligoamides. The aromatic oligoamides are highly abiotic and can form extremely stable helical conformations. Recently, such aromatic foldamers, attached to a Gly-^{L}Phe-CME spacer, were successfully loaded on a $tRNA^{fMet}_{CAU}$ using flexizyme and were incorporated during the initiation event into peptides.[49, 50] Foldamers of different lengths (composed of P & Q units, as illustrated in Fig. 5f) and side chains were tested, representing the largest entities accepted by the ribosome but also those most chemically remote from what the ribosome has been optimized to accept. Installation of a ClAc-group, combined with a downstream Cys, led to the generation of a novel class of hybrid macrocycles, composed of a rigid helical entity on one side and a peptide molecule rigidly projected in three-dimensional space. These studies demonstrated that the ribosome exit tunnel tolerates highly non-peptidic polymers, even those which, like proteins, have a propensity to fold.

Conclusions

Translation initiation is a fundamental step in gene expression, the significance of which can be nicely inlaid in Aristotle's proverb "well begun is half done". The conservation of methionine and its derivatives as the first amino acid has been extensively studied, questioned and challenged by the scientific community for over 60 years. This endeavor has led to significant expansion of our understanding, not only for the mechanisms governing the initiation event, but also for the tolerance of the various steps for non-Met building blocks. As described above, MetRS will accept certain non-proteinogenic amino acids, providing they are structurally and electronically related to Met. MTF will formylate a broader range of proteinogenic and non-proteinogenic amino acids loaded onto $tRNA^{fMet}_{CAU}$ but, for example, will not accept D-stereochemistry. IF2 and the ribosome initiation complexes appear least discriminating, accepting a very broad range of non-Met substrates of diverse size, stereochemistry, structure and functionality. The only general rule being that higher efficiencies are observed when N- of amino acid substrates are acylated. In all, provided loaded $tRNA^{fMet}_{CAU}$ can be made, and Met initiation suppressed, a huge range of structural and functional space can be accessed. Efficient and general methods for producing loaded $tRNA^{fMet}_{CAU}$, such as the flexizyme protocol, have and will help explore this space.

Applications stemming from the expanded registry of structural and functional elements include protein expression bearing heavy atoms for structural elucidation, inclusion of fluorescence tags, orthogonal handles for site-specific labeling, and intramolecular macrocyclization. The latter is particularly valuable in the formation of large cyclic peptide libraries and the discovery of novel highly-functional, drug-like peptides.[51] Exemplified by the foldamer studies, reprogramming initiation can lead to the generation of intriguing part-peptide hybrid molecules, the characteristics of which remain largely unexplored.

All in all, we have witnessed the development of N-terminally reprogrammed biopolymers resulting from single, site-specific incorporation of exotic building blocks into proteins and peptides, allowing the construction of macromolecules with unprecedented chemical and structural diversity. In the future, research will be directed towards the exploration and exploitation of both structural and functional characteristics of the resulting novel molecules. Furthermore, investigation of whether the recently expanded ribosomal tolerance can be extrapolated to translation elongation may take us even further towards harvesting the potential of the ribosome as an assembly line for ever more 'artificial' polymers.

Acknowledgements

We thank to all members of the Suga group who are involved in the engineering of translation initiation, and the work was supported by various research supports from JSPS, JST, and AMED grants as well as other private funding sources.

References

1. H. Neumann, *Synth. Biol.*, 2012, **586**, 2057-2064.
2. M. C. T. Hartman, K. Josephson and J. W. Szostak, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 4356-4361.
3. O. M. Rennert and H. S. Anker, *Biochemistry*, 1963, **2**, 471-476.
4. G. Baldini, B. Martoglio, A. Schachenmann, C. Zugliani and J. Brunner, *Biochemistry*, 1988, **27**, 7951-7959.
5. M. Kozak, *Microbiol. Rev.*, 1983, **47**, 1-45.

