

Since this genetic code reprogramming system lacks competitors for the incorporations of nonproteinogenic amino acids, it is possible to perform multiple incorporations or the polymerization of monomers that would alter the backbone from the ordinary peptide bond to other types of bonds. For instance, we have recently performed mRNA-directed polyester synthesis using seven varieties of α -hydroxy acids that were each individually assigned to one of seven codons (Ohta et al. 2007). This work represents the first demonstration of up to 12 consecutive additions of α -hydroxy acids with various different compositions designated by codons on mRNA. We also recently performed the mRNA-directed synthesis of linear and cyclic *N*-methyl-peptides (Kawakami et al. 2008). As a proof-of-concept experiment, up to 10 successive incorporations of *N*²-methylated amino acids with six different kinds of side-chains were performed. In this work, we also combined this method with the aforementioned thioether cyclization method using the ClAc group, showing that cyclic *N*-methyl-peptides with three or four *N*-methylated backbone residues could be synthesized. This work has opened the possibility of the mRNA-encoded synthesis of *N*-methyl-peptide libraries for the screening of biologically active molecules.

Conclusion and perspectives

Since the advent of the combinatorial synthesis of peptides, peptides have drawn significant attention with the anticipation of finding novel drugs. However, only a few therapeutic peptides have reached the market because of problems associated with poor stability and cell-membrane permeability. A lesson from peptide-like natural products has given a possible direction; the peptide should have non-standard structures with not only side-chains but a cyclic and *N*²-methylated backbone, as well. We now have a new tool for synthesizing such nonstandard peptides in a cell-free translation system so that their sequences can be encoded by oligonucleotides (mRNA or cDNA) that are readily amplifiable and sequence readable by conventional molecular biological techniques.

A controllable translation system and a flexible tool for the preparation of a wide variety of nonstandard aminoacyl-tRNAs are essential to genetic code reprogramming. As described here, the combined use of wPURE system and the flexizyme technology is a reliable methodology for the ribosomal syntheses of nonstandard peptides containing the desired modifications (Fig. 3). By means of genetic code reprogramming, we can expect a whole new era of peptide chemistry that is compatible not only with cell-based high-throughput screening, but with in vitro selection methods, such as mRNA or ribosome display against isolated targets, as well. We are on the verge of witnessing the discovery of active nonstandard peptides against various therapeutic targets.

Acknowledgements

This work was supported by grants from the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) and the US National Institutes of Health (GM59159).

References

- Ames, R.S., Sarau, H.M., Chambers, J.K., Willette, R.N., Alyar, N.V., Romanic, A.M., et al. 1999. Human uterotensin-II is a potent vasoconstrictor and agonist for the orphan receptor GPR14. *Nature*, **401**: 282–286. doi:10.1038/45809. PMID:10499587.
- Arkin, M.R., and Wells, J.A. 2004. Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nat. Rev. Drug Discov.* **3**: 301–317. doi:10.1038/nrd1343. PMID:15060526.
- Chapeville, F., Ehrenstein, G.V., Benzer, S., Weisblum, B., Ray, W.J., and Lipmann, F. 1962. On role of soluble ribonucleic acid in coding for amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **48**: 1086–1092. doi:10.1073/pnas.48.6.1086. PMID:13878159.
- Coulouarn, Y., Lihmann, I., Jegou, S., Anouar, Y., Tostivint, H., Beauvillain, J.C., et al. 1998. Cloning of the cDNA encoding the uterotensin II precursor in frog and human reveals intense expression of the uterotensin II gene in motoneurons of the spinal cord. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 15803–15808. doi:10.1073/pnas.95.26.15803. PMID:9861051.
- Ehrenstein, G., Weisblum, B., and Benzen, S. 1963. Function of sRNA as amino acid adaptor in synthesis of hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **49**: 669–675. doi:10.1073/pnas.49.5.669. PMID:16591086.
- Fischbach, M.A., and Walsh, C.T. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **106**: 3468–3496. doi:10.1021/cr0503097. PMID:16895337.
- Forster, A.C., Tan, Z.P., Nalam, M.N.L., Lin, H.N., Qu, H., Cornish, V.W., and Blacklow, S.C. 2003. Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 6353–6357. doi:10.1073/pnas.1132122100. PMID:12754376.
- Goto, Y., Ohta, A., Sako, Y., Yamagishi, Y., Murakami, H., and Suga, H. 2008. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**: 120–129. doi:10.1021/cb700233t. PMID:18215017.
- Hahn, M., and Stachelhaus, T. 2006. Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 275–280. doi:10.1073/pnas.0508409103. PMID:16407157.
- Hecht, S.M., Alford, B.L., Kuroda, Y., and Kitano, S. 1978. "Chemical aminoacylation" of tRNA's. *J. Biol. Chem.* **253**: 4517–4520. PMID:248056.
- Heckler, T.G., Chang, L.H., Zama, Y., Naka, T., Chorghade, M.S., and Hecht, S.M. 1984. T4 RNA ligase mediated preparation of novel "chemically miscacylated" tRNAs. *Biochemistry*, **23**: 1468–1473. doi:10.1021/bi00302a020. PMID:6372858.
- Hohsaka, T., Ashizuka, Y., Murakami, H., and Sisido, M. 1996. Incorporation of nonnatural amino acids into streptavidin through in vitro frame-shift suppression. *J. Am. Chem. Soc.* **118**: 9778–9779. doi:10.1021/ja9614225.
- Johnson, D.L., and Jolliffe, L.K. 2000. Erythropoietin mimetic peptides and the future. *Nephrol. Dial. Transplant.* **15**: 1274–1277. doi:10.1093/ndt/15.9.1274. PMID:10978375.
- Josephson, K., Hartman, M.C.T., and Szostak, J.W. 2005. Ribosomal synthesis of unnatural peptides. *J. Am. Chem. Soc.* **127**: 11727–11735. doi:10.1021/ja0515809. PMID:16104750.
- Kawakami, T., Murakami, H., and Suga, H. 2008. Messenger RNA-programmed incorporation of multiple *N*-methyl-amino acids into linear and cyclic peptides. *Chem. Biol.* **15**: 32–42. doi:10.1016/j.chembiol.2007.12.008. PMID:18215771.
- Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., et al. 2002. An engineered *Escherichia coli* tyrosyl-

- tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 9715–9720. doi:10.1073/pnas.142220099. PMID:12097643.
- Leader, B., Baca, Q.J., and Golan, D.E. 2008. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* **7**: 21–39. doi:10.1038/nrd2399. PMID:18097458.
- Lo Conte, L., Chothia, C., and Janin, J. 1999. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**: 2177–2198. doi:10.1006/jmbi.1998.2439. PMID:9925793.
- Murakami, H., Ohta, A., Ashigai, H., and Suga, H. 2006. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods*, **3**: 357–359. doi:10.1038/nmeth877. PMID:16628205.
- Naranda, T., Wong, K., Kaufman, R.I., Goldstein, A., and Olsson, L. 1999. Activation of erythropoietin receptor in the absence of hormone by a peptide that binds to a domain different from the hormone binding site. *Proc. Natl. Acad. Sci. U.S.A.* **96**: 7569–7574. doi:10.1073/pnas.96.13.7569. PMID:10377456.
- Noren, C.J., Anthonycahill, S.J., Griffith, M.C., and Schultz, P.G. 1989. A general method for site-specific incorporation of unnatural amino-acids into proteins. *Science*, **244**: 182–188. doi:10.1126/science.2649980. PMID:2649980.
- Ohta, A., Murakami, H., Higashimura, E., and Suga, H. 2007. Synthesis of polyester by means of genetic code reprogramming. *Chem. Biol.* **14**: 1315–1322. PMID:18096500.
- Ohtsuki, T., Manabe, T., and Sisido, M. 2005. Multiple incorporation of non-natural amino acids into a single protein using tRNAs with non-standard structures. *FEBS Lett.* **579**: 6769–6774. doi:10.1016/j.febslet.2005.11.010. PMID:16310775.
- Parmley, S.F., and Smith, G.P. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, **73**: 305–318. doi:10.1016/0378-1119(88)90495-7. PMID:3149606.
- Pero, S.C., Oligino, L., Daly, R.J., Soden, A.L., Liu, C., Roller, P.P., et al. 2002. Identification of novel non-phosphorylated ligands, which bind selectively to the SH2 domain of Grb7. *J. Biol. Chem.* **277**: 11918–11926. doi:10.1074/jbc.M111816200. PMID:11809769.
- Robertson, S.A., Noren, C.J., Anthonycahill, S.J., Griffith, M.C., and Schultz, P.G. 1989. The use of 5'-phospho-2 deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of transfer-RNA. *Nucleic Acids Res.* **17**: 9649–9660. doi:10.1093/nar/17.23.9649. PMID:2602139.
- Saito, H., Kourouklis, D., and Suga, H. 2001. An in vitro evolved precursor tRNA with aminoacylation activity. *EMBO J.* **20**: 1797–1806. doi:10.1093/emboj/20.7.1797. PMID:11285242.
- Schwarzer, D., Finking, R., and Marahiel, M.A. 2003. Nonribosomal peptides: from genes to products. *Nat. Prod. Res.* **20**: 275–287. doi:10.1039/b111145k.
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. 2001. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**: 751–755. doi:10.1038/90802. PMID:11479568.
- Turk, B.E., and Cantley, L.C. 2003. Peptide libraries: at the crossroads of proteomics and bioinformatics. *Curr. Opin. Chem. Biol.* **7**: 84–90. doi:10.1016/S1367-5931(02)00004-2. PMID:12547431.
- Wang, L., and Schultz, P.G. 2005. Expanding the genetic code. *Angew. Chem. Int. Ed.* **44**: 34–66. doi:10.1002/anie.200460627.
- Weiss, G.A., and Chamberlin, R. 2003. Bridging the synthetic and biopolymer worlds with peptide-drug conjugates. *Chem. Biol.* **10**: 201–202. doi:10.1016/S1074-5521(03)00056-5. PMID:12670531.
- Wrighton, N.C., Farrell, F.X., Chang, R., Kashyap, A.K., Barbone, F.P., Mulcahy, L.S., et al. 1996. Small peptides as potent mimetics of the protein hormone erythropoietin. *Science*, **273**: 458–463. doi:10.1126/science.273.5274.458. PMID:8662529.
- Yoo, T.H., and Tirrell, D.A. 2007. High-throughput screening for methionyl-tRNA synthetases that enable residue-specific incorporation of noncanonical amino acids into recombinant proteins in bacterial cells. *Angew. Chem. Int. Ed.* **46**: 5340–5343. doi:10.1002/anie.200700779.
- Zhang, Y., and Gladyshev, V.N. 2007. High content of proteins containing 21st and 22nd amino acids, selenocysteine and pyrrolysine, in a symbiotic deltaproteobacterium of gutless worm *Olavius algarvensis*. *Nucleic Acids Res.* **35**: 4952–4963. doi:10.1093/nar/gkm514. PMID:17626042.
- Zhang, Z., Alfonta, L., Tian, F., Bursulaya, B., Uryu, S., King, D.S., and Schultz, P.G. 2004. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **101**: 8882–8887. doi:10.1073/pnas.0307029101. PMID:15187228.

Ribosomal Synthesis of Bicyclic Peptides via Two Orthogonal Inter-Side-Chain Reactions

Yusuke Sako,^{1,†} Jumpei Morimoto,^{1,‡} Hiroshi Murakami,¹ and Hiroaki Suga^{*1,†,§}

Research Center for Advanced Science and Technology, The University of Tokyo, Komaba, Meguro, Tokyo 153-8904, Japan, Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo, Tokyo 113-8656, Japan, and Department of Advanced Interdisciplinary Studies, Graduate School of Engineering, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo 153-8904, Japan

Received February 6, 2008; E-mail: hsuga@rcast.u-tokyo.ac.jp

Many naturally occurring peptides possess two pairs of disulfide bonds, representing a bicyclic structural feature.¹ Such a structure rigidifies the peptide conformation into a specific active form and consequently plays a crucial role in the biological function of the peptide. To form the two disulfide bonds, two particular pairs of cysteines (Cys) out of four Cys residues must selectively react with each other, avoiding two other possible combinations² (Figure 1A, a–d = Cys). Specific formation of such disulfide bonds in naturally occurring peptides is generally dictated by the proximity of the Cys residues in the tertiary structure³ or in cells it is occasionally driven by a disulfide isomerase that rearranges kinetically trapped incorrect disulfide pairs to the correct pairs.⁴ Although it is of great interest to generate libraries of such bicyclic peptides using the translation apparatus and subsequently screen their biological functions, upon randomization of the sequence, the disulfide bond formation would very likely be scrambled (Figure 1A); thus the 4-Cys-containing peptide scaffold cannot be utilized as a reliable bicyclic platform. We here report a new peptide scaffold containing a single Cys and three nonproteinogenic amino acids (Figure 1B), selectively forming the desirable cross-links. Significantly, this peptide scaffold can be expressed by using a reconstituted *Escherichia coli* cell-free translation system (called PURE⁵) under the reprogrammed genetic code. Therefore, it is readily applicable to the preparation of libraries of bicyclic peptides with a uniform skeleton.

We designed and synthesized two pairs of amino acids containing orthogonal functional groups (Figure 1B). 4-(2-Chloroacetyl)aminobutyric acid (Cab) bears the thiol-reactive functional group in its side chain. We have previously reported that, upon the site-specific incorporation of Cab paired with a Cys into the peptide chain, these two side chains spontaneously react with each other to give a thioether linkage.⁶ Because this chemistry requires no additional reagent, the Cab–Cys pair is a convenient and reliable amino acid pair to generate the desired monocyclic peptides closed by a nonreducible bond. The functional groups of the other pair, azidohomoalanine (Aha) and propargylglycine (Pgl), are known to react with each other via Cu(I)-catalyzing azide–alkyne cycloaddition, and this type of chemistry has been extensively used in bioconjugate applications.⁷ Despite the fact that the individual molecules were shown to be compatible with translation,⁸ their site-specific double incorporations have not yet been demonstrated. Therefore, we first attempted to establish the assignment of Aha and Pgl to certain codons by genetic code reprogramming and demonstrate their double incorporations into the nascent chain of a model peptide, and then perform the intramolecular cyclization.

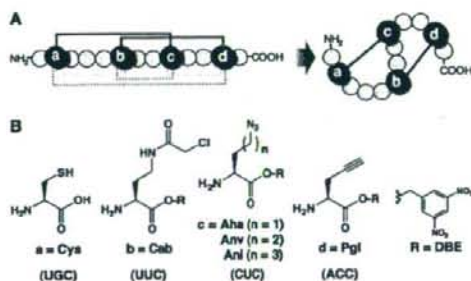


Figure 1. Ribosomal synthesis of bicyclic peptides. (A) Schematic representation of a bicyclic peptide closed by two orthogonal inter-side-chain reactions. (B) Structure and assignment of amino acids to the reprogrammed genetic code. Cab, 4-(2-Chloroacetyl)aminobutyric acid; Aha, azidohomoalanine; Anv, azidonorvaline; Anl, azidonorleucine; Pgl, propargylglycine; DBE, 3,5-dinitrobenzyl ester. Codon assignment of each amino acid is shown in parentheses.

