READING OF THE LYSINE CODONS IN THE MS 2 COAT PROTEIN CISTRON DURING PROTEIN SYNTHESIS IN VITRO

Per ELIAS, Florentyna LUSTIG, Torsten AXBERG, Bengt ÅKESSON and Ulf LAGERKVIST
Department of Medical Biochemistry, University of Gothenburg, S-400 33 Gothenburg, Sweden

Received 13 October 1978

1. Introduction

The translation of the genetic information contained in the messenger RNA is based on the recognition between codon and anticodon. The first two positions of the codon are read by the anticodon strictly according to the rules of classic base pairing. However, in the reading of the third codon position by the wobble nucleotide of the anticodon a greater degree of freedom must be postulated in order to allow for the discrepancy between the large number of codons in a degenerate code and the limited number of anticodons available for reading them. In his wobble hypothesis [1] Crick has laid down the restrictions which presumably govern the reading of the third codon position. Obviously, the reading of this position must be restricted in all cases where there would otherwise be a mistake in the synthesized protein. For intance, in the phenylalanine/leucine codon group (UUU, UUC, UUA and UUG), an anticodon with G in the wobble position cannot recognize the codons UUA and UUG because that would lead to the introduction of phenylalanine instead of leucine. We have introduced the term codon family to denote a group of four codons which all code for the same amino acid and we have asked ourselves if the rules of the wobble hypothesis apply as strictly here as in the phenylalanine/leucine case, in spite of the fact that it makes no difference to translational fidelity how the third position of the codon is read because the first two codon nucleotides are enough to specify the amino acid. To answer this question, the codon–anticodon recognition in the valine codon family was investigated by using an in vitro system from *Escherichia coli* programmed with MS2-RNA, in which under normal conditions most of the protein synthesized is MS2 coat protein. The primary structure of the coat protein cistron in MS2-RNA has been determined and can be compared to the known amino acid sequence of the coat protein [2]. By measuring the incorporation of labeled valine from valyl-tRNAs with different anticodons into peptide positions corresponding to the four valine codons (GUU, GUC, GUA and GUG) we could show that each of the valine anticodons tested (U*AC, GAC and IAC) recognized all four valine codons [3]. On the basis of these results we postulated the existence of an alternative reading method which we referred to as reading 'two out of three', disregarding the third codon nucleotide. We also suggested that 'two out of three' misreading could take place in all codon families and a theory was developed which implied that the genetic code is organized in such a way as to prevent the 'two out of three' method from being used when it might compromise translational fidelity, i.e., anywhere outside the codon families [4]. This hypothesis predicts, for instance, that while misreading by 'two out of three' can be tolerated in the valine codon family it is prohibited in the reading of the lysine codons.

In a recent publication [5] Mitra has used the

Abbreviations: $tRNA^{Lys}$, lysine tRNA; U*, 5-oxy-acetic acid uridine monophosphate; s^2U, 5-methoxycarbonylmethyl-2-thio uridine monophosphate. MSH, mercaptoethanol; PCA, perchloric acid;
A triplet of nucleotides in parentheses after a tRNA denotes the anticodon: $tRNA^{Lys}$ (CUU), lysine tRNA with anticodon CUU

MS2-RNA programmed system to investigate the ability of $tRNA_1^{Lys}$ (CUU) to read the codon AAA, the predominant lysine codon in the MS2 coat protein cistron, which contains five AAA codons and only one AAG. No attempt was made to authenticate the product of the reaction but Mitra nevertheless claims that >90% of the protein formed in his experiment, measured as acid-precipitable counts, is coat protein and he concludes that $tRNA_1^{Lys}$ (CUU) recognizes the codon AAA under his conditions of in vitro protein synthesis. In what follows we will show that this claim is unfounded and the conclusion erroneous. With lysyl-$tRNA_1^{Lys}$ (CUU) as the only source of lysine for protein synthesis there is no detectable formation of coat protein in the system and we must consequently conclude that $tRNA_1^{Lys}$ (CUU) cannot read the codon AAA.