6. C. O. Gualerzi and C. L. Pon, *Biochemistry*, 1990, **29**, 5881-5889.
7. R. Chattapadhyay, H. Pelka and L. H. Schulman, *Biochemistry*, 1990, **29**, 4263-4268.
8. R. Giegé, J. P. Ebel and B. F. C. Clark, *FEBS Lett.*, 1973, **30**, 291-295.
9. R. M. Sundari, E. A. Stringer, L. H. Schulman and U. Maitra, *J. Biol. Chem.*, 1976, **251**, 3338-3345.
10. C. Mayer, C. Köhrer, E. Kenny, C. Prusko and U. L. RajBhandary, *Biochemistry*, 2003, **42**, 4787-4799.
11. J. Trupin, H. Dickerman, M. Nirenberg and H. Weissbach, *Biochem. Biophys. Res. Commun.*, 1966, **24**, 50-55.
12. S. S. Kerwar and H. Weissbach, *Arch. Biochem. Biophys.*, 1970, **141**, 525-532.
13. K. P. McConnell and J. L. Hoffman, *FEBS Lett.*, 1972, **24**, 60-62.
14. J. L. Hoffman, K. P. McConnell and D. R. Carpenter, *Biochim. Biophys. Acta*, 1970, **199**, 531-534.
15. T.-Y. Cheng and K. A. Habtman, *J. Mol. Biol.*, 1968, **31**, 191-207.
16. N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann and R. Huber, *Eur. J. Biochem.*, 1995, **230**, 788-796.
17. W. A. Hendrickson, J. R. Horton and D. M. LeMaster, *EMBO J.*, 1990, **9**, 1665-1672.
18. K. L. Kiick, J. C. M. van Hest and D. A. Tirrell, *Angew. Chem. Int. Ed.*, 2000, **39**, 2148-2152.
19. K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 19-24.
20. K. L. Kiick, R. Weberskirch and D. A. Tirrell, *FEBS Lett.*, 2001, **502**, 25-30.
21. D. R. Liu and P. G. Schultz, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 4780-4785.
22. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498.
23. K. Nozawa, P. O'Donoghue, S. Gundllapalli, Y. Araiso, R. Ishitani, T. Umehara, D. Söll and O. Nureki, *Nature*, 2009, **457**, 1163-1167.
24. J. C. W. Willis and J. W. Chin, *Nat. Chem.*, 2018, **10**, 831-837.
25. A. J. Link, M. K. S. Vink, N. J. Agard, J. A. Prescher, C. R. Bertozzi and D. A. Tirrell, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 10180-10185.
26. D. M. Abdeljabbar, T. J. Klein, S. Zhang and A. J. Link, *J. Am. Chem. Soc.*, 2009, **131**, 17078-17079.
27. D. M. Abdeljabbar, T. J. Klein and A. J. Link, *ChemBioChem*, 2011, **12**, 1699-1702.
28. J. M. Tharp, O. Ad, K. Amikura, F. R. Ward, E. M. Garcia, J. H. D. Cate, A. Schepartz and D. Söll, *Angew. Chem. Int. Ed.*, 2020, **59**, 3122-3126.
29. C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, 182.
30. B. McIntosh, V. Ramachandiran, G. Kramer and B. Hardesty, *Biochimie*, 2000, **82**, 167-174.
31. S. Gite, S. Mamaev, J. Olejnik and K. Rothschild, *Anal. Biochem.*, 2000, **279**, 218-225.
32. S. Mamaev, J. Olejnik, E. K. Olejnik and K. J. Rothschild, *Anal. Biochem.,* 2004, **326**, 25-32
33. Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa and T. Ueda, *Nat. Biotechnol.*, 2001, **19**, 751-755.
34. A. C. Forster, Z. Tan, M. N. L. Nalam, H. Lin, H. Qu, V. W. Cornish and S. C. Blacklow, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 6353-6357.
35. K. Josephson, M. C. T. Hartman and J. W. Szostak, *J. Am. Chem. Soc.*, 2005, **127**, 11727-11735.
36. D. E. Hacker, J. Hoinka, E. S. Iqbal, T. M. Przytycka and M. C. T. Hartman, *ACS Chem. Biol.*, 2017, **12**, 795-804.
37. J. M. Rogers and H. Suga, *Org. Biomol. Chem.*, 2015, **13**, 9353-9363.
38. Y. Goto, A. Ohta, Y. Sako, Y. Yamagishi, H. Murakami and H. Suga, *ACS Chem. Biol.*, 2008, **3**, 120-129.
39. E. Schmitt, M. Panvert, S. Blanquet and Y. Mechulam, *EMBO J.*, 1998, **17**, 6819-6826.
40. Y. Goto and H. Suga, *J. Am. Chem. Soc.*, 2009, **131**, 5040-5041.
41. R. Takatsuji, K. Shinbara, T. Katoh, Y. Goto, T. Passioura, R. Yajima, Y. Komatsu and H. Suga, *J. Am. Chem. Soc.*, 2019, **141**, 2279-2287.

42. Y. Ohshiro, E. Nakajima, Y. Goto, S. Fuse, T. Takahashi, T. Doi and H. Suga, *ChemBioChem*, 2011, **12**, 1183-1187.
43. K. Torikai and H. Suga, *J. Am. Chem. Soc.*, 2014, **136**, 17359-17361.
44. T. Kawakami, K. Ogawa, T. Hatta, N. Goshima and T. Natsume, *ACS Chem. Biol.*, 2016, **11**, 1569-1577.
45. O. Ad, K. S. Hoffman, A. G. Cairns, A. L. Featherston, S. J. Miller, D. Söll and A. Schepartz, *ACS Cent. Sci.*, 2019, **5**, 1289-1294.
46. J. Lee, K. E. Schwieter, A. M. Watkins, D. S. Kim, H. Yu, K. J. Schwarz, J. Lim, J. Coronado, M. Byrom, E. V. Anslyn, A. D. Ellington, J. S. Moore and M. C. Jewett, *Nat. Commun.*, 2019, **10**, 5097.
47. Y. Yin, N. Ochi, T. W. Craven, D. Baker, N. Takigawa and H. Suga, *J. Am. Chem. Soc.*, 2019, 141, 19193-19197.
48. G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933-5941.
49. J. M. Rogers, S. Kwon, S. J. Dawson, P. K. Mandal, H. Suga and I. Huc, *Nat. Chem.*, 2018, **10**, 405-412.
50. C. Tsiamantas, S. Kwon, C. Douat, I. Huc and H. Suga, *Chem. Comm.*, 2019, 55, 7366--7369.
51. Y. Yamagishi, I. Shoji, S. Miyagawa, T. Kawakami, T. Katoh, Y. Goto and H. Suga, *Chem. Biol.*, 2011, **18**, 1562-1570.