In the present study, we chose the Leu codon (CUC) for the assignment of Aha and the Thr codon (ACC) for Pgl. Each amino acid activated with a 3,5-dinitrobenzyl ester (DBE) (Figure 1B) was charged onto orthogonal tRNA^{Aaa-E1} using flexizyme,⁹ a tRNA acylation ribozyme (Supporting Information Figure S1). Then, these aminoacyl-tRNAs, Aha-tRNA^{Aaa-E1}_{GAG} and Pgl-tRNA^{Aaa-E1}_{GGU}, were added in a PURE system in which both Thr and Leu were withdrawn (referred to as wPURE system). We observed double incorporations of ⁵Aha and ¹¹Pgl into the nascent chain of a model peptide 1 (Figure 2) in 36% yield relative to the yield of wild-type expressed under the normal genetic code (Supporting Information Figure S2A,B); this value translated to the quantity of 1 with an approximately 0.64 pmol/μL compared with 1.75 pmol/μL of the wild-type.¹⁰ MALDI-TOF analysis of the peptide showed a single major peak corresponding to 1 accompanied by a minor peak corresponding to an amine molecule 3 (see more details in the legend of Figure 2).¹¹ We also demonstrated that the replacement of Aha with its derivatives bearing a longer arm (Figure 1B, Anv, n = 2, and Anl, n = 3) yielded the desired peptides with nearly equivalent quantities and qualities to 1 (Supporting Information Figure S2B,C). Thus, the genetic code reprogramming approach enables us to prepare Aha/Pgl-containing peptides.

To test if the inter-side-chain closure of ⁵Aha/¹¹Pgl-containing peptide 1 could be achieved by the cycloaddition, 1 was treated with CuSO₄ and ascorbate (Figure 2, 1→2). Because we expected that both the unreacted linear form and the cyclic form generated by the cycloaddition would show no change in molecular weight, the occurrence of cycloaddition was confirmed by reducing the azide

[†] Research Center for Advanced Science and Technology.

[‡] Department of Chemistry and Biotechnology.

[§] Department of Advanced Interdisciplinary Studies.

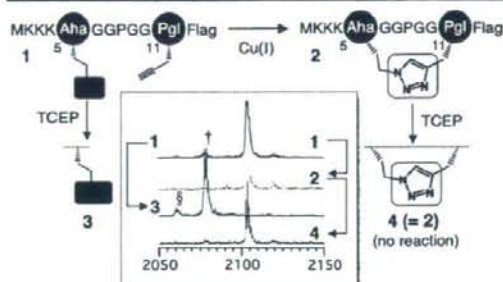


Figure 2. Cyclization via the Cu(I)-catalyzed azide-alkyne cycloaddition. The observed molecular mass (m_{obs}) of 1 is 2103.52 Da, while the calculated molecular mass (m_{calc}) is 2102.96 Da. Likewise, 2 is $m_{\text{obs}} = 2103.47$ and $m_{\text{calc}} = 2102.96$, while 3 is $m_{\text{obs}} = 2076.54$ Da and $m_{\text{calc}} = 2076.97$ Da. Upon treatment of 2 with TCEP, no change in m_{obs} was detected ($m_{\text{obs}} = 2102.56$). The minor peak shown by † in the spectrum of 1 is consistent with m_{obs} of 3, possibly generated by the reduction of the azide group by DTT and/or mercaptoethanol contained in the translation buffer.¹¹ ‡ corresponds to a peak of an imido ester likely produced by a similar mechanism reported by Back et al.¹²

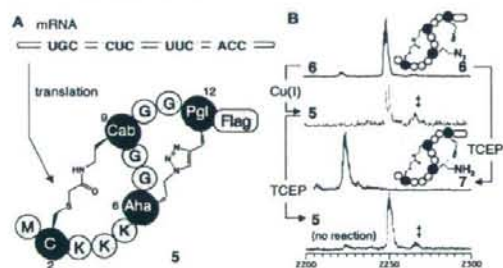


Figure 3. Ribosomal synthesis of bicyclic peptides. (A) Schematic structure of the bicyclic peptide 5 containing ⁹⁹Cab, ⁶Aha, and ¹²Pgl. (B) MALDI-TOF analysis of the nascent peptide 6 ($m_{\text{obs}} = 2248.29$ Da, $m_{\text{calc}} = 2248.97$ Da), which was spontaneously cyclized by the thioether bond between Cab and Cys side chains. TCEP reduction of peptide 6 resulted in the formation of the reduced peptide 7 ($m_{\text{obs}} = 2222.94$ Da, $m_{\text{calc}} = 2222.98$ Da), while bicyclic peptide 5 produced by the copper-catalyzed cycloaddition of 6 was not reduced ($m_{\text{obs}} = 2248.36$). † denotes a peak of peptide 5 in which Met was oxidized during the MALDI-TOF analysis.

group by incubating with tris(carboxyethyl)phosphine (TCEP). Prior to the cycloaddition, the azide group of 1 was readily reduced by TCEP to yield the corresponding amine 3, consistent with its MALDI-TOF analysis (Figure 2, 1–3). In contrast, the TCEP treatment of the product generated by incubation of 1 with Cu(I) showed no change in the molecular weight (Figure 2, 2–4). This result confirmed that 1 was converted to 2 by the cycloaddition. Likewise, ⁵Anv/¹¹Pgl- or ⁵Anl/¹¹Pgl-containing peptide, in which the ring structure was assisted by a longer arm, also afforded the corresponding cyclic peptide upon the Cu(I)-catalyzed reaction (Supporting Information Figure S4).

With the above cyclization methodology in our hands, we next applied it to the synthesis of bicyclic peptides in combination with the intramolecular thioether formation between Cab and Cys. We prepared an mRNA template for the expression of peptide 5 containing ²Cys, ⁶Aha, ⁹⁹Cab, and ¹²Pgl in the peptide sequence (Figure 3A). ⁹⁹Cab was assigned to a Phe codon (UUC) by charging onto tRNA^{Asn-E1}_{GAA} using flexizyme. We then performed the triple sense suppressions of ⁶Aha, ¹²Pgl, and ⁹⁹Cab using a wPURE system lacking Leu, Thr, and Phe. MALDI-TOF analysis of the peptide

showed a major peak corresponding to the monocyclic 6 closed by a thioether bond between the side chains of ²Cys and ⁹⁹Cab (Figure 3B, 6). These results indicated that the "first" eight-membered ring cyclization occurred in the translation mixture in situ cleanly. 6 was then subjected to the "second" seven-membered ring cyclization step in the presence of Cu(I) to yield 5, which was further treated with TCEP. As expected, the molecular weight of the product (5) did not show any changes before or after TCEP treatment (Figure 3B, 6→5), as opposed to the observation that the Cu(I)-untreated 6 was reduced to the corresponding amine 7 by TCEP (6→7). This result indicates that peptide 5 has the expected bicyclic structure with an eight/seven-membered ring fusion.

Using the same DNA template that expressed 5, Aha was substituted with Anv or Anl simply altering the use of the corresponding Anv- or Anl-tRNA^{Asn-E1}_{GAA} in the translation. The respective precursor peptide containing ²Cys, ⁹⁹Cab, and ¹²Pgl along with ⁶Anv or ⁶Anl was expressed under the reprogrammed genetic code and treated with Cu(I). Again, to ensure the cycloaddition, these peptides as well as the precursor monocyclic peptides were treated with TCEP, and the resulting peptides were analyzed by MALDI-TOF (Supporting Information Figure S5). The data were indeed consistent with the idea that the desired bicyclic peptides were produced upon the Cu(I) treatment.

Encouraged by the above observations, we next attempted to apply this methodology to the expression of peptides composed of different sequences with other combinations of ring sizes. We designed three DNA templates (see Tables 1 and 2 in Supporting Information) that would produce precursor peptides 8, 11, and 14 (Figure 4A–C) under the assignments of Aha, Cab, and Pgl to the Leu, Phe, and Thr codons. It was expected that these precursor peptides were spontaneously closed by the first sulfhydryl–chloroacetyl cyclization upon the translation and then converted to bicyclic peptides 10, 13, and 16, respectively, via the second azide–alkyne cycloaddition catalyzed by Cu(I) (Figure 4A–C, left panels). Consequently, peptide 10 had the same bicyclic structure as 5 with an eight/seven-membered ring fusion, but its sequence composition of amino acids in each ring differed from that of 5. Moreover, peptides 13 and 16 were composed of not only different amino acid sequences but also different bicyclic structures with eight/eight- and seven/seven-membered ring fusions, respectively. To confirm the occurrence of azide–alkyne cycloaddition, the precursor as well as cyclized peptides were treated with TCEP, and these peptides were analyzed by MALDI-TOF (Figure 4A–C, right panels). As expected, the precursor peptides 8, 11, and 14 were reduced by TCEP to yield the corresponding monocyclic amino peptides, 9, 12, and 15. In contrast, no change in molecular weight was observed for the Cu(I)-treated peptides, indicating that the cyclization of 8, 11, and 14 took place to yield bicyclic peptides 10, 13, and 16, respectively. These results have clearly shown the versatility of this methodology to the synthesis of bicyclic peptides with various sequences and ring sizes in the 7/7, 8/7, and 8/8 fusions.

In conclusion, we have developed a new methodology to construct bicyclic scaffolds by the combination of peptide translation under a reprogrammed genetic code with post-translational cyclizations involving two pairs of amino acids, Cys–Cab and Aha–Pgl. Because the sulfhydryl–chloroacetyl reaction and the azide–alkyne cycloaddition are orthogonal, they act as the specific pairs to yield only the desirable bicyclic peptide with fairly clean conversions of each step. Most importantly, we have shown the versatility of this methodology to the peptide synthesis of various bicyclic peptides bearing fused ring scaffolds, which are controlled by the side chain length of the azide-containing nonproteinogenic amino acids and even the amino acid sequence compositions between the cyclizing

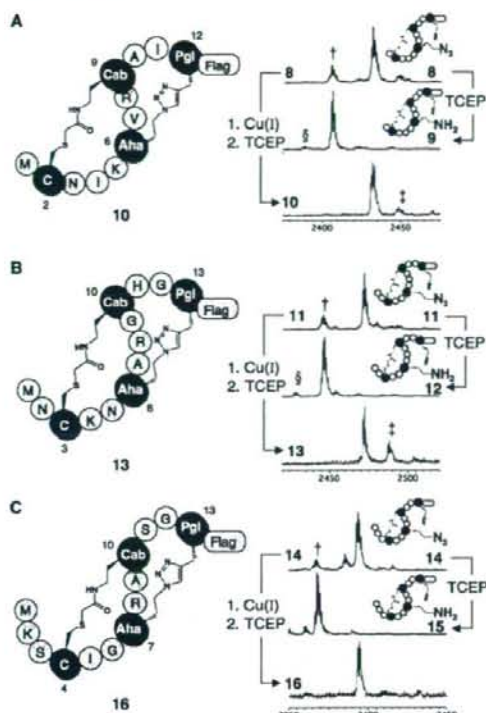


Figure 4. Ribosomal synthesis of bicyclic peptides from three different template DNAs. Left panel shows the schematic structures of bicyclic peptides 10, 13, and 16. Right panel shows MALDI-TOF analysis of the precursor monocyclic peptides 8, 11, and 14, the TCEP-reduced peptides 9, 12, and 15, and Cu(I) followed by TCEP-treated peptides 10, 13, and 16. (A) Formation of bicyclic peptide 10. Molecular weights of the observed peaks are as follows: 8, $m_{\text{obs}} = 2430.58$ Da ($m_{\text{calc}} = 2431.12$); 9, $m_{\text{obs}} = 2405.82$ Da ($m_{\text{calc}} = 2405.13$); bicyclic peptide 10, $m_{\text{obs}} = 2431.46$ Da ($m_{\text{calc}} = 2431.12$). The minor peak indicated by † ($m_{\text{obs}} = 2404.57$ Da) was possibly generated due to the reduction of the azide group by DTT and/or mercaptoethanol contained in the translation buffer.¹¹ The minor peak indicated by ‡ ($m_{\text{obs}} = 2448.52$ Da) was likely generated by the oxidation of Met of 10 during the MALDI-TOF analysis; § denotes an imido ester produced by a similar mechanism reported by Back et al.¹² (B) Formation of bicyclic peptide 13. Molecular weights of the observed peaks are as follows: 11, $m_{\text{obs}} = 2470.67$ Da ($m_{\text{calc}} = 2471.02$); 12, $m_{\text{obs}} = 2444.85$ Da ($m_{\text{calc}} = 2445.03$); bicyclic peptide 13, $m_{\text{obs}} = 2471.23$ Da ($m_{\text{calc}} = 2471.02$). The minor peaks shown by † and ‡ are $m_{\text{obs}} = 2444.58$ and 2487.75 Da, respectively. (C) Formation of bicyclic peptide 16. Molecular weights of the observed peaks are as follows: 14, $m_{\text{obs}} = 2392.76$ Da ($m_{\text{calc}} = 2393.03$); 15, $m_{\text{obs}} = 2366.95$ Da ($m_{\text{calc}} = 2367.04$); bicyclic peptide 16, $m_{\text{obs}} = 2393.16$ Da ($m_{\text{calc}} = 2393.03$). The minor peak indicated by † corresponds to $m_{\text{obs}} = 2366.79$ Da, whereas the other minor peak observed near 2390 Da was unknown.

residues. Notably, the libraries of bicyclic peptides can be easily constructed by synthesis of the corresponding mRNA libraries. Screening of such peptide libraries with uniform bicyclic scaffolds, perhaps in combination with an appropriate *in vitro* peptide-display

technique, will quickly lead us to discover unique peptidyl drugs against various therapeutic targets.