2. Materials and methods

2.1. *Preparation of antiserum*

Lysine:tRNA ligase from *E. coli* was purified to homogeneity as in [6,7]. The enzyme was dissolved in 0.5 ml 0.05 M K-phosphate buffer (pH 7.5) to give 1 mg enzyme/ml final conc. This solution was mixed with an equal volume of Freunds complete adjuvants. Female albino rabbits were injected subcutaneously in the back with this solution. The injection was repeated 3–4 times, using Freunds incomplete adjuvants, with 2 weeks between each injection. The rabbits were bled by cardiac puncture and the immunoglobulins were purified essentially as in [8] and dialysed against 0.01 M NaCl–0.01 M Tris–HCl (pH 7.5) prior to use.

2.2. *Purification of tRNA*

Crude tRNA from yeast was obtained from Boehringer Mannheim. Chromatography on benzoylated DEAE-cellulose as in [9] separated $tRNA_1^{Lys}$ and $tRNA_2^{Lys}$ from each other. These tRNAs were further purified by using chromatography on benzoylated DEAE-cellulose after phenoxyacetylation of the esterified lysyl-tRNAs [10] as in [3,11]. $tRNA_1^{Lys}$ (anticodon CUU) and $tRNA_2^{Lys}$ (anticodon s^2UUU) were also identified in a ribosome binding assay [12] using their respective codon triplets, ApApG and ApApA [5]. The structure of the $tRNA_1^{Lys}$ was determined in [13] and that of $tRNA_2^{Lys}$ in [14]. The acceptor capacity of the tRNA fractions used was ~1 nmol/A_{260} unit. Esterification of $tRNA^{Lys}$ was carried out essentially as in [3].

2.3. *Conditions for in vitro protein synthesis*

Preparation of the S30-extract from *E. coli* Q_{13} cells was essentially as in [3]. The in vitro protein synthesizing system was as in [15] with the following modifications. Magnesium acetate was not included in mix I but added separately to give 11 mM final conc. in the incubation mixture. A mixture of 19 L-amino acids (without lysine) was added to give 0.08 mM for each amino acid. Following the addition of S30 extract corresponding to ~200 μg protein, 1 A_{260} unit of crude tRNA from *E. coli*, 1 A_{260} unit of MS2 RNA and 0.1 μCi [^{14}C]lysine (350 mCi/mmol) the total reaction volume was adjusted with water to 50 μl final vol. When protein synthesis was carried out in the presence of antibodies the complete reaction mixture, which contained [^{14}C]lysyl-tRNA (instead of [^{14}C]lysine) and a 1000-fold excess of unlabeled L-lysine, was preincubated on ice for 2–3 min followed by incubation at 37°C for 30 min. The reaction was stopped by the addition of 1 ml 0.4 M PCA and the sample was hydrolysed for 20 min at 90°C. The sample was chilled and the precipitate was collected on Whatman GF/C 2.5 cm glass fiber discs and rinsed with 0.4 M PCA. Determinations of radioactivity were as in [3].

For identification of the products formed under these conditions the following procedures were used. The reaction mixtures above and in table 1 were scaled up 100 times. The material insoluble in 0.4 M PCA was isolated as in [16], dissolved in 8 M urea–0.1 M MSH–0.1 M NH_4HCO_3 (pH 8.0) and chromatographed on a calibrated column of Sephadex G-75 (2 × 100 cm). The eluate was monitored for radioactivity and the combined fractions corresponding to the coat protein peak were dialysed against distilled water and lyophilized. This material was then digested with trypsin and the peptides were separated using chromatography and high-voltage electrophoresis in two dimensions as in [3]. The radioactive spots were visualized by autoradiography.

