Synthesis of Nonnatural mutants of \( \lambda \)-Cro repressor protein that contain an electron-accepting amino acid

By Masahiko Sisido, Shunsuke Tokunaga, and Takahiro Hohsaka

Department of Bioscience and Biotechnology, Faculty of Engineering, Okayama University, 3-1-1, Tsushimanaka, Okayama 700–8530
(Communicated by Seizo Okamura, M.J.A., Sept. 12, 2000)

Abstract: The protein biosynthetic system has been extended to incorporate nonnatural amino acids in addition to the 20 naturally occurring ones. Transfer RNAs were chemically aminoacylated with nonnatural amino acids and the aminoacylated tRNAs were added to the cell lysate of E. coli, together with the target mRNA. The positions of the nonnatural amino acids were directed by the 4-base codon/anticodon pairs that are orthogonal to the existing 3-base codon/anticodon pairs. Here, the position-specific incorporation of nonnatural amino acids was applied to incorporate an electron-accepting amino acid, 2-anthraquinonylalanine (angAla) into a DNA-binding protein, \( \lambda \)-Cro repressor. The position of angAla was directed by a CGGG 4-base colon introduced at several different positions on the mRNA. Among the several mutant proteins, the 64angAla Cro showed a strong binding to the consensus double-stranded DNA.

Key words: Nonnatural amino acids; 4-base codons; \( \lambda \)-Cro repressor; 2-anthraquinonylalanine; DNA binding.

Introduction. Incorporation of nonnatural amino acids into proteins has been carried out through extension of the genetic engineering. In 1989, Schultz’s group\(^1\) and Chamberlin’s group\(^2\) independently reported successful incorporation of a nonnatural amino acid into a specific position of proteins. They synthesized an aminoacylated tRNA by enzymatic ligation of a truncated tRNA with a chemically synthesized 3’-aminoacylated oligonucleotide. The aminoacylated tRNA was put into an E. coli lysate together with a mRNA that encodes the target protein. The position of the nonnatural amino acid was assigned by a stop codon, UAG, on the mRNA and an anticodon, CUA, on the tRNA.

We have further extended the above system to include nonnatural amino acids that carry a variety of functional side groups.\(^3\) Nonnatural amino acids with fluorescent groups, electron donors and acceptors, a photoisomerizable group, and metal ligands have been synthesized and examined if they are incorporated into proteins. In short, aromatic amino acids with straight configurations, such as 2-anthrylalanine and 2-pyrenylalanine could not be incorporated.\(^4\) The incorporation of the amino acids of specialty functions into proteins opens a way to synthesize new class of engineered proteins (nonnatural mutants) that carry a specific functional group at specific positions.

So far, the number of nonnatural amino acids that can be incorporated into a single protein has been limited to one, because only a single codon (UAG) could be assigned to the nonnatural amino acid. Multiple incorporation of nonnatural amino acids will become possible if we extend codon/anticodon pairs that work independently from the existing codon/anticodon pairs. A mathematical terminology, orthogonality, represents the condition required for the additional codon/anticodon pairs.\(^5,6\) The additional codon/anticodon pair must work as a member of the extended codon/anticodon system but must be independent from other members. We have shown that several 4-base codon/anticodon pairs satisfy the orthogonality condition.\(^5,6\) Fig. 1 illustrates how the 4-base codon/anticodon strategy works. When a CGGG sequence was successfully translated as the 4-base codon by the tRNA\(\text{cccG}\) that contains a 4-base anticodon CCCG, the protein synthesis continues to the end,
resulting in the synthesis of a full-length protein that contains a nonnatural amino acid. If the CGGG sequence was translated as the CGG 3-base codon by the endogenous tRNA_{CGG}^{Arg}, the reading frame is shifted backward by one base, resulting in an encounter of a UAA stop codon. Thus, the unsuccessful translation as the 3-base codon results in the termination of the protein synthesis. Since the CGG 3-base codon is not used frequently in E. coli system, the efficiency of the 4-base translation is as high as 70% in the case of simple nonnatural amino acids, like p-nitrophenylalanine. In this article, an attempt to incorporate electron-accepting amino acids into a DNA-binding protein Cro, by using the 4-base codon strategy will be described.