Acknowledgment. We thank Dr. P. C. Reid for critical proofread. This work was supported by grants of Japan Society for the promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H.S., Grants-in-Aid for JSPS Fellows to Y.S. (17-11918), and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Supporting Information Available: Experimental details and additional data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) (a) Forte, L. R.; London, R. M.; Freeman, R. H.; Krause, W. *J. Am. Chem. Soc.* **2000**, *122*, F180–F191. (b) Sakurai, T.; Ameyama, A.; Ishii, M.; Matsuzaki, I.; Chemelli, R. M.; Tanaka, H.; Williams, S. C.; Richardson, J. A.; Kozlowski, G. P.; Wilson, S.; Arch, J. R.; Buckingham, R. E.; Haynes, A. C.; Carr, S. A.; Annan, R. S.; McNulty, D. E.; Liu, W. S.; Terrett, J. A.; Elshourbagy, N. A.; Bergsma, D. J.; Yanagisawa, M. *Cell* **1998**, *92*, 573–585. (c) Yanagisawa, M.; Kurihara, H.; Kimura, S.; Tomobe, Y.; Kobayashi, M.; Mitsui, Y.; Yazaki, Y.; Goto, K.; Masaki, T. *Nature* **1988**, *332*, 411–415.
- (2) Badock, V.; Raida, M.; Adermann, K.; Forstmann, W. G.; Schrader, M. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1952–1956.
- (3) (a) Thornton, J. M. *J. Mol. Biol.* **1981**, *151*, 261–287. (b) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167–339.
- (4) (a) Ellgaard, L.; Ruddock, L. W. *EMBO Rep.* **2005**, *6*, 28–32. (b) Gilbert, H. F. *J. Biol. Chem.* **1997**, *272*, 29399–29402.
- (5) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001**, *19*, 751–755.
- (6) Sako, Y.; Goto, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 241–249.
- (7) (a) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192–3193. (b) Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686–4687. (c) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793–16797. (d) Dieterich, D. C.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9482–9487. (e) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164–11165.
- (8) (a) Link, A. J.; Vink, M. K.; Tirrell, D. A. *J. Am. Chem. Soc.* **2004**, *126*, 10598–10602. (b) Forster, A. C.; Tan, Z.; Nalam, M. N.; Lin, H.; Qu, H.; Cornish, V. W.; Blocker, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6353–6357.
- (9) (a) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357–359. (b) Ohuchi, M.; Murakami, H.; Suga, H. *Curr. Opin. Chem. Biol.* **2007**, *11*, 135–144. (c) Kang, T.-J.; Suga, H. *Biochem. Cell Biol.* **2008**, *86*, 92–99. (d) Ohta, A.; Yamagishi, Y.; Suga, H. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159–167. (e) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 120–129. (f) Kawakami, T.; Murakami, H.; Suga, H. *Chem. Biol.* **2008**, *15*, 32–42. (g) Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. *Chem. Biol.* **2007**, *14*, 1315–1322.
- (10) Our general method for the peptide quantification was as follows: Radioisotope (RI) counts of [¹⁴C]-Asp were plotted against its known concentrations to determine the calibration line. Then, RI counts of the expressed wild-type peptide and cyclic peptides were fitted to the calibration line to estimate the individual concentration of peptides. We found that this quantification method was reliable to assess the quantity of peptide previously proven by performing activity assays using a known bioactive peptide.⁵
- (11) During the translation, the azide group of Aha, Anv, and Anl was very likely reduced by DTT and/or mercaptoethanol included in the translation buffer. See the following reference as an example of chemistry: Meijnjohanns, E.; Meldal, M.; Jensen, T.; Werdelin, O.; GalliStampino, L.; Mouritsen, S.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1997**, 871–884.
- (12) Back, J. W.; David, O.; Kramer, G.; Masson, G.; Kasper, P. T.; de Koning, L. J.; de Jong, L.; van Maarseveen, J. H.; de Koster, C. G. *Angew. Chem., Int. Ed.* **2005**, *44*, 7946–7950. We found that the reported mechanism herein also proceeds in the translated peptides containing appropriate nonproteogenic amino acids (Nakajima, E.; et al. unpublished results). More detailed studies of this work will be reported elsewhere.

JA800953C

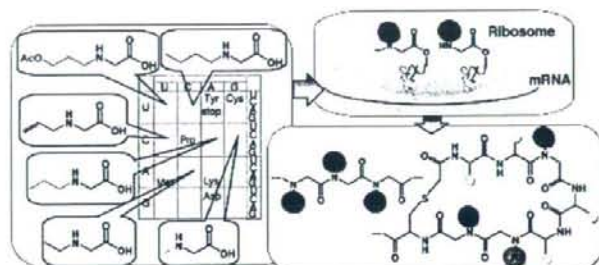
Communication

Ribosomal Synthesis of Polypeptoids and Peptoid#Peptide Hybrids

Takashi Kawakami, Hiroshi Murakami, and Hiroaki Suga

J. Am. Chem. Soc., 2008, 130 (50), 16861-16863 • DOI: 10.1021/ja806998v • Publication Date (Web): 19 November 2008

Downloaded from <http://pubs.acs.org> on February 19, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



ACS Publications
High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society, 1155 Sixteenth Street N.W., Washington, DC 20036

Ribosomal Synthesis of Polypeptoids and Peptoid–Peptide Hybrids

Takashi Kawakami,[†] Hiroshi Murakami,[‡] and Hiroaki Suga^{*,†,‡}

Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan, and Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

Received September 4, 2008; E-mail: hsuga@rcast.u-tokyo.ac.jp

Peptoids are artificially designed synthetic oligomers composed of *N*-substituted glycine building blocks.¹ It has been shown that peptoids exhibit unique pharmacokinetic properties, different from ordinary peptides, with greater protease resistance and membrane permeability being attributed to the lack of protons on the amide bond.² The *N*-substituted groups represent the secondary structural diversity of peptoids like peptide sidechains, allowing for construction of libraries and leading to the discovery of aptamers (or inhibitors) against therapeutic targets.³ However, peptoids have intrinsic poor plasticity to form a variety of 3D structures, and therefore the 3D structural diversity in the peptoid library would be limited. Alternatively, a peptoid–peptide hybrid would be an attractive framework to improve the plasticity perhaps without sacrificing the major pharmacokinetics of the peptoid. However, such an approach to date relied on a semirational design of hybrids based on an active peptide followed by their activity validation,⁴ giving only a limited number of successes. Clearly, a new synthetic strategy of peptoids or peptoid–peptide hybrids, which is readily extendable to their diverse library construction and screening (or selection) of active ligands, is awaited for the discovery of potent molecules. We here report a novel method to express polypeptoids and peptoid–peptide hybrids by means of translation machinery under the reprogrammed genetic code, allowing for the mRNA-directed synthesis of such oligomers with linear and cyclic scaffolds.

Despite the fact that *N*-methyl- α -amino acids are difficult to efficiently incorporate without prematurely truncated products or oligomerize into a nascent peptide chain using the classical amber suppression method,⁵ we and other groups recently found that nearly a dozen *N*-methyl- α -amino acids derived from proteinogenic or nonproteinogenic amino acids could be incorporated by means of translation machinery using genetic code reprogramming.⁶ Success, at least in our case, relied on the integration of two systems. One is a reconstituted *E. coli* cell-free translation system in which certain proteinogenic amino acids and/or cognate aminoacyl-tRNA synthetases are withdrawn (referred to as *withdrawn* PURE system; wPURE), allowing us to diminish the competing background incorporation of the proteinogenic amino acids in the translation elongation event.⁷ The other is a system involving artificial tRNA acylation ribozymes, referred to as flexizymes.⁸ This has greatly facilitated the preparation of a wide variety of tRNAs charged with nonproteinogenic amino acids.^{9e,9c,9} During our studies, we found the flexizyme system was able to charge *N*-substituted glycines (rGly) onto tRNAs efficiently (Supporting Information, Figure S1). We thus became interested in investigating if rGly could be incorporated into a nascent peptide chain under the reprogrammed genetic code.

To investigate the incorporation efficiencies of rGly into a peptide chain by translation machinery, we chose a systematic set of rGly

bearing various *N*-substituted groups (Figure 1A). The *N*-substituted groups were divided into three families, (1) nonbranched alkyl

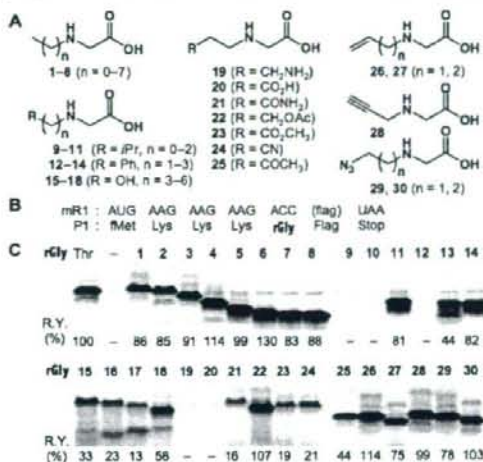


Figure 1. Tolerance of *N*-substituted glycines (rGly) in ribosomal peptide synthesis. (A) Chemical structure of rGly used in this study. 1, *N*-methyl-glycine (Nme); 2, *N*-ethyl-glycine (Nme); 3, *N*-propyl-glycine (Npr); 4, *N*-butyl-glycine (Nbu); 5, *N*-pentyl-glycine; 6, *N*-hexyl-glycine; 7, *N*-heptyl-glycine; 8, *N*-octyl-glycine; 9, *N*-isopropyl-glycine; 10, *N*-isobutyl-glycine; 11, *N*-isopentyl-glycine; 12, *N*-benzyl-glycine; 13, *N*-(2-phenylethyl)-glycine; 14, *N*-(3-phenylpropyl)-glycine; 15, *N*-(3-hydroxypropyl)-glycine; 16, *N*-(4-hydroxybutyl)-glycine; 17, *N*-(5-hydroxypentyl)-glycine; 18, *N*-(6-hydroxyhexyl)-glycine; 19, *N*-(3-aminopropyl)-glycine; 20, *N*-(2-carboxylethyl)-glycine; 21, *N*-(2-carbamoyl-ethyl)-glycine; 22, *N*-(3-acetoxypropyl)-glycine (Naco); 23, *N*-(2-methoxycarbonyl-ethyl)-glycine; 24, *N*-(2-cyanoethyl)-glycine; 25, *N*-(3-oxobutyl)-glycine; 26, *N*-allyl-glycine (Nall); 27, *N*-3-butenyl-glycine; 28, *N*-propargyl-glycine; 29, *N*-(2-azidoethyl)-glycine; 30, *N*-(3-azidopropyl)-glycine. (B) Sequences of mRNA (mR1) that express an rGly-containing-peptide (P1). Flag in parentheses indicates the RNA sequence encoding the Flag peptide (DYKDDDDK). (C) Tricine-SDS-PAGE analysis of the expressed peptides labeled with [¹⁴C]-Asp detected by autoradiography. Lane 1, the wildtype peptide expressed in the Thr-containing ordinary PURE system where ACC assigns Thr; lane 2, a negative control using Thr-withdrawn PURE (wPURE) system in the presence of uncharged tRNA^{Met}_{UAG}; lanes 3–32, expression of the peptide containing a single rGly in the presence of designated rGly-tRNA^{Asp}_{GCU} prepared by flexizyme system. Each expression yield (R.Y.) relative to wildtype was determined by a mean score of triplicates.

chains (1–8), (2) branched alkyl chains (9–14), and (3) functionalized alkyl chains (15–30). These rGly were derived to 3,5-dinitrobenzyl esters or cyanomethyl esters, making them compatible to the flexizyme system,^{9c} and then charged onto suppressor tRNA^{Asp-E2}_{GCU}^{9a} (Figure S1). The individual rGly-tRNA^{Asp-E2}_{GCU} were added to a wPURE system lacking threonine (Thr) and surveyed for their single incorporation into a model peptide upon

[†] Department of Chemistry and Biotechnology.

[‡] Research Center for Advanced Science and Technology.

suppression of the Thr ACC codon in the designed mRNA sequence (Figure 1B, mR1 and P1). Tricine-SDS-PAGE analysis of the resulting [¹⁴C]-Asp-labeled peptides showed that the first family of rGly bearing a nonbranched alkyl chain were all incorporated as efficiently as Thr (Figure 1C, 1–8). The incorporation of the designated rGly was also confirmed by MALDI-TOF analysis, giving the expected molecular mass in all cases (Figure S2).

In the second family, three rGly (Figure 1C, 11, 13, and 14) were efficient elongators similar to those in the first family, whereas the other three (Figure 1C, 9, 10, and 12) having the branched methyl or phenyl group closer to the α-amino group were not. Most rGly in the third family exhibited acceptable incorporation efficiencies (Figure 1C, 15–18 and 21–30), in which the MALDI-TOF was consistent with the expected molecular mass (Figure S2). The exceptions were those with a negatively or positively charged alkyl group (Figure 1C, 19 and 20). In summary, our survey has shown that a wide variety of rGly, in which the *N*-substituent has no bulky group near the α-amino group and no charged functional group, act as elongators for translation machinery. This observation is similar to that for *N*-methyl-α-amino acids in which bulky or charged side chains were unfavorable elongators;^{6b} a short list of unfavorable elongators recently reported by Zhang et al. also fell into this trend, where *N*-substituted α-amino acids were bulky near the α-amino group due to the combination between a *N*-substituent and α-carbon side chain.^{6c} Therefore, this seems to be a common trend for *N*-substituted α-amino acids.

We next attempted consecutive elongations of rGly to express a short stretch of polypeptoid sequences fused with a KK-Flag peptide tag to facilitate its isolation. In this study, we chose three rGly, *Net*, *Naco*, and *Nall* (Figure 1A, 2, 22, and 26) that were assigned to arbitrarily chosen ACC, UUC, and CUC codons, respectively, with aid of the flexizyme system. Five mRNA templates (Figure 2A, mR2–6) were designed to elongate the above rGly in succession ranging from 2 to 6 times (Figure 2A, P2–P6). We also expressed peptides (Figure 2A, wt2–6) from the respective mRNA templates using the ordinary PURE system to compare the expression level of peptoids. Tricine-SDS-PAGE analysis of [¹⁴C]-Asp-labeled peptoids and peptides indicated that the expression level of peptoids gradually decreased with increasing peptoid length (lanes 2, 4, 6, 9, and 11), in contrast to the observation that peptides were expressed at a nearly consistent level independent from the length (Figure 2B, lanes 1, 3, 5, 8, and 10). Although this observation is somewhat similar to the successive elongation of *N*-methyl-α-amino acids,^{6b} the expression level of peptoids was hampered more seriously by an increase in length. In fact, when we expressed a tetra-*N*-methyl-peptide using the mR4 template (Me4, see the box in Figure 2A and B), the expression level of Me4 was nearly double that of P4 (Figure 2B, lane 6 vs 7). Both ribosome and EF-Tu may have an effect on their elongation efficiencies.¹⁰ In particular, peptidyl elongation may be attributed to steric hindrances of the peptidyl-tRNA in ribosome peptidyl-tRNA (P) site or/and the rGly-tRNA in aminoacyl-tRNA (A) site, probably making the peptidyl elongation less efficient than *N*-methyl-peptidyl elongation.