3. Results

3.1. *Total protein synthesis measured as acid precipitable radioactivity*

In order to compare the ability of lysine tRNAs with different anticodons to recognize the lysine codons AAA and AAG we have used an in vitro protein synthesizing system programmed with MS2-RNA. Under normal conditions this system synthesizes most of its protein as coat protein by preferential translation of the MS2 coat protein cistron which contains five AAA and only one AAG. The system used should preferably be strictly dependent on added lysyl-tRNA so as to rule out competition between lysyl-tRNAs as well as complications caused by transaminoacylation and re-esterification [3]. To accomplish this we used antibodies against lysine: tRNA ligase from *E. coli*. The antibody preparation was found to completely abolish any incorporation of free lysine into protein in the system (table 1). The system was also free of transaminoacylation and re-esterification. Two $tRNA^{Lys}$ isoacceptors from yeast, $tRNA_1^{Lys}$ (CUU) and $tRNA_2^{Lys}$ (s^2UUU), were esterified with ^{14}C-lysine and their ability to sustain protein synthesis in the lysyl-tRNA-dependent system was investigated. The results showed that $tRNA_1^{Lys}$ (CUU) was less effective in this respect than $tRNA_2^{Lys}$ (s^2UUU) or unfractionated lysyl-tRNA from *E. coli* (table 1). A time curve (fig.1) showed that the incorporation of lysine into protein using lysyl-$tRNA_1^{Lys}$ (CUU) was considerably slower than when lysyl-$tRNA_2^{Lys}$ (s^2UUU) was used and the same result was obtained when the reaction rate was studied as a function of the concentration of lysyl-$tRNA^{Lys}$ in the incubation (fig.2).

3.2. *Recognition of the lysine codons by $tRNA_1^{Lys}$ (CUU) and $tRNA_2^{Lys}$ (s^2UUU)*

The results presented in section 3.1. only show that polypeptide material, large enough to be acid precipitable, was formed in our in vitro system when lysyl-$tRNA_1^{Lys}$ (CUU) or lysyl-$tRNA_2^{Lys}$ (s^2UUU) was the only source of lysine for protein synthesis. In order to draw any conclusions about the ability of the two isoacceptors to read the lysine codons in the coat protein cistron, the product formed in the incubation

Table 1
Protein synthesis in vitro dependent on added lysyl-$tRNA^{Lys}$

Source of lysine	Antibodies against lysine:tRNA ligase	MS2-RNA	pmol lysine incorporated	Yield (%)
[^{14}C]Lysine	–	–	0.10	
[^{14}C]Lysine	–	+	19.5	
[^{14}C]Lysine	+	–	0.09	
[^{14}C]Lysine	+	+	0.24	
[^{14}C]Lysyl-$tRNA_1^{Lys}$	+	–	0.07	
[^{14}C]Lysyl-$tRNA_1^{Lys}$	+	+	2.13	10
[^{14}C]Lysyl-$tRNA_2^{Lys}$	+	–	0.08	
[^{14}C]Lysyl-$tRNA_2^{Lys}$	+	+	7.71	30
Unfractionated [^{14}C]Lysyl-$tRNA^{Lys}$	+	–	0.15	
Unfractionated [^{14}C]Lysyl-$tRNA^{Lys}$	+	+	5.00	22

For information on the tRNAs used see sections 3 and 2.2. Conditions were as in section 2 with the modifications indicated in the table. Yields were calculated on the basis of [^{14}C]lysyl-$tRNA^{Lys}$ added. The antibody preparation used caused a general depression of protein synthesis (not dependent on the inhibition of the lysine:tRNA ligase) of ~50%

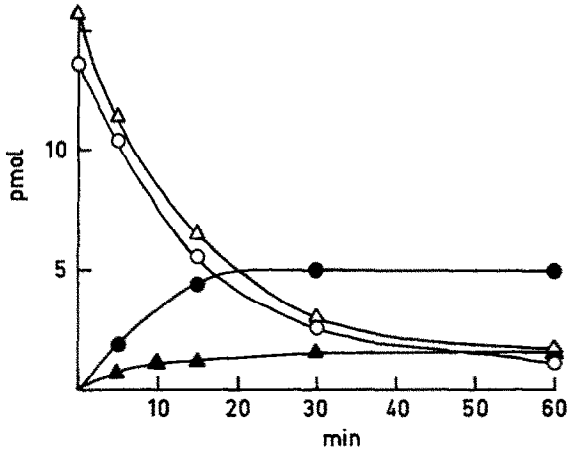

Fig.1. Incorporation of lysine from lysyl-tRNA^{Lys} into protein and stability of lysyl-tRNA^{Lys} during protein synthesis in vitro. The incubation was as in section 2 and table 1. The amount of remaining [^{14}C]lysyl-tRNA$_1^{Lys}$ (△) and [^{14}C]lysyl-tRNA$_2^{Lys}$ (○) was calculated as the difference between radioactivity insoluble in cold and hot PCA. Lysine incorporation into protein from [^{14}C]lysyl-tRNA$_1^{Lys}$ (▲) and [^{14}C]lysyl-tRNA$_2^{Lys}$ (●) was determined as radioactivity insoluble in hot 0.4 M PCA.