**Results and discussion.** Synthesis of various amino acids that carry electron-accepting groups and their incorporation efficiencies. The incorporation efficiencies of nonnatural amino acids markedly depend on the structure of the side chains. In designing strongly electron-accepting amino acids that can be incorporated into proteins in high yield, several different types of amino acids were synthesized and tested for their incorporation efficiency into a model protein, streptavidin, in the *E. coli in vitro* biosynthesizing system. The procedures for the preparation of nonnatural mutants and the evaluation of the incorporation efficiency are the same as described previously. The chemical structures of the amino acids are listed in Chart 1.

The aminoacylation of the tRNA with the above amino acids was carried out according to the method originally developed by Hecht and coworkers. A nonnatural amino acid was linked to the 3′-terminal of the mixed dinucleotide pdCpA to synthesize the aminoacylated pdCpA. A tRNA that lacks the 3′-terminal pCpA dinucleotide unit, tRNA(-CA), was synthesized through a run-off transcription of the corresponding DNA with T7 RNA polymerase. The aminoacylated pdCpA and the tRNA(-CA) were linked together by T4 RNA ligase. The product is the full-length tRNA aminoacylated with a nonnatural amino acid.

The nonnatural mutants of streptavidin were synthesized in *E. coli S30 in vitro* translation system that contains the aminoacylated tRNA and the target mRNA of streptavidin. In the target mRNA, a CGGG 4-base codon was introduced at the 83rd position. The translation mixture was applied to a SDS polyacrylamide gel electrophoresis, followed by the transfer to a PVDF membrane. Streptavidins were detected by using anti T7-tag monoclonal antibody and alkaliphostatase-labeled anti-Mouse IgG (Western blotting). The results are shown in Fig. 2. Of the electron-accepting amino acids examined, 2-anthraquinonylalanine showed the highest efficiency, whereas others in which the electron-accepting groups are linked with spacer chains showed very small efficiencies, except for the case of NI-1. The results indicate that only anqAla can be used as an electron-accepting amino acid for position specific incorporation into proteins.
Construction of a plasmid DNA that encodes Cro.

The DNA and amino acid sequences of wild-type Cro are shown in Chart 2.

The sequence includes a T7 tag for probing the protein by using anti-T7 tag antibody and a His6 tag for an affinity chromatography with a Co-resin (TALON™ resin) that selectively binds to the His6 tag. As described above, the histidine hexamer will appear only when the CGGG 4-base codon was successfully translated to a nonnatural amino acid and, therefore, the affinity chromatography to the His6 tag will be useful to purify the nonnatural mutants.

The gene of the above sequence was constructed by multiple PCR method using short DNA fragments as the primers. The product was isolated on a 6% PAGE, and digested with EcoRI and HindIII. The fragment was purified on a 6% PAGE, then cloned into pUC-EH to give pCro. The plasmid was introduced into a competent cell and a colony was taken up. A colony PCR was carried out and the product was analyzed on a DNA sequence.

Alternatively, the fragment was cloned into pU12 that contains a T7 tag upstream of the EcoRI site. The plasmid was introduced into a competent cell and the cell was cloned to give colonies and the target gene was obtained from the plasmid. Similarly, the target genes for mutant Cro’s that contain a CGGG 4-base codon at different positions were obtained by using different primers and identified by the sequence analysis. The DNA was transcribed to mRNA by using T7 RNA polymerase. The mRNA was stored at -30 °C.

Expression of the wild-type λ-Cro and the mutants that contain 2-anthraquinonylalanine at different positions. The mRNA and the tRNA aminoacylated with anqAla were added to the E. coli S30 in vitro biosynthesizing system. And the wild-type Cro and the nonnatural mutants incorporated with anqAla at different positions were synthesized. The anqAla was incorporated at the 10th, 26th, 32nd, 36th, 51st, 56th, 64th, 65th, and 66th positions. Western blotting indicated that the yields of the mutants markedly depended on the positions of anqAla. High yields were observed for the 32-, 56-, 64-, and 65-anqAla mutants, but only very small amount of proteins were produced for other mutants. The small amount of the latter mutants may suggest that those proteins are misfolded and quickly decomposed by proteases in the S30 system.