To demonstrate the expression of designer peptoids, we prepared two sequence variants of the tetrapeptoid P4 (Figure 3A, P4a and P4b) and expressed them from the respective mRNA templates (Figure 3A, mR4a and mR4b). The expression levels of these tetrapeptoids were similar to each other, and MALDI-TOF analysis of each peptoid showed a single major peak that is consistent with its expected molecular mass according to the corresponding mRNA templates (Figure 3B). To the best of our knowledge, this result

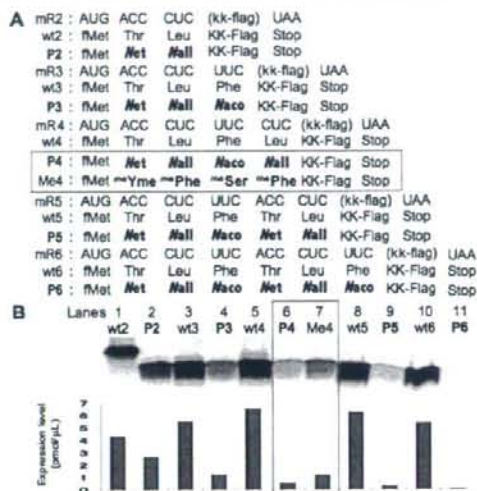


Figure 2. mRNA-programmed synthesis of di-, tri-, tetra-, penta-, and hexapeptoid. (A) Sequences of mRNA templates (mR2–6), control wildtype peptides (wt2–6), a tetra-*N*-methyl-peptide (Me4), and peptoids containing *Net*, *Nall*, and *Naco* (P2–6). ¹⁴C-Yme, *N*-methyl-*L*-*p*-methoxyphenylalanine; ¹⁴C-Phe, *N*-methyl-*L*-phenylalanine; ¹⁴C-Ser, *N*-methyl-*L*-serine. (B) Tricine-SDS-PAGE analysis of control wildtype peptides, a tetra-*N*-methyl-peptide, and peptoids expressed from the respective mRNA. The wildtype peptides were expressed in the ordinary PURE system, while peptoids were expressed in the wPURE system containing *Net*-tRNA^{Asp-E2}, *Nall*-tRNA^{Asp-E2}, and *Naco*-tRNA^{Asp-E2}. Expression level of each peptide labeled with [¹⁴C]-Asp based on its observed radioisotope counts is shown in the graph.

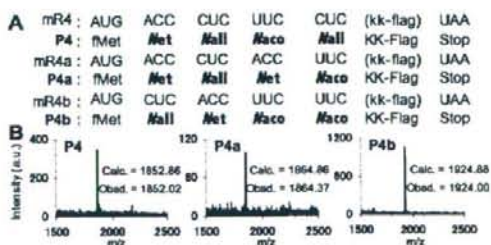


Figure 3. mRNA-programmed synthesis of linear tetrapeptoids (A) Sequences of mRNA templates (mR4, mR4a, and mR4b) and encoded linear peptoids containing *Net*, *Nall*, and *Naco* (P4, P4a, and P4b). The *kk-flag* in parentheses indicates the RNA sequence encoding a KK-Flag peptide (KKDYKDDDDK). The expression level of each tetrapeptoid was estimated by tricine-SDS-PAGE as follows: P4, 0.52 pmol/μL; P4a, 0.37 pmol/μL; P4b, 0.51 pmol/μL. (B) MALDI-TOF-MS spectra of the Flag-purified translation products. The calculated mass (Calc) and observed mass (Obsd) for singly charged species [M+H]⁺ are shown in each spectrum.

records the first demonstration of mRNA-directed expression of polypeptoids using the translation machinery.

Due to the above success, we extended our methodology to the ribosomal synthesis of peptoid-peptide hybrids. With a future application to drug discovery in mind, we attempted to express peptoid-peptide hybrids with a framework in which peptoid bonds are scattered in cyclic peptides closed by a nonreducible thioether bond.^{9b} We reprogrammed the genetic code in which CGC, ACC, CAC, and UCC were assigned to *Nme*, *Net*, *Npr*, and *Nbu* (Figure 1A, 1–4), respectively, for the peptoid synthesis and AUG was

assigned to *N*-chloroacetyl- α -¹Phe or *N*-chloroacetyl- α -⁶Phend (ClAc-¹Phe or ClAc-⁶Phe) for the formation of cyclic scaffolds (Figure S3B). Four mRNA templates (Figure S3A, mR7–10) were designed to express cyclic peptide–peptide hybrids with different sequence compositions attached to the C-terminal KK-Flag (Figure 4A, P7–P10). MALDI-TOF analysis of each cyclic peptide–peptide

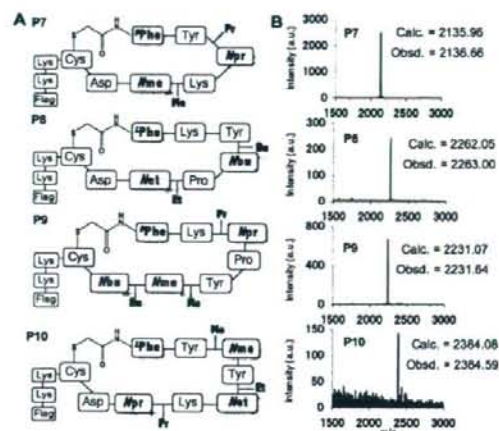


Figure 4. mRNA-programmed synthesis of cyclic peptide–peptide hybrids. (A) Schematic structure of cyclic peptide–peptide hybrids containing *N*me, *N*et, *N*pr, and *N*bu (P7–10). The locations of *N*-substituted groups on the peptide bond are highlighted in gray. The expression level of each peptide–peptide hybrid was estimated by tricine-SDS-PAGE as follows: P7, 1.97 pmol/ μ L; P8, 2.63 pmol/ μ L; P9, 0.21 pmol/ μ L, and P10, 0.84 pmol/ μ L. (B) MALDI-TOF-MS spectra of the Flag-purified translation products. The calculated mass (Calc) and observed mass (Obsd) for singly charged species $[M+H]^+$ are shown in each spectrum.

hybrid expressed from the respective mRNA template exhibited the expected molecular mass as a single major peak (Figure 4B, P7–P10). This firmly demonstrates the mRNA-directed expression of cyclic peptide–peptide hybrids under the reprogrammed genetic code.

In conclusion, we have demonstrated mRNA-directed synthesis of linear polypeptides and cyclic peptide–peptide hybrids. The translation machinery surprisingly accepts a variety of rGly for elongation, including those with alkyl chains and functional groups that are orthogonal^{7d,e,11} to proteinogenic side chains, such as the ketone, alkene, azide, and alkyne. We have shown consecutive elongations of rGly and synthesis of cyclic peptide–peptide hybrids. Particularly, the latter hybrids contain two or more peptide bonds dispersed in the sequence and consist of cyclic structures closed by a physiologically stable thioether bond. Since we have already shown that the thioether cyclization is spontaneously formed and applicable to a variety of ring sizes independent from the sequence compositions of peptide,^{9g,9h,c,12} the methodology reported herein offers us a new means of the mRNA-programmed library synthesis of cyclic peptide–peptide hybrids. This has already directed our current research to running screenings or selections of a new class of peptidic drugs against therapeutic targets by the integration with appropriate techniques.¹³ Such efforts are underway in our laboratory.

Acknowledgment. We thank Dr. P. C. Reid for critical proofread. This work was supported by grants of Japan Society for the promotion of Science Grants-in-Aid for Scientific Research

(S) (16101007) to H.S., Grants-in-Aid for JSPS Fellows to T.K. (20-664), and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Supporting Information Available: Experimental details and additional data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) (a) Simon, R. J.; Kania, R. S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367–9371. (b) Patch, J. A.; Barron, A. E. *Curr. Opin. Chem. Biol.* **2002**, *6*, 872–877. (c) Shin, S. B.; Yoo, B.; Todor, L. J.; Kirshenbaum, K. J. *Am. Chem. Soc.* **2007**, *129*, 3218–3225. (d) Lee, B. C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N. *J. Am. Chem. Soc.* **2008**, *130*, 8847–8855. (e) Kwon, Y. U.; Kodadek, T. *J. Am. Chem. Soc.* **2007**, *129*, 1508–1509.
- (2) (a) Zuckermann, R. N.; et al. *J. Med. Chem.* **1994**, *37*, 2678–2685. (b) Alluri, P. G.; Reddy, M. M.; Baehrawat-Sikder, K.; Olivos, H. J.; Kodadek, T. *J. Am. Chem. Soc.* **2003**, *125*, 13995–14004. (c) Wrenn, S. J.; Weisinger, R. M.; Halpin, D. R.; Harbury, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 13137–13143. (d) Udugamasooriya, D. G.; Dincen, S. P.; Brekken, R. A.; Kodadek, T. *J. Am. Chem. Soc.* **2008**, *130*, 5744–5752.
- (3) (a) Nguyen, J. T.; Turck, C. W.; Cohen, F. E.; Zuckermann, R. N.; Lim, W. A. *Science* **1998**, *282*, 2088–2092. (b) Kruijzer, J. A.; Nijenhuis, W. A.; Wanders, N.; Gispens, W. H.; Liskamp, R. M.; Adan, R. A. *J. Med. Chem.* **2005**, *48*, 4224–4230.
- (4) (a) Bain, J. D.; Wacker, D. A.; Kuo, E. E.; Chamberlin, A. R. *Tetrahedron* **1991**, *47*, 2389–2400. (b) Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197–200. (c) Short, G. F., 3rd; Laikhter, A. L.; Lodder, M.; Shayo, Y.; Arslan, T.; Hecht, S. M. *Biochemistry* **2000**, *39*, 8768–8781. (d) Frankel, A.; Millward, S. W.; Roberts, R. W. *Chem. Biol.* **2003**, *10*, 1043–1050. (e) McCreary, C.; Green, R. *Chem. Biol.* **2004**, *11*, 575–582. (f) Tan, Z.; Forster, A. C.; Blacklow, S. C.; Cornish, V. W. *J. Am. Chem. Soc.* **2004**, *126*, 12752–12753. (g) Tan, Z.; Blacklow, S. C.; Cornish, V. W.; Forster, A. C. *Methods* **2005**, *36*, 279–290. (h) Hartman, M. C.; Josephson, K.; Lin, C. W.; Szostak, J. W. *PLoS ONE* **2007**, *2*, e972. (i) Zhang, B.; Tan, Z.; Dickson, L. G.; Nalam, M. N.; Cornish, V. W.; Forster, A. C. *J. Am. Chem. Soc.* **2007**, *129*, 11316–11317. (j) Kawakami, T.; Murakami, H.; Suga, H. *Chem. Biol.* **2008**, *15*, 32–42. (k) Ohta, A.; Yamagishi, Y.; Suga, H. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159–167. (l) Kang, T. J.; Suga, H. *Biochem. Cell. Biol.* **2008**, *86*, 92–99. (m) Subtelny, A. O.; Hartman, M. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2008**, *130*, 6131–6136. (n) Sando, S.; Masu, H.; Furutani, C.; Aoyama, Y. *Org. Biomol. Chem.* **2008**, *6*, 2666–2668.
- (5) (a) Forster, A. C.; Weissbach, H.; Blacklow, S. C. *Anal. Biochem.* **2001**, *297*, 60–70. (b) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001**, *19*, 751–5. (c) Forster, A. C.; Tan, Z.; Nalam, M. N.; Lin, H. Q.; Hsueh, C.; Cornish, V. W.; Blacklow, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6353–6357. (d) Josephson, K.; Hartman, M. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 11727–11735. (e) Shimizu, Y.; Kanamori, T.; Ueda, T. *Methods* **2005**, *36*, 299–304. (f) Shimizu, Y.; Kuruma, Y.; Ying, B. W.; Umekage, S.; Ueda, T. *FEBS J.* **2006**, *273*, 4133–4140.
- (6) (a) Xiao, H.; Murakami, H.; Suga, H.; Ferre-D'Amare, A. R. *Nature* **2008**, *454*, 358–361. (b) Ohuchi, M.; Murakami, H.; Suga, H. *Curr. Opin. Chem. Biol.* **2007**, *11*, 537–542. (c) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357–359.
- (7) (a) Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. *Chem. Biol.* **2007**, *14*, 1315–1322. (b) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 120–129. (c) Sako, Y.; Morimoto, J.; Murakami, H.; Suga, H. *J. Am. Chem. Soc.* **2008**, *130*, 7232–7234. (d) Goto, Y.; Murakami, H.; Suga, H. *RNA* **2008**, *14*, 1390–1398.
- (8) LaRiviere, F. J.; Wolfson, A. D.; Uhlenbeck, O. C. *Science* **2001**, *294*, 165–168.
- (9) (a) Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 5652–5653. (b) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 11782–11783. (c) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164–11165. (d) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56–61. (e) Chang, P. V.; Prescher, J. A.; Hangauer, M. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2007**, *129*, 8400–8401. (f) Tam, A.; Soellner, M. B.; Raines, R. T. *J. Am. Chem. Soc.* **2007**, *129*, 11421–11430. (g) Slavoff, S. A.; Chen, I.; Choi, Y. A.; Ting, A. Y. *J. Am. Chem. Soc.* **2008**, *130*, 1160–1162.
- (10) Sako, Y.; Goto, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 241–249.
- (11) (a) Hanes, J.; Pluckthun, A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4937–4942. (b) He, M.; Taussig, M. J. *Nucleic Acids Res.* **1997**, *25*, 5132–5134. (c) T. Nemoto, N.; Miyamoto-Sato, E.; Humisi, Y.; Yanagawa, H. *FEBS Lett.* **1997**, *414*, 405–408. (d) Roberts, R. W.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12297–12302. (e) Mastrobattista, E.; Taly, V.; Chamudet, E.; Treacy, P.; Kelly, B. T.; Griffiths, A. D. *Chem. Biol.* **2005**, *12*, 1291–1300. (f) Wrenn, S. J.; Harbury, P. B. *Annu. Rev. Biochem.* **2007**, *76*, 331–349.

JA806998V

Expression of Histone H3 Tails with Combinatorial Lysine Modifications under the Reprogrammed Genetic Code for the Investigation on Epigenetic Markers

Taek Jin Kang,^{1,3} Satoshi Yuzawa,^{2,3} and Hiroaki Suga^{1,2,*}

¹Research Center for Advanced Science and Technology, The University of Tokyo, 153-8904 Tokyo, Japan

²Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 113-8656 Tokyo, Japan

³These authors contributed equally to this work.

*Correspondence: hsuga@rcast.u-tokyo.ac.jp

DOI 10.1016/j.chembiol.2008.09.014

SUMMARY

We report the ribosomal synthesis of N-terminal peptides of histone H3, so-called H3 tail (H3t), with combinatorial methyl and acetyl modifications of selected lysine residues, and the application of such peptides to studying the influence of lysine modification on H3t binding to chromodomain of heterochromatin protein 1 (chromoHP1). Genetic code reprogramming was employed to reassign four codons to acetylated, mono-, di-, and trimethylated lysines, and 38-mer H3t peptides containing modified lysines at designated sites were expressed from the corresponding mRNA sequences. Using a series of H3t constructs, we show complex crosstalk among methylated lysine 9 and 27, and acetylated lysine 14 for binding to chromoHP1. This proof-of-concept study offers a unique means for the synthesis of not only an H3t library containing modified lysines but also other classes of peptides bearing posttranslational methylation and acetylation.