must be identified as coat protein. Figure 3 shows the results obtained when the reaction products were chromatographed on a calibrated Sephadex G-75 column. When free [^{14}C]lysine was used (without

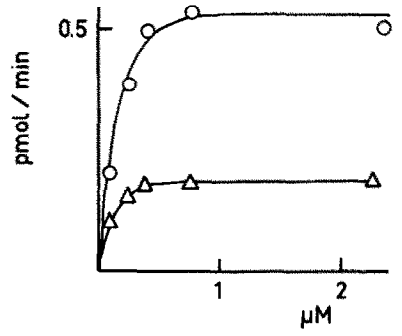

Fig.2. Rate of [^{14}C]lysine incorporation into protein as a function of [^{14}C]lysyl-tRNA^{Lys} concentration. Incubation was as in section 2 and table 1. For each concentration a time curve was obtained and the rate of incorporation was determined from the linear part of the curve. The rate of incorporation from [^{14}C]lysyl-tRNA$_1^{Lys}$ (△) and [^{14}C]lysyl-tRNA$_2^{Lys}$ (○) was then plotted against the concentration of lysyl-tRNA^{Lys}. For further information see the legend to fig.1.

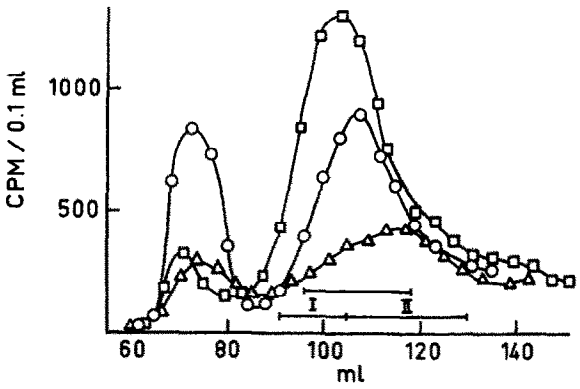

Fig.3. Chromatography of protein labeled with [^{14}C]lysine in vitro on a Sephadex G-75. The chromatographic profiles represent experiments where free [^{14}C]lysine (□), [^{14}C]lysyl-tRNA$_1^{Lys}$ (△) and [^{14}C]lysyl-tRNA$_2^{Lys}$ (○) have been used. For further experimental details see section 2. The bars indicate how the chromatographic fractions were pooled for further identification in the experiments with [^{14}C]lysyl-tRNA$_1^{Lys}$ (lower bars denoted I and II) and [^{14}C]lysyl-tRNA$_2^{Lys}$ (upper bar).

antibodies) ~80% of the labeled protein formed was eluted in the position to be expected for authentic coat protein while with [^{14}C]lysyl-tRNA$_2^{Lys}$ (s^{2}UUU) ~50% of the labeled material appeared in this position. When [^{14}C]lysyl-tRNA$_1^{Lys}$ (CUU) was used, however, the resulting chromatographic peak was very broad with a maximum that appeared significantly later than expected for coat protein. These results raised the possibility that with lysyl-tRNA$_1^{Lys}$ (CUU) we were producing some polypeptide material not related to coat protein. This suspicion was confirmed when the labeled protein synthesized in the different experiments was digested with trypsin and the resulting peptides separated by a combination of paper chromatography and high-voltage electrophoresis [3]. Inspection of fig.4 shows that when [^{14}C]lysyl-tRNA$_2^{Lys}$ (s^{2}UUU) was used the resulting protein fractions (pooled as indicated in fig.3) gave the pattern of lysine peptides to be expected for authentic coat protein. An identical pattern (not shown) was obtained when free [^{14}C]lysine was used. However, with [^{14}C]lysyl-tRNA$_1^{Lys}$ (CUU) as the source of lysine for protein synthesis an entirely different result was obtained. In this case the tryptic peptides had no resemblance to the pattern expected for coat protein.