The mutants that are produced in high yields were purified by an affinity chromatography using Co-resin and used for the binding experiment to the DNA.

DNA-binding of the wild-type and the mutant
Cro is known to bind to three operator sequences (OR1, OR2, OR3) with the smallest dissociation constant to OR3 (2 × 10^{-12}M). The consensus sequence of the three operator sequences as shown above exhibits a little smaller dissociation constant (1.2 × 10^{-12}M). The binding of the wild-type and the 64-angAla mutants to the consensus sequence was examined from the change of polarization of the FITC fluorescence covalently attached to the ds DNA, at pH 7.4 and 20 °C. The titration curves for the wild-type and the 64angAla Cro in the presence of 1 nM of the FITC-labeled ds DNA are shown in Fig. 3.

The wild-type Cro shows a typical saturation curve indicating that high binding activity is retained in the *in vitro* translation product, despite the presence of a T7 tag and a His_{6} tag at the N- and C-end, respectively. Similarly, the 64angAla mutant exhibits the saturation curve, despite a T7 tag, a His_{6} tag, and the incorporation of large angAla unit at the 64th position. It must be noted that the Cro binds to the ds DNA in a dimeric form. The

*Cro’s to the consensus sequence of double-stranded DNA.* The binding of the wild-type and the mutant Cro’s to double-stranded (ds) DNA was examined by using an FITC-labeled DNA of the following sequence:

FITC-5’-TCTATCACCGGGTGATAAA-3’
3’-AGATAGTGCCGCCACTATT-5’

Fig. 2. Relative incorporation efficiencies of various nonnatural amino acids carrying electron-accepting groups into streptavidin. The yield of the wild-type streptavidin is taken to be 100%.

Fig. 3. Fluorescence polarization of the FITC-labeled ds DNA of the consensus sequence in the presence of different amount of protein solution that contains the wild-type Cro or the 64angAla mutant. [FITC-dsDNA] = 1 nM, 20 °C, pH=7.4. The protein concentrations can be evaluated from the turning points of the titration curve. In the control experiment, the reaction mixture of the *in vitro* system that contain no mRNA was added.

![Fig. 4. The computer-predicted conformation of dimeric 64angAla Cro bound to a consensus sequence of double-stranded DNA.](image)
binding of the wild-type and the 64anqAla Cro suggests that the proteins form the active dimer, despite the presence of the tags and the nonnatural amino acids.

The 32-, 56-, and 65-anqAla mutants induced small increase of the fluorescence polarization without the saturation, indicating weak binding to the consensus ds DNA. In these cases, the large side group of the anqAla unit may be disadvantageous for the dimer formation or for the DNA binding.

The positions and orientations of the anqAla units in the dimer of 64anqAla Cro bound to the ds DNA were predicted from molecular mechanics calculations. The starting conformation was taken from the X-ray crystallographic structure of wild-type Cro bound to the ds DNA. The energy minimization was made varying only the side groups of the anqAla units and of the neighboring amino acids. The computer-predicted conformation is illustrated in Fig. 4. The anthraquinonyl groups are located close to the DNA main chain. It is particularly interesting that the consensus sequence contains a GGG sequence that is known to deliver an electron to an anthraquinonyl group in the photoexcited state.

Therefore, the 64anqAla Cro is a potential photocleaving enzyme that is very specific to the consensus sequence. The alteration of enzyme function by the incorporation of nonnatural amino acid as proposed in this paper will open a new field of protein engineering.

**Experimental section.** The nonnatural amino acids were synthesized in our laboratory. AQ-m’s (m=1–4) were synthesized by the coupling of α-Boc, α,ω-diaminoalkanoic acids with 2-chlorocarbonyl-anthraquinone in DMF. Similarly, NI-m’s (m=1–4) were synthesized by the reaction of α-Boc, α,ω-diaminoalkanoic acids with 1,8-naphthalic anhydride in DMF, followed by the imide formation at elevated temperature (100 °C). DL-2-Anthraquinonylalanine was synthesized as before.

Synthesis and purification of nonnatural mutants in the E. coli in vitro system has been described before.4)

**Acknowledgment.** This work has been supported by a Grand-in-Aid for Specially Promoted Research from the Ministry of Education, Science, Sports and Culture of Japan (No. 11102003).

**References**