INTRODUCTION

A nucleosome consists of 146 base pairs of DNA wrapped around an octamer of core histone proteins, forming the basic repeating unit of chromatin. Each of the four core histones, H2A, H2B, H3, and H4, contains a structural domain consisting of three α helices and an unstructured domain at the N terminus or C terminus, referred to as the histone tail. Enzyme-catalyzed chemical modification of certain residues in the histone tails serves to regulate the local structure of chromatin, altering the accessibility of various proteins to histone tails as well as to DNA. Thus, histone tail modification represents one of the important epigenetic marks linked to various cellular processes, such as replication, transcription, and DNA repair (Kouzarides, 2007).

The histone tails of core histones, especially that of H3 (H3t), consists of about 40 amino acid residues and are subjected to various chemical modifications such as acetylation, methylation, phosphorylation, and possibly ubiquitylation (Kouzarides, 2007; Peterson and Lanier, 2004). Lysines (Ks) in H3t can be acetylated (ac; acetylated lysine, acK) or methylated (me), and the methylation can be one of mono-, di-, or trimethylation (me1, me2, and me3, respectively; Figures 1A and 1B). In general, hyper- and

hypoacetylation of H3t are linked with transcriptional activation and repression, respectively. Methylation can influence transcription in various ways, depending on the site of the modification (Kouzarides, 2007; Martin and Zhang, 2005). Linking a particular modification or set of modifications to a particular cellular function involving effector proteins postulated a "histone code" hypothesis (Jenuwein and Allis, 2001; Strahl and Allis, 2000). Thus, modifications at H3t can be viewed as special marks modulating the interaction among histones or between the histone and nonhistone proteins. Certain modifications at H3t can also affect the consequences of others, and evidence of such crosstalk between modifications has been accumulating in the past few years (Fischle et al., 2005; Kouzarides, 2007; Latham and Dent, 2007; Seet et al., 2006). Moreover, recent findings of the combinatorial linkages among modifications in H3t also suggested possible long-range crosstalk among modifications of H3t (Garcia et al., 2007; Taverna et al., 2007). However, our current understandings of their interplay at the molecular level are limited, partly because there is no reliable and readily accessible method to prepare the full-length H3t with multiple modifications in a defined and controlled manner.

Genetic code reprogramming is a technique where codons usually assigned to proteinogenic amino acids are reassigned to nonproteinogenic amino acids, including those with unique side chains (Forster et al., 2003; Josephson et al., 2005; Murakami et al., 2006). This enables the expression of peptides containing multiple nonproteinogenic amino acids using a translation system. We envisioned that this technique is suitable for the synthesis of H3t with sparsely positioned multiple modifications. We here report the mRNA-directed synthesis of 38-mer H3t containing Ks modified with *N*-mono-, di-, trimethyl and/or acetyl group(s) and its application to studies on the specific interaction between chromoHP1 and H3t modified at positions 9, 14, and 27. This proof-of-concept study offers a unique means for the synthesis of an H3t library containing modified Ks to investigate the role of combinatorial modifications in histone tail structure and function.

RESULTS

Translatability of Methylated or Acetylated Lysine upon Genetic Code Reprogramming

The 38-mer H3t is composed of 10 members of the proteinogenic amino acids (alanine, arginine, glutamine, glycine, leucine,

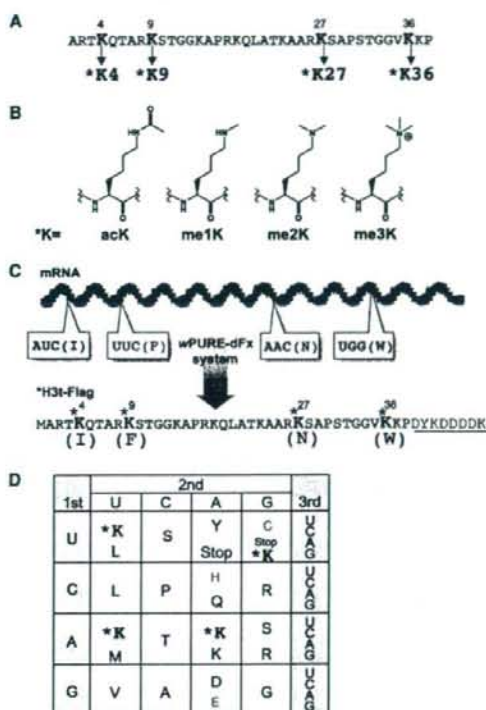


Figure 1. Genetic Code Reprogramming for the Preparation of *H3t
(A) The primary sequence of H3t used in this study. *Ks were incorporated into positions indicated.

(B) *Ks used in this study. Note that both me1K and me2K are protonated in physiological conditions, and thus all the meKs are positively charged.

(C) An example of codon assignments for the genetic code reprogramming. Codons for K4, K9, K27, and K36 were changed to those for I, F, N, and W, respectively. Color codes used here represent the codons in use throughout the paper.

(D) A representative reprogrammed genetic code table. Four reprogrammed codons are in colors, whereas three unused codon boxes, shown in gray, were left vacant.

lysine, proline, serine, threonine, and valine), including eight Ks. Among these Ks, four Ks at positions 4, 9, 27, and 36 (K4, K9, K27, and K36, respectively) were often found methylated *in vivo*; therefore, we were interested in synthesizing H3t peptides containing Ks specifically modified at these positions (abbreviated to *Ks; Figure 1A). In addition to the these 10 amino acids, methionine, aspartate, and tyrosine (M, D, Y, respectively) were included in the peptide chain to assign the initiation (M) and Flag peptide tag (DYKDDDDK; D could be [¹⁴C]-D) at the C terminus for purification, immobilization, and radiolabeling of the expressed H3t peptides. Consequently, seven amino acids were yet unused, and four of these codons were reassigned to *Ks modified with me1, me2, me3, or ac on the ε-amino group

(Figure 1B). We chose codons of AUC (isoleucine, I), UUC (phenylalanine, F), AAC (asparagine, N), and UGG (tryptophan, W) and reassigned these codons to code for *Ks (Figures 1C and 1D). Thus, these amino acids as well as other unused amino acids were withdrawn from the PURE (protein synthesis using recombinant elements) system to generate a wPURE system that contained all the aminoacyl-tRNA synthetases but only the essential 13 amino acids, for genetic code reprogramming. Derivatization of *Ks to 3,5-dinitrobenzyl esters (DBE) made these amino acids compatible with a dinitro-flexizyme, a ribozyme-based tRNA acylating catalyst (dFx; see Murakami et al., 2006). dFx was able to aminoacylate tRNA^{Asn-E2}_{GAU} with the respective modified *K-DBE (see Figure S1 available online). Because dFx is inherently able to recognize tRNAs bearing any anticodon, this finding ensured that any pair of *K and tRNA^{Asn-E2}_{NNN} (NNN denotes an anticodon complementary to the reprogrammed codon) could be prepared by the dFx technology.

Though it was known that acK could be incorporated into a nascent peptide chain efficiently by translation under the reprogrammed genetic code (Murakami et al., 2006), no information was available for the incorporation efficiency of meKs. To verify the translatability of the individual *Ks at each position, we designed a template containing codons for I, F, N, and W at positions 4, 9, 27, and 36, respectively (Figure 1C). Prior to the suppression experiments, two control experiments were performed: as a positive control, the corresponding DNA template was transcribed and translated in the wPURE system that was supplemented with I, F, N, and W. An intense band of a [¹⁴C]-labeled peptide was observed on SDS-tricine-PAGE (Figure 2A, lane 1), and MALDI-TOF analysis of the Flag-purified peptide revealed a molecular weight (MW) consistent with that of the expected peptide containing I, F, N, and W at the designated positions instead of Ks (Figure 2B, P-WT). As a negative control to verify the background expression at position 4 (I) without *K4 suppression, the same template was translated in a wPURE system supplemented with F, N, and W only. A faint band with nearly the same mobility as the P-WT was observed (Figure 2A, lane 2), and the MALDI-TOF analysis of the Flag-purified peptide suggested the misincorporation of other amino acids (Figure S2). However, this background expression could be readily competed out by suppression using *K-tRNA^{Asn-E2}_{GAU}; upon the addition of the respective *K-tRNA^{Asn-E2}_{GAU} to the wPURE system supplemented with F, N, and W, an intense band appeared irrespective of *K used (lanes 3–6), and the MALDI-TOF analysis of the respective Flag-purified peptide confirmed a peak of the main product with the expected MW containing the designated *K (Figure 2B, P-K4me1, P-K4me2, P-K4me3, and P-K4ac). Likewise, the background peptide expressions at the other three positions, K9, K27, and K36 (lanes 8, 14, and 20), were also suppressed by the addition of the respective *K-tRNA^{Asn-E2}_{NNN} (NNN denotes the corresponding anticodon) to afford an intense band in the tricine-SDS PAGE (Figure 2A, lanes 9–12, 15–18, and 21–24), and the MW of the respective peptides were also confirmed (Figure 2B, P-KXme1, P-KXme2, P-KXme3, and P-KXac; X = 9, 27, or 36). These findings showed all *Ks were compatible with the ribosomal peptide synthesis upon genetic code reprogramming, yielding the designated peptide as a main product.

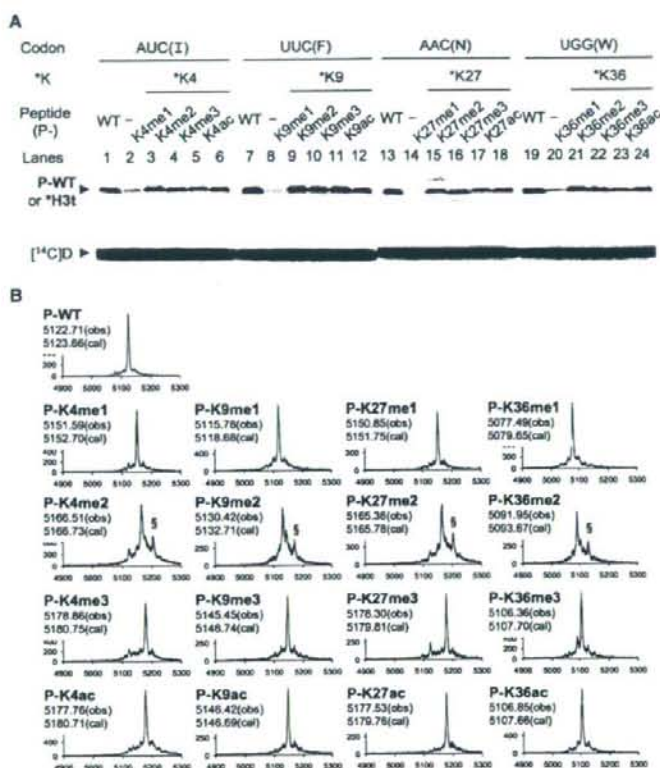


Figure 2. Single Incorporations of *Ks into H3t

(A) Analysis of expressed peptides containing a single *K substitution. Incorporation efficiencies were examined on the 15% SDS-tricine-PAGE using the wild-type peptide that has I4, F9, N27, and W36 as a control. Because the wPURE system used in this study did not contain any of these amino acids, a proper combination of amino acids were exogenously added to the translation; for example, F, N, and W were added to the translation for I codon suppressions. Unincorporated [¹⁴C]-D is shown at the bottom of the gel.

(B) MALDI-TOF analysis of each *H3t peptide after anti-Flag antibody affinity purification. § indicates a peak of the potassium adduct. Note that an intense peak corresponding to the potassium adduct was present in the series of me2K-containing *H3t.

The immobilized molecules were acid eluted from the resin and analyzed by immunoblot using anti-Flag and anti-polyhistidine polyclonal antibodies (Figure 3B). The *H3t peptides were detected only with the addition of charged tRNA^{Asn-E2}_{GAA} (Figure 3B, lanes 2–6 versus lane 1), confirming that the *H3t peptide expression was solely the consequence of genetic code reprogramming. Significantly, chromoHP1 was detected only when *H3t containing meK9 was used as bait (Figure 3B, lanes 2 and 3 versus lanes 4–6; see also Figure S3 for product verification by mass analysis). We clearly show that *H3tK9me3 was

Interaction of HP1 with the Translated H3t Peptides Containing a Methylated Lysine at the Ninth Position

The next critical question was whether the translated H3t peptides containing the modified Ks (*H3t peptides) were able to exhibit the expected modification-specific interactions with effector proteins. To confirm this, we chose heterochromatin protein 1 (HP1), of which the chromodomain is known to interact with K9-modified *H3t peptides in a methylation-dependent manner (Jacobs et al., 2001; Lachner et al., 2001). Two well-established analytical methods for detecting protein-protein interactions, pull-down analysis and surface plasmon resonance (SPR), were employed to show the meK9-specific recruitment of the chromodomain of HP1 (chromoHP1) by the *H3t peptide.

A template sequence that contained the UUC(F) codon at position 9 was translated to Flag-tagged H3t, *H3tK9ac, *H3tK9me1, *H3tK9me2, and *H3tK9me3 in the wPURE system, in the presence of the corresponding *K-tRNA^{Asn-E2}_{GAA} (Figure 3A). As a negative control, the same template was translated in the absence of aminoacyl-tRNA^{Asn-E2}_{GAA} in the same wPURE system. *H3t peptides were displayed on an agarose resin through covalently immobilized anti-Flag antibody, and the resulting resin was used directly as bait to pull down chromoHP1.

able to pull down chromoHP1 the most effectively among the *H3t peptides. Our findings are in good agreement with an earlier report in which the affinity of 15-mer synthetic H3t peptides (Ala1–Ala15) containing K9me1, k9me2, and k9me3 were in the order of K9me1 < K9me2 < K9me3 (Fischle et al., 2003).

To quantitatively assess the affinity of *H3t peptides, we next performed SPR where the expressed *H3t peptides were immobilized on a sensor chip that was covalently coupled with an anti-Flag antibody. When 2 μM of chromoHP1 was flowed over the respective *H3t peptides on the chip, we reproducibly observed that the *H3tK9me3 exhibited a resonance unit at equilibrium (ΔReq) value 5–20 times higher than other *H3t peptides, indicating its strong affinity to chromoHP1 (Figure 3C). We then varied the concentrations of chromoHP1 using *H3tK9me3 as a binding partner to obtain the dissociation constant of 0.76 ± 0.05 μM (Figure 3C, inset). This value was slightly lower but similar to the reported values 2.4–4.0 μM using the aforementioned short H3t peptide with the K9me3 modification. Collectively, these findings show that the *H3t peptides prepared by genetic code reprogramming can be used to detect interactions with a known effector protein qualitatively and quantitatively.