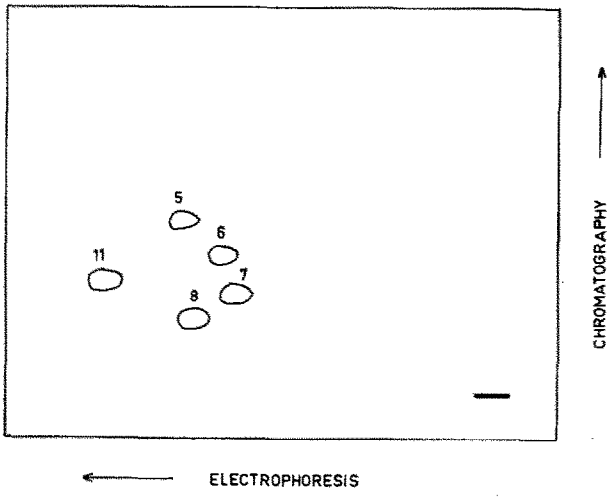

Fig.4a

PEPTIDE	SEQUENCES
5	UAC · ACC · AUC · AAA Tyr – Thr – Ile – Lys 58 61
6	GCA · AUG · CAA · CGU · CUC · CUA · AAA Ala – Met – Gln – Gly – Leu – Leu – Lys 107 113
7	GUC · GAG · GUG · GCU · AAA Val – Glu – Val – Pro – Lys 62 66
8	UCA · CAG · GCU · UAC · AAA Ser – Gln – Ala – Tyr – Lys 39 43
11	AAA · UAC · ACC · AUC · AAA Lys – Tyr – Thr – Ile – Lys 57 61

Fig.4b

Fig.4c

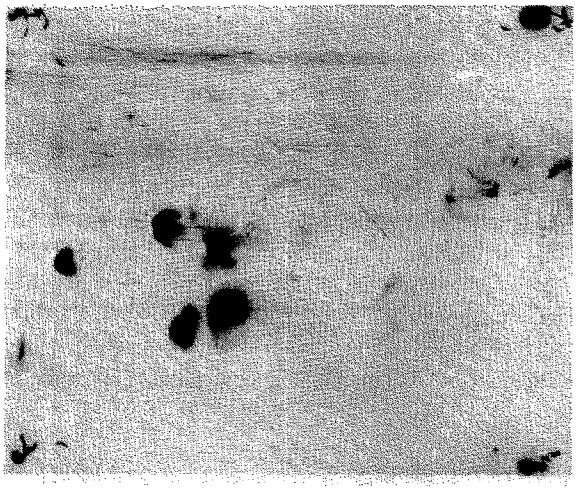

Fig.4d

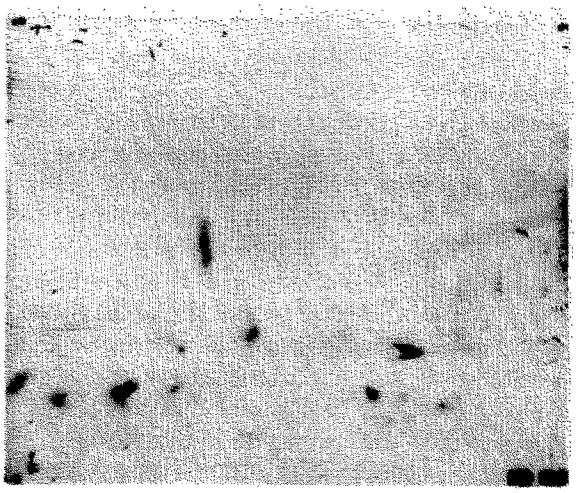

Fig.4. Finger printing of protein labeled in vitro with [^{14}C]lysine. For experimental details see section 2.