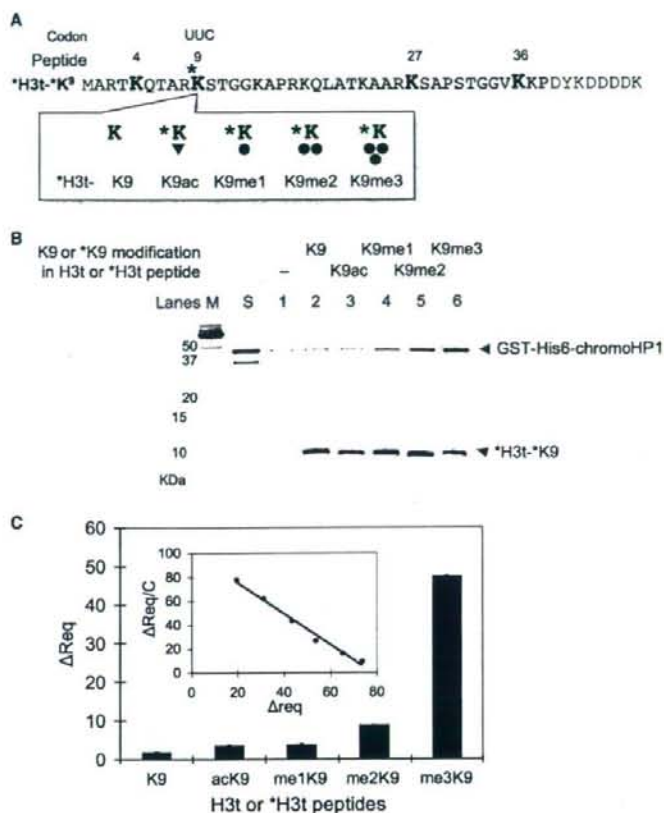


Figure 3. Modification-Specific Interaction of *H3t with the Chromodomain of HP1

(A) *H3t peptides used for the pull-down or the sensor chip for SPR analysis. Each peptide contains *K at position 9.

(B) Pull-down analysis of *H3t interaction with chromoHP1. Pulled-down prey (chromoHP1) and the baits (*H3t peptides) were probed using anti-polyhistidine and anti-Flag antibodies, respectively. M, protein size marker; S, chromoHP1 (2% of input); 1, pull-down with the bait prepared from the translation in the absence of both F and aminocyl-tRNA^{Asn-E2}_{NNN}; 2-6, pull-down with *H3t-K9, -K9ac, -K9me1, -K9me2, and -K9me3, respectively.

(C) SPR analysis of *H3t interaction with chromoHP1. Respective peptides were immobilized on the sensor chip through preimmobilized anti-Flag tag antibody. Resonance units in the equilibrium with 2 μM of chromoHP1 were measured and normalized against the amount of *H3t bound to the chip. Mean values of triplicate measurements of two independent sets of experiments and the SD are shown. A scatchard plot of *H3tK9me3-chromoHP1 with varying concentrations of chromoHP1 is shown in the inset.

MALDI-TOF analysis of each Flag-purified peptide showed that the observed major peak was consistent with the expected MW for the multiple suppressions (Figure 4A, P-D, P-T, and P-Q).

To perform the ribosomal synthesis of *H3t peptides, two approaches are applicable. One is to prepare different *H3t peptides from the same template using reprogrammed genetic tables with customized *K assignments for each *H3t peptide (Figure 4B). This is possible because the dFx technology enables us to readily prepare any desired *K-tRNA^{Asn-E2}_{NNN} to

create a new reprogrammed genetic table. The other approach is to use a unique genetic table (e.g., as shown in Figure 1D), assigning each respective *K and then expressing the desired *H3t peptides from different mRNA sequences. In this method, swapping the positions of reassigned codons to *K in mRNA dictates the positions of *K incorporation in each *H3t peptide (Figure 4C). To show the versatility of our methodology, we employed both approaches to synthesize *H3t peptides.

In the first approach, each reprogrammed codon was assigned to K, meK, me2K, me3K, or acK by charging the respective *K onto tRNA^{Asn-E2}_{NNN} (Figure 4B). In this series of experiments, the incorporation of K at preselected positions was also performed by the suppression using K-tRNA^{Asn-E2}_{NNN} like other *K incorporations (as indicated by the color-coded Ks in Figure 4B). Upon expression of each peptide by the respective reprogrammed genetic table, we confirmed the product by MALDI-TOF analysis (see *H3t-II, *H3t-III, and *H3t-IV). The observed MW of the major peak was consistent with the expected values for the respective peptides in all cases, indicating that the desired *H3t peptides were synthesized.

Synthesis of H3t Peptides by Multiple Incorporations of Modified Lysines

We next directed our investigation to multiple incorporations of *Ks into H3t. Using the template shown in Figure 1C, codons at positions 4, 9, 27, and 36, coding for I, F, N, and W, respectively, were reprogrammed to assign me1K, me3K, acK, and me2K, respectively. We then suppressed the reprogrammed codons using the designated *K-tRNA^{Asn-E2}_{NNN} in the wPURE system with proper supplementation of the cognate amino acids. The template sequence was expressed in the wPURE system supplemented with all amino acids to yield P-WT as a positive control, and the same template was translated in the wPURE system to show no detectable background expression (Figure 4A, lanes 1 and 2). When the suppressions were performed using the corresponding *K-tRNA^{Asn-E2}_{NNN}, a clear band appeared in each experiment (Figure 4A, lanes 3-6). Increasing the number of suppressions resulted in a concomitant decrease in the overall expression level, yet even the quadruple suppression was achieved at an expression level of 37% to that of P-WT (Figure 4A, lane 6, P-Q versus lane 1, P-WT). Most importantly,

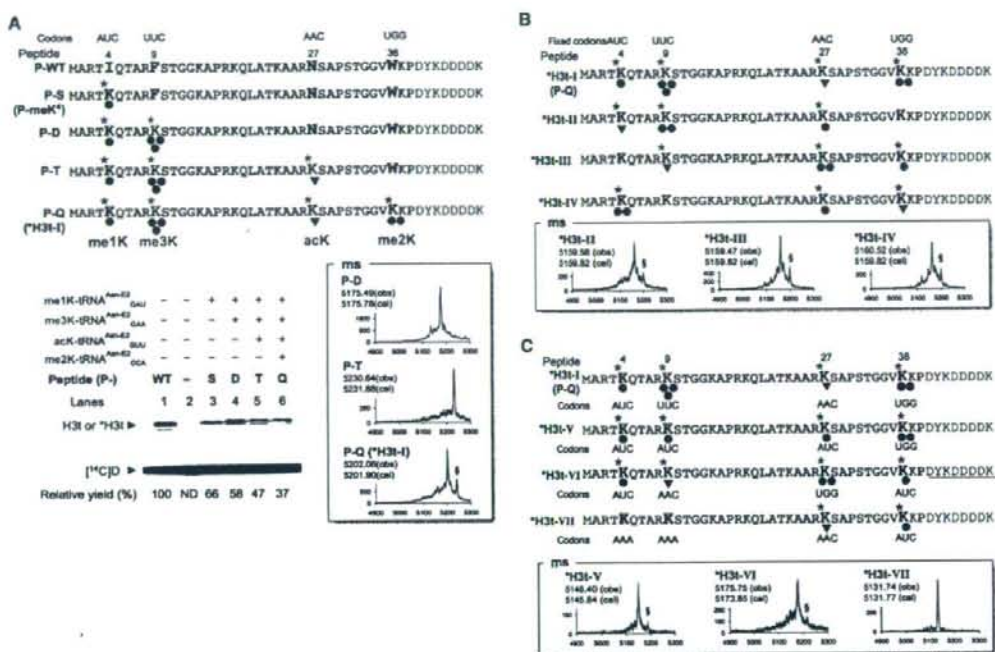


Figure 4. Multiple Incorporations of *Ks into H3t

(A) Demonstration of multiple incorporations of *Ks into H3t. Single to quadruple modifications can be coded in mRNA using the reassignment scheme shown in Figure 1C. Multiple incorporations were assessed quantitatively (left) and qualitatively (right) using 15% tricine-SDS-PAGE and MALDI-TOF, respectively. (B) Multiple incorporations of *Ks into H3t peptides using codon reassignments using the single mRNA sequence. (C) Multiple incorporations of *Ks into H3t peptides using codon arrangements in mRNA. ‡ indicates a peak of the potassium adduct of me2K-containing *H3t peptides. The triangle and circle denote acetylation and methylation, respectively.

In the second approach, we prepared three mRNA templates encoding *H3t peptides (Figure 4C). Because the reprogrammed genetic codon table was fixed in this series of experiments (Figure 1D), each mRNA sequence dictated the identity and the position of *Ks incorporated, giving three different *H3t peptides with different MW. Unmodified K was assigned by the cognate AAA codon in mRNA (as indicated by black K in Figure 4C). The respective translated peptides were analyzed by MALDI-TOF, showing that the major peak of each peptide was consistent with the expected MW. Thus, this approach is also reliable for the synthesis of *H3t peptides with various combinations of K and *Ks at any designated positions.

Interplay among Lysine Modifications at Positions 9, 14, and 27 of H3t Peptides in Its Binding to chromoHP1

HP1-histone H3 binding is mostly attributed to H3K9me, and, in many cases, the modification can be connected to transcriptional silencing or heterochromatin formation. However, it has also been shown that H3K27me3 can recruit HP1 to H3t in vitro with a reduced affinity compared with that of an H3K9me3 peptide (Fischle et al., 2003), even though the neighboring amino

acids of K9 and K27 are similar (Figure 5A). However, acetylation at K14 has been generally conceived to be a transcription-facilitating epigenetic mark as opposed to K9me. Mass analyses of in vivo modifications of histone H3 revealed that these two epigenetically opposite modifications coexist in the same molecule of histone H3 (Bonenfant et al., 2007; Garcia et al., 2007; Thomas et al., 2006). Furthermore, a significant portion of histone H3 was shown to contain K9meK27me double modification or K9meK14acK27me triple modification (Garcia et al., 2007). In terms of the role of these modifications on the interaction between *H3t and HP1, it has been reported using synthetic *H3t peptides lacking the K27 modification that *H3tK9me3K14ac did not show an appreciable difference from *H3tK9me3 in qualitative pull-down experiments (Hirota et al., 2005; Mateescu et al., 2004). However, there might be interplay between the modifications of K9 and K14 along with K27 in the HP1-H3t interaction, which had not been investigated. To address the above question, we performed SPR experiments where *H3t peptides containing combinatorial modifications at the above K residues were tested for binding to chromoHP1.

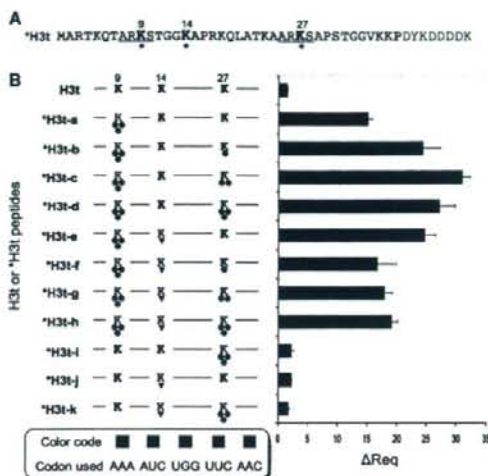


Figure 5. Interplay among Lysine Modifications at Positions 9, 14, and 27 in chromoHP1-H3t Binding

(A) The sequence of H3t peptides used. Positions 9, 14, and 27 contained modified Ks in various combinations. The consensus sequences in the *K9 and *K27 regions are underlined.

(B) SPR analysis of chromoHP1 interaction with H3t peptides modified at positions 9, 14, and 27 in combinatorial manners. Modifications are summarized in the left panel. *Ks were incorporated into the peptide sequence using the codon assignments shown in color codes in the bottom left panel. The right panel shows resonance units observed in the equilibrium with 0.25 μ M of chromoHP1 normalized against the amount of H3t bound to the chip. Mean values and SD are shown in each bar, which were generated from triplicate measurements of two independent sets of experiment (*H3t-a~h) or a single set of experiment (H3t and *H3t-i~k). The triangle and circle denote acetylation and methylation, respectively.

We prepared a series of H3t peptides with a total of 10 combinations of the methyl and acetyl modifications at positions 9, 14, and 27 (Figure 5A, *H3t-b~k; see also Figure S4 for the product verification by mass analyses) using the reprogrammed codons shown in Figure 1C (the summary of the reprogrammed codons used was also shown in the bottom panel of Figure 5B). The concentration of chromoHP1 in the mobile phase was kept at 0.25 μ M to prevent saturation binding, where the amount of bound chromoHP1 (Δ Req) was expected to be proportional to the affinity of respective peptides. This let us assess the effect of modifications on chromoHP1-H3t binding semiquantitatively. As shown earlier in Figure 3, K9me3 modification promoted the chromoHP1-H3t binding (H3t versus *H3t-a, Figure 5B). K9me3K27me double modification resulted in elevated binding over K9me3 modification (*H3t-b~d versus *H3t-a), though K27me3 alone (*H3t-i) did not show an appreciable positive effect on chromoHP1 binding compared with H3t. However, the K14ac modification on *H3tK9me3 (*H3t-a \rightarrow *H3t-e) increased the binding to a similar level observed in *H3tK9me3K27me1 (*H3t-e versus *H3t-b), whereas the same modification on *H3tK9me3K27me1 (*H3t-b \rightarrow *H3t-f) suppressed chromoHP1 binding down to the level of *H3tK9me3 (*H3t-f versus *H3t-a).

Once K14ac modification was placed in *H3tK9me3-K14acK27me, further methylation(s) on K27 did not affect the affinity (*H3t-f \rightarrow *H3t-g or *H3t-h). Finally, neither K14ac modification alone on H3t (*H3t-j) nor K27me3 modification on *H3tK27me3 (*H3t-k) showed a significant elevation in chromoHP1 affinity from that of H3t.

DISCUSSION

In this report, we show the preparation of H3t with multiple modifications using genetic code reprogramming and the functional use of these peptides in elucidating histone modification-effector binding relationships. Because of the biological significance of histone modifications and our insufficient knowledge of their roles in epigenetic control, it has been of great interest to develop methods to prepare histones or histone tails containing *Ks. One such method is to use enzymes that are known to modify histones (Kouzarides, 2007). This method allows for the preparation of not only modified histone tails but also the full-length histones, which represents the most significant advantage of this method. However, many of such enzymes would modify multiple sites in the target histones with incomplete modifications, leaving certain heterogeneities in the modified histones (Winter et al., 2008). This may impose difficulties in investigating specific interactions of histones or histone tails with effector proteins in certain cases. An elegant alternative approach was recently reported where a cysteine residue was introduced to H3 and then chemically modified with 2-halo-ethylamine derivatives to install a *K analog (Simon et al., 2007). The virtue of this method is its simplicity for the synthesis of the full-length H3 analogs, whereas an inherent limitation is that only a single site modification or multiple but homogeneous *K modifications (albeit this has yet to be demonstrated) is possible; therefore it is not applicable to addressing questions regarding the importance of combinatorial modifications at the arbitrary K residues.

Chemically synthesized 20- to 30-amino-acid-long peptides containing *K at a certain specific sites has been widely and successfully used to elucidate the molecular interactions taking place on histones (Fischle et al., 2003; Jacobs et al., 2001; Lachner et al., 2001; Matthews et al., 2007; Meehan et al., 2003). Such short peptides can also be ligated to the C-terminal fragment of H3 or H4 to construct the full-length histones (Fischle et al., 2003; Shogren-Knaak et al., 2003, 2006). However, to elucidate the role of sparsely located modifications, the length of greater than 40 amino acids that span the tail would be preferable. This task is not easy to achieve using chemical synthesis, particularly if multiple *Ks with various kinds were necessary to be placed in the sequence.

Comparatively, ribosomal synthesis can be a practical and reliable method for the preparation of relatively long peptides with versatility. However, for the preparation of H3t with multiple modifications, two prerequisites must be met: (i) the ribosome should use *Ks and (ii) the positions of these amino acids should be coded for in the mRNA precisely. Because all *Ks, including the bulkiest me3K, were incorporated into the peptide efficiently (Figure 2), the first prerequisite was easily met. The second prerequisite was solved by the use of genetic code reprogramming combined with the wPURE and flexizyme systems. This is, to the best of our knowledge, the first report concerning the ribosomal incorporation of *Ks with various degrees of methyl

modifications into a peptide chain. Combining these two findings, we could prepare a variety of ^3H 3t peptides with four kinds of modifications at four different positions 4, 9, 27, and 36 in the desired combinations (Figure 4). The 47-mer Flag-tagged ^3H 3t peptides with sparsely located such modifications were easily prepared using our approach.

Taking advantage of the fact that any kind of peptide tag can be appended to ^3H 3t by expressing mRNA that encodes such a tag at the N or C terminus, the crude ^3H 3t-Flag peptides from the translation can be directly immobilized on an anti-Flag antibody-coated resin or sensor chip; conveniently, the immobilized peptides can be used to confirm K9 modification-specific binding of ^3H 3t to chromoHP1 (Figure 3). The quantitative analysis of the interaction between ^3H 3tK9me3 and chromoHP1 using SPR showed tight binding with a dissociation constant of $0.76 \mu\text{M}$. This value is slightly lower than a previously reported value ($2.5\text{--}4.0 \mu\text{M}$) using a short synthetic ^3H 3tK9me3 peptide consisting of A1–A15 (Fischle et al., 2003; Jacobs and Khorasanizadeh, 2002). Although the exact reason for a decrease in the dissociation constant of our ^3H 3tK9me3 is unclear, it is possible that the full-length ^3H 3tK9me3 might act differently from the shorter peptide for the binding to chromoHP1. Nonetheless, the resin-bound ^3H 3t could be used to pull down chromoHP1 (Figure 3), demonstrating the feasibility of our method for an application that allows us to pull down various and possibly novel effector proteins from cell lysates.

Finally, we have shown the use of our methodology in unveiling the possible interplay among the modifications at K9, K14, and K27. Resonance unit at equilibrium (ΔReq) of ^3H 3t-chromoHP1 binding increased by an approximately 2-fold when the second modification of K27me was added to ^3H 3tK9me3 (Figure 5, ^3H 3t-a \rightarrow ^3H 3t-c). K27me3 modification alone on H3t did not positively influence binding under the conditions used here (Figure 5, ^3H 3t-i). Despite that the neighboring residues of K9me3 and K27me3 are similar (see underlines in Figure 5A), the X-ray structure of the complex of ^3H 3tK9me3 with chromoHP1 has revealed that T6 in the K9me3 region plays a critical role in fortifying their interaction, resulting in lower affinity of chromoHP1 to K27me3 (Fischle et al., 2003). However, a 2-fold enhancement in binding observed for ^3H 3tK9me3K27me2 (Figure 5, ^3H 3t-c) suggests that methylation(s) of K27 may contribute to the binding once chromoHP1 is recruited to the site of K9me3. Moreover, K14ac modification in ^3H 3tK9me3K27me3 cancels the observed enhancement in affinity (Figure 5, ^3H 3t-d \rightarrow ^3H 3t-h), exhibiting the binding ability similar to ^3H 3tK9me3 (Figure 5, ^3H 3t-a). These observations allow us to propose a model where methylation(s) on K27 retards the dissociation of chromoHP1 from the peptide by "shuttling" chromoHP1 between the K9me3 region and the K27me3 region of H3; when K14ac modification is added, the shuttling is disrupted, presumably by a conformational change of the peptide, neutralizing the enhancement by K27me3. However, this shuttling model fails to explain another observation made for the K14ac modification on ^3H 3tK9me3 that enhances binding to chromoHP1 (Figure 5, ^3H 3t-a \rightarrow ^3H 3t-e). Because ^3H 3tK14ac does not bind to chromoHP1 (Figure 5, ^3H 3t-j), this effect is clearly cooperative with the K9me3 modification. We therefore propose an alternative model that K14ac or K27me3 modification alone on ^3H 3tK9me3 "locks" the peptide conformation with a favorable state for binding to chromoHP1. However, the simultaneous modifications in both sites unlock the favorable

conformation, thereby canceling the enhancement. More biochemical and structural studies are required to reveal the molecular mechanism of the above two models in the future.

Regardless of the exact mechanism, in human cells contradictory epigenetic marks of K9me3 and K14ac do coexist in the same molecule of histone H3, and this double modification occasionally accompanies the third modification of K27me (Garcia et al., 2007). Also, K14ac was shown to be necessary together with phosphorylation (ph) at S10 to eject HP1 from histone H3K9me, during the G2/M phase of the human cell cycle (Mateescu et al., 2004); in other cases, S10ph alone was enough to break that binding (Fischle et al., 2005; Hirota et al., 2005). Thus, it can be speculated at this point that the effect of K14ac on HP1 binding to histone H3 would be K27me dependent and specific for the cell cycle and/or the position on the chromatin. Even though the crosstalk among K9me, S10ph, and K14ac modifications in histone H3 have been investigated using short peptides, our work suggests that the crosstalk of these modifications in the context with the K27 modification would be worthy to reinvestigate in a more extensive manner.

SIGNIFICANCE

Posttranslational methylation of lysines in histones poses significant impacts on several cellular functions, thus constituting one of important epigenetic marks. The effect of lysine methylation is dependent on the position and the degree (mono-, di-, and trimethylation) of modification, and can be augmented or nullified by (an)other modifications in the same and/or neighboring histone (Kouzarides, 2007). Because over 150 different combinations of lysine methylation and acetylation are known in histone H3 (Garcia et al., 2007), and such modifications are mostly found in the unstructured N terminus of histone H3 (H3t), it is required to prepare peptides with posttranslational modifications of which positions span the entire region of H3t. We here showed that the ribosome could use methylated lysine (KXme, X denotes the number of position) efficiently (regardless of its methylation state) as well as acetylated lysine (KXac). Upon genetic code reprogramming, we could prepare H3t peptides with lysine methylation and acetylation at designated sites in a combinatorial manner. Using this method, we also unveiled possible crosstalk among K9me, K14ac, and K27me of H3t upon binding to chromodomain of heterochromatin protein 1; K9me3 had a positive effect on binding, K27me augmented the positive effect, and the third modification, K14ac, nullified the augmenting effect of K27me only. Similarly, K14ac augmented K9me3-induced binding, and K27me nullified the effect of K14ac on binding. Being rapid and simple, this method would help to decipher the effect of combinatorial modifications in histones, and it can be extended to the investigations of other classes of peptides, such as p53, which contains posttranslational modifications, including lysine methylation (Chuilikov et al., 2004).

EXPERIMENTAL PROCEDURES

Preparation of the Template DNA

A DNA sequence of the human H3.1 was referred to for the PCR preparation of the H3t peptide template with the addition of a C-terminal Flag tag. Briefly, K9F

(5'-CAGAC TGCCG GCTTC TCGAC CGGTG GTAAA GCA-3') was annealed to Linker (5'-GCGAG CGGCT TTTGT AGCCA GTTGC TTCTC GGGTG CTTTA CC ACC GG-3') and extended by Taq DNA polymerase. The extension mixture was diluted 20 times into the PCR mixture and amplified using K41 (5'-AAGAA GGAGA TATAC ATATG GCTCG TACAA TCCAG ACTGC CGGC-3') and K27N (5'-CACCC CTCCA GTAGA GGGCG CACTG TTGCG AGCGG CTTTT GT-3') as the 5' and 3' primer respectively. Similarly, the product was further extended twice using 5'T7 (5'-GTAAT ACGAC TCACT ATAGG GTTTA ACTTT AAGAA GG AGA TATAC AT-3') and K36W (5'-CTTGT CGTCA TCGTC TTTGT AGTCA GGT TT CCACA CCCCT CGAGT AGA-3') as the first primer pair, and 5'T7 and 3'FLAG (5'-CGAAG CTTAC TTGTC GTCAT CGTCT TTGTA-3') as the second primer pair.

Syntheses of *K-DBEs

N^ε-Boc-*N*^ε-methyl-L-lysine was synthesized as an *N*^ε-Boc derivative from *N*^ε-Boc-L-lysine by (i) benzoylation using benzaldehyde and NaBH₄, (ii) methylation using formaldehyde and NaBH₄, (iii) debenzoylation with hydrogenation using Pd/C, and (iv) protection of the ε-amino group with Boc group using Boc anhydride according to the reported procedure (Andruszkiewicz, 1988). Similarly, *N*^ε-Boc-*N*^ε,*N*^ε-dimethyl-L-lysine was synthesized from *N*^ε-Boc-L-lysine using formaldehyde and NaBH₄ for the reductive amination. *N*^ε-Boc-*N*^ε,*N*^ε,*N*^ε-trimethyl-L-lysine was synthesized using methyl iodide as a methylating reagent to quaternize the ε-amino of *N*^ε-Boc-L-lysine according to the reported procedure (Chen and Benoit, 1976).

Carboxylic acids of *N*^ε-Boc-L-lysine derivatives were activated to 3,5-dinitrobenzyl esters, and finally, *N*^ε-protecting groups were removed as described previously (Murakami et al., 2006).

Aminoacylation of tRNA^{Acn-E2} with Amino Acids

tRNA^{Acn-E2} and flexizyme dFx were prepared by the runoff transcription of appropriate templates as described previously (Murakami et al., 2006). Acceptor stem sequences of tRNA were changed from authentic *E. coli* tRNA^{Acn} to enhance transcription (T1G) and the orthogonality toward aminoacyl-tRNA synthetases (C2G). Aminoacylation was accomplished by incubating heat-denatured/rnated tRNA with equimolar of flexizyme dFx in the presence of 0.1 M HEPES-KOH (pH 7.5), 0.6 M MgCl₂, and 5 mM *K-DBE for 3 hr on ice. Reactions were quenched by acidifying the reaction with NaOAc (pH 5.2), and aminoacylated tRNAs were recovered by repeated ethanol precipitations.

Translation

The PURE system (Shimizu et al., 2001) was purchased from Post Genome Institute Company, Ltd. (Tokyo, Japan) and was used according to the manufacturer's guide. The wPURE system contained all the aminoacyl-tRNA synthetases, but contained only 13 amino acids (C, E, F, H, I, N, and W were withdrawn, where C, cysteine; E, glutamate; H, histidine). When necessary, one or more of these amino acids were added to the translation mixture. Usually, a 3–5 μl scale reaction with the addition of 0.2 mM of amino acids of choice was used for analysis. [¹⁴C]D was added to the reaction for the gel analysis of the product. Each aminoacylated tRNA^{Acn-E2} was added to a final concentration of 50 μM as a mixture with the flexizyme dFx: so that the total concentration of exogenous RNAs in the translation was 100 μM.

Analyses of the Ribosomally Synthesized H3t

*H3t peptides from the 5 μl scale translation reaction were immobilized on anti-Flag-M2 agarose (Sigma-Aldrich; St. Louis, MO) by incubating 1 hr in TBS (50 mM 0.1 M HEPES-KOH, 150 mM NaCl [pH 8.0]) with rotation at room temperature. After washing the resin with TBS briefly, peptides were eluted from the resin with 0.2% TFA by incubating 30 min at room temperature. Peptides in the eluate were bound to C18 resin (ZipTip; Millipore; Billerica, MA), washed and desalted with 0.1% TFA, and eluted directly onto the MALDI target plate with saturated matrix (*N*-cyano-4-hydroxycinnamic acid (Bruker Daltonics; Billerica, MA) in the 1:1 mixture of acetonitrile and 0.2% TFA. Average molecular masses were recorded using Autoflex III (Bruker) in a linear positive mode. The instrument was calibrated externally with peptide and protein standards (Bruker).

Preparation of chromoHP1

The chromodomain (residues 17–76) of *Drosophila* HP1 (SWISS-PROT accession code P05205) was subcloned into *Nco* I/BamHI sites of pET42a vector

(Novagen; Madison, WI) and expressed in *E. coli* strain BL21(DE3) with an N-terminal GST and hexa-His tags. The chromodomain alone was shown to be active enough to form a complex with synthetic H3t peptide (Jacobs and Khorasanizadeh, 2002). chromoHP1 in the form of N-terminal GST-His6 fusion was purified by Talon metal affinity resins (Kontech; Mountain View, CA) and dialyzed into binding buffer (20 mM imidazole, 25 mM NaCl, 2 mM DTT [pH 7.6]).

Pull-Down Analysis

A 15 μl scale translation reaction was primed with the template encoding H3t shown in Figure 4A and one of K- or *K-tRNA^{Acn-E2}_{GAA} in the absence of phenylalanine in the translation system. A translation without aminoacyl-tRNA^{Acn-E2}_{GAA} was prepared as a negative control for the pull-down analysis. *H3ts from the translation reaction were immobilized on 2.5 μl of the anti-Flag-M2 agarose by incubating 1 hr in TBS with rotation at room temperature. After brief washing with TBS, the resin was blocked with 3% BSA in TBS (1 hr at room temperature with rotation), washed again with TBS, and the resin immersed in 3 μl of 25 μM chromoHP1 in binding buffer for 1 hr at room temperature. The resin was washed twice with washing buffer (20 mM imidazole, 150 mM NaCl, 0.05% Tween20 [pH 7.5]) and once again with low Tris buffer (4 mM Tris-HCl, 10 mM NaCl [pH 8.0]). Elution was performed twice (10 min) using 0.1 M Gly-HCl (pH 3.0), and combined eluates were dried under reduced pressure and analyzed by western blot. Both anti-polyhistidine (Sigma) and anti-Flag tag antibodies (Sigma) were used in the same blot to visualize chromoHP1 and *H3t peptides, respectively.