(a) Tryptic fingerprint showing the lysine-containing peptides from MS2 coat protein which can be separated by a combination of chromatography and high-voltage electrophoresis. Peptide 11, which contains an N-terminal lysine, results from the preferential cleavage between Arg 56 and Lys 57, retaining the bond between Lys 57 and Tyr 58 [17].

(b) Sequences of the lysine containing peptides shown in fig.4a and the corresponding sequences in MS2-RNA [2].

(c) Autoradiogram showing the radioactive tryptic peptides labeled from [^{14}C]lysyl-tRNA$_2^{Lys}$. The chromatographic fractions containing coat protein were pooled as indicated in fig.3.

d) Autoradiogram showing the radioactive tryptic peptides labeled from [^{14}C]lysyl-tRNA$_1^{Lys}$. The chromatographic fractions were pooled as indicated in fig.3.

Both fraction I and II (fig.3) from this experiment gave the same fingerprint with only slight variations in the intensities of the different spots in the autoradiogram.

From these results we conclude that $tRNA_1^{Lys}$ (CUU) cannot read the AAA codons in the MS2 coat protein cistron. The limited amount of polypeptide material formed with this isoacceptor is probably the result of reading the AAG-rich stretch of the replicase cistron. On the other hand, $tRNA_2^{Lys}$ (s^2UUU) would seem to be able to read not only the AAA codons but also the single AAG codon, which in the coat protein cistron immediately precedes the codons corresponding to peptide 6 (fig.4). The fact that this peptide is labeled, strongly indicates that AAG must have been read. Direct proof of the reading of AAG must await suitable methods for cleaving the large and insoluble peptide corresponding to this codon.

4. Discussion

The results presented above clearly show that when lysyl-$tRNA_1^{Lys}$ (CUU) is the only source of lysine for protein synthesis in an in vitro system programmed with MS2-RNA, there is no detectable formation of MS2 coat protein. This is in contrast to the result when lysyl-$tRNA_2^{Lys}$ (s^2UUU) or unfractionated lysyl-tRNA is the source of lysine. We must therefore conclude that under our conditions of in vitro protein synthesis $tRNA_1^{Lys}$ (CUU) does not read the codon AAA, which makes up five of the six lysine codons in the cistron. To put it in another way, the lysine codon AAA cannot be misread by the 'two out of three' method, disregarding the third nucleotide.

One may then ask why 'two out of three' misreading is not allowed in the case of the lysine codons when it is clearly permitted under the same conditions in the valine codon family [3]. To answer this question we must briefly consider a hypothesis [4] to account for the organization of the genetic code in terms of the necessity of restricting 'two out of three' misreading to such situations where it could not compromise translational fidelity. The hypothesis is based on the assumption that the probability of reading a codon by the 'two out of three' method is a function of the strength of the interaction between the anticodon and the first two codon nucleotides. It is also assumed that, in codon–anticodon recognition, an interaction of the G · C type, involving three hydrogen bonds, is stronger than an A · U interaction with only two bonds. The hypothesis further points out that the codons of the genetic code are laid out in such a way that codons that represent a high probability of reading 'two out of three' are strictly confined to the codon families in which the 'two out of three' method can be used with impunity. On the other hand, those places in the code where this method could lead to translational errors are exclusively occupied by low-probability codons. This organization of the code and the competition with tRNAs having anticodons able to read all three positions of the codon would effectively prevent the 'two out of three' method from being used when it might compromise translational fidelity. A prediction of this hypothesis is that misreading of the lysine codons by the 'two out of three' method should not be possible since these codons make only weak A · U type interactions in the first two positions. On the other hand, 'two out of three' misreading would be allowed in, for instance, the valine codon family.

An interesting question remains to be answered. Are there any restrictions in the ability of $tRNA_2^{Lys}$ (s^2UUU), to recognize the AAG codon in our system? It has been claimed that the introduction of an SH-group in the 2-position of the wobble nucleotide U restricts its ability to wobble with G in the third codon position [18,19]. We are trying to answer this question in experiments where $tRNA_2^{Lys}$ (s^2UUU) and $tRNA_1^{Lys}$ (CUU) compete with each other for the codon AAG in the MS2 coat protein cistron.

Acknowledgement

We are indebted to