In Vitro Binding Analysis by SPR

SPR assays were performed on a Biacore 2000. The instrument was maintained at 15°C, and the flow rate was 20 μl/min throughout the assay. Anti-Flag M2 antibody (Sigma) was diluted to 45 μg/ml in 10 mM sodium acetate (pH 5.0) (GE Healthcare; Buckinghamshire, England) and covalently immobilized on a sensor chip CM5 (GE Healthcare) using the Amine Coupling Kit (GE Healthcare). Translation reactions were diluted 100 times in HBS-EP (GE Healthcare) and captured on anti-Flag tag antibody-coated flow cell for 2 min. Then, varying concentrations of chromoHP1 (specified in figure legends) were injected and allowed to interact with antibody-bound *H3t for 2 min, resonance units in equilibrium were normalized by captured amounts, and molecular weights (ΔR_{eq}) were evaluated. Ten millimolar Glycine-HCl (pH 2.0; GE Healthcare) was used to regenerate the antibody-coated sensor chip.

SUPPLEMENTAL DATA

Supplemental Data include four figures and can be found with this article online at <http://www.cell.com/chembiol/content/full/15/11/1166/DC1/>.

ACKNOWLEDGMENTS

ack-DBE was a kind gift from H. Murakami in the Suga Group. We thank P.C. Reid for critical proofreading. This work was supported by a grant from the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) and a research and development project in the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO) to H.S.

Received: August 22, 2008

Revised: September 25, 2008

Accepted: September 29, 2008

Published: November 21, 2008

REFERENCES

- Andruszkiewicz, R. (1988). Synthesis of N-omega-tert-butoxycarbonyl, N-alpha-methylidiamino acids. *Pol. J. Chem.* 62, 257–261.
- Bonenfant, D., Towbin, H., Coulot, M., Schindler, P., Mueller, D.R., and van Costrum, J. (2007). Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. *Mol. Cell. Proteomics* 6, 1917–1932.

- Chen, F.C.M., and Benoit, N.L. (1976). New method of quaternizing amines and its use in amino-acid and peptide chemistry. *Can. J. Chem.* 54, 3310–3311.
- Chulkov, S., Kurash, J.K., Wilson, J.R., Xiao, B., Justin, N., Ivanov, G.S., McKinney, K., Tempst, P., Prives, C., Gambin, S.J., et al. (2004). Regulation of p53 activity through lysine methylation. *Nature* 432, 353–360.
- Fischle, W., Wang, Y.M., Jacobs, S.A., Kim, Y.C., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* 17, 1870–1881.
- Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438, 1116–1122.
- Forster, A.C., Tan, Z.P., Nalam, M.N.L., Lin, H.N., Qu, H., Cornish, V.W., and Blacklow, S.C. (2003). Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6353–6357.
- Garcia, B.A., Pesavento, J.J., Mizzen, C.A., and Kelleher, N.L. (2007). Pervasive combinatorial modification of histone H3 in human cells. *Nat. Methods* 4, 487–489.
- Hirota, T., Lipp, J.J., Toh, B.H., and Peters, J.M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* 438, 1176–1180.
- Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295, 2080–2083.
- Jacobs, S.A., Taverna, S.D., Zhang, Y.N., Briggs, S.D., Li, J.M., Eisenberg, J.C., Allis, C.D., and Khorasanizadeh, S. (2001). Specificity of the HP1 chromodomain for the methylated N-terminus of histone H3. *EMBO J.* 20, 5232–5241.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074–1080.
- Josephson, K., Hartman, M.C.T., and Szostak, J.W. (2005). Ribosomal synthesis of unnatural peptides. *J. Am. Chem. Soc.* 127, 11727–11735.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.
- Lachner, M., O'Carroll, N., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
- Latham, J.A., and Dent, S.Y.R. (2007). Cross-regulation of histone modifications. *Nat. Struct. Mol. Biol.* 14, 1017–1024.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838–849.
- Mateescu, B., England, P., Halgand, F., Yaniv, M., and Muchardt, C. (2004). Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep.* 5, 490–496.
- Matthews, A.G.W., Kuo, A.J., Ramon-Maiques, S., Han, S.M., Champagne, K.S., Ivanov, D., Gallardo, M., Carney, D., Cheung, P., Giocone, D.N., et al. (2007). RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450, 1106–1110.
- Meehan, R.R., Kao, C.F., and Pennings, S. (2003). HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. *EMBO J.* 22, 3164–3174.
- Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006). A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* 3, 357–359.
- Peterson, C.L., and Laniel, M.A. (2004). Histones and histone modifications. *Curr. Biol.* 14, R546–R551.
- Seet, B.T., Dikic, I., Zhou, M.-M., and Pawson, T. (2006). Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* 7, 473–483.
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755.
- Shogren-Knaak, M.A., Fry, C.J., and Peterson, C.L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 278, 15744–15748.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844–847.
- Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* 128, 1003–1012.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45.
- Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabanowitz, J., Chait, B.T., Hunt, D.F., and Allis, C.D. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N terminus. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2086–2091.
- Thomas, C.E., Kelleher, N.L., and Mizzen, C.A. (2006). Mass spectrometric characterization of human histone H3: a bird's eye view. *J. Proteome Res.* 5, 240–247.
- Winter, S., Simboeck, E., Fischle, W., Zupkovitz, G., Dohnal, I., Mechtler, K., Ammerer, G., and Selsner, C. (2008). 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. *EMBO J.* 27, 88–99.

DOI: 10.1002/cbic.200800439

Polymerization of α -Hydroxy Acids by Ribosomes

Atsushi Ohta,^[a] Hiroshi Murakami,^[b] and Hiroaki Suga^{*[a, b]}

Over 30 years ago, Fahnestock and Rich reported intriguing data showing the capability of the ribosome to polymerize phenyllactic acid. Although the polymerization was initiated and terminated randomly on polyuridylic acids, the given data convincingly suggested that the generated polymer was composed of an approximately 7:3 mixture of phenyllactic acid and phenylalanine. Despite the fact that Fahnestock's conclusion was very likely correct, there have been no reports to follow up the ribosome-catalyzed polymerization of α -hydroxy acids until very recently. At the

end of 2007, we reported messenger RNA (mRNA)-directed polyester synthesis by using the new emerging method of genetic-code reprogramming in which α -hydroxy acids with various kinds of side-chains are assigned to arbitrarily chosen codons. In this work, we have achieved the ribosomal synthesis of polyesters with the sequence composition and length in a fully controlled manner according to the sequence of mRNA. This Concept article describes the background of the method development and its application to the synthesis of polyesters.

Introduction

Can ribosomes polymerize an α -hydroxy acid?

Over 30 years ago, Fahnestock and Rich published a report entitled "Ribosome-catalyzed polyester formation".^[1] This report describes a landmark experiment in which it was shown that the ribosome is capable of polymerizing a non- α amino acid substrate, phenyllactic acid (F^{bc}). The classic Nirenberg method, which led to decoding the universal genetic code, was modified for their experiment. Instead of translating a synthetic polyuridylic (poly-U) acid by using Phe-tRNA^{Phe}, [¹⁴C]-labeled F^{bc} -tRNA^{Phe} was added to the translation system, which was composed of a cell-free *E. coli* S-100 and minimal organic/inorganic components, such as GTP and buffer (Figure 1A). The [¹⁴C]- F^{bc} -tRNA^{Phe} was chemically prepared by deamination of [¹⁴C]-Phe-tRNA^{Phe} by using nitrous acid. After the translation reaction, the polymerized products were precipitated with trichloroacetic acid (TCA) and the radioactivity of the precipitates was counted to quantify the yield of polymer. It was shown that the TCA-insoluble matter was formed only in the presence of all translation components, that is, S-100, poly-U, and GTP, and its recovery yield was nearly 10% of the poly-Phe synthesis. Upon treating the precipitate with alkaline, the resulting products were analyzed by paper electrophoresis. Supposedly, alkaline digestion would cleave ester bonds between F^{bc} - F^{bc} , but not amide bonds such as F^{bc} -Phe or Phe-Phe. In fact, the paper electrophoresis separated the products that originated from F^{bc} -Phe and F^{bc} with a ratio of approximately 3:7. The authors suggested that Phe likely originated in S-100, and thus such a minimal contamination of amino acids could not be avoided. Nonetheless, this observation gave indirect but convincing support for the idea that the polymerization of F^{bc} occurred consecutively three or four times followed by random incorporation of Phe into the poly- F^{bc} chain.

Despite the limitations of the analytical techniques available at that time, the data convincingly suggested that the ribosome is capable of catalyzing the polymerization of F^{bc} ; how-

ever, many technical as well as scientific questions were left unresolved by this demonstration. Regarding technical issues, the polyester formation was only confirmed by indirect evidence, that is, by detecting the [¹⁴C]-radioisotope of the alkaline-digested sample of the acid-insoluble precipitants. Therefore, it was still unknown how long F^{bc} could be consecutively polymerized. Ideally, the polyester should be directly detected as an intact polymer. Moreover, due to unavoidable contamination of Phe in S-100, deacylated tRNA^{Phe} generated by the hydrolysis of [¹⁴C]- F^{bc} -tRNA^{Phe} would be recharged by phenylalanyl-tRNA synthetase (PheRS). This resulted in the random incorporation of Phe into the poly- F^{bc} chain after every three or four residues. Thus, the contamination-free polymerization of F^{bc} was not, unfortunately, achieved. Moreover, it is of scientific interest whether other types of α -hydroxy acids (α -ha) can be accepted by ribosomes for polymerization. Along the same line, instead of the random initiation and termination demonstrated by Fahnestock's experiment, it is critical to utilize the full capability of the translation system, that is, to show fully controlled initiation, elongation, and termination along the mRNA sequence that leads to the generation of a defined length of polyester. Finally, it would be the most critical challenge to demonstrate the ribosomal synthesis of a variety of polyesters that contain distinct side-chains designated by mRNA templates. To this end, we set a project to challenge the ribosomal polymerization of α -hydroxy acids that would simultaneously address all of the above issues.^[2]

[a] A. Ohta, Prof. H. Suga
Department of Chemistry and Biotechnology
Graduate School of Engineering, University of Tokyo
113-8656 Tokyo (Japan)
Fax: (+81) 3-5452-5495
E-mail: hsuga@rcast.u-tokyo.ac.jp

[b] Dr. H. Murakami, Prof. H. Suga
Research Center for Advanced Science and Technology
University of Tokyo, 153-8904 Tokyo (Japan)

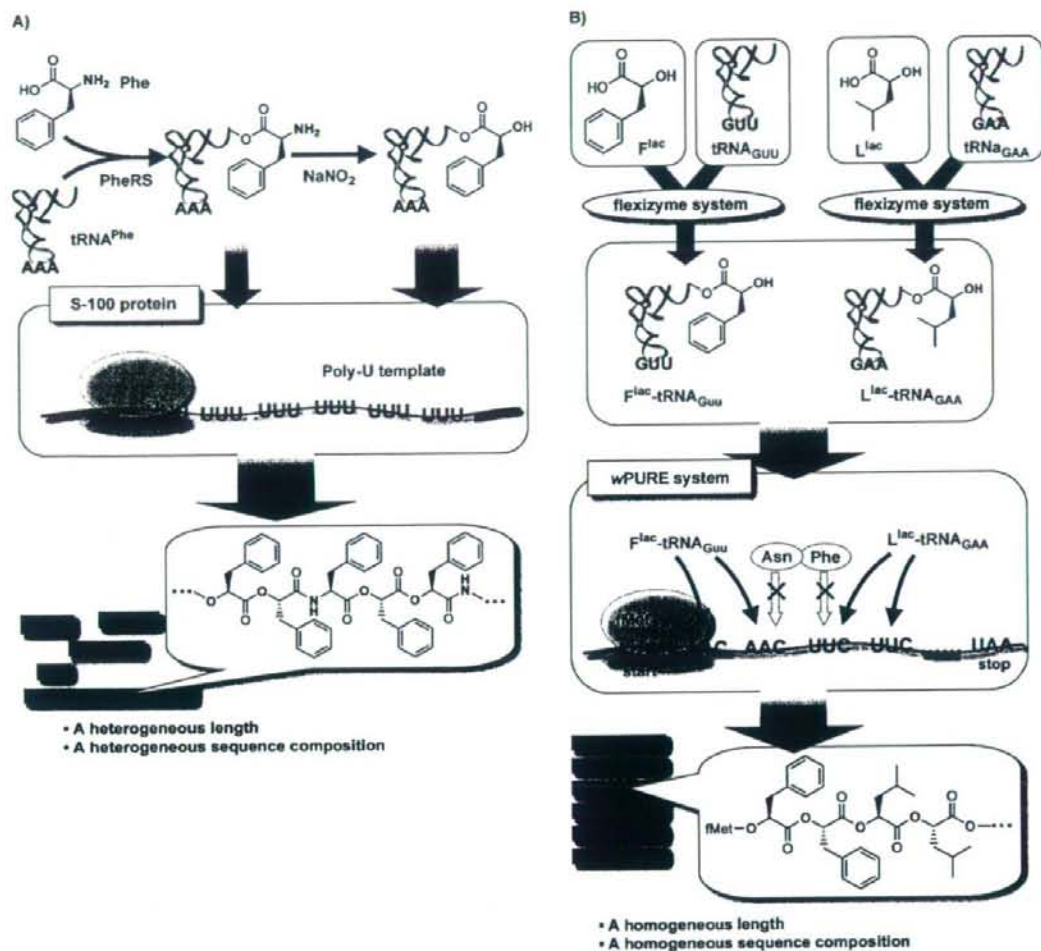


Figure 1. Two different approaches for ribosomal polyester synthesis. A) In Fahnestock's approach, in which F^{lac} -tRNA^{Phe} was prepared by chemical deamination of Phe-tRNA^{Phe}, polymerization was randomly initiated and terminated on the poly-U template. Due to the lack of an open-reading frame in the poly-U template, the resulting product had nonhomogeneous lengths of polyesters. Moreover, a trace amount of Phe-tRNA^{Phe} resulted in the random incorporation of Phe into the polyester chain; B) mRNA-directed polyester synthesis by using genetic-code reprogramming. The flexizyme system facilitates the hydroxyacylation of tRNAs, and these hydroxyacyl-tRNAs were added to a wPURE system for the polymerization of the α -hydroxy acids; L^{lac}: isopropylactic acid; F^{lac}: phenyllactic acid.

Genetic-Code Reprogramming for Polyester Synthesis

About ten years after Rich's report, a new methodology was developed for introducing nonproteinogenic α -amino acids into peptide chains.^[3,4] A mischarged tRNA_{CUA} with a nonproteinogenic amino acid was used to suppress a UAG stop codon (amber codon) and the amino acid was incorporated into a specific site of the peptide chain by using a cell-free translation system. This method is also applicable for the incor-

poration of α -hydroxy acid^[5-9] to generate an ester bond in the peptide chain. In fact, it has been utilized to disrupt the backbone hydrogen-bonding network in a protein of interest and to specifically cleave the peptide chain at the ester site. Unfortunately, the mischarged tRNA_{CUA} must inherently compete with the release factor present in the translation apparatus, that is, translation termination, and therefore the incorporation efficiency heavily depends on the kind of side chains in the α -hydroxy acid. This fact also prohibits us from performing consecutive multiple incorporations of an α -hydroxy acid(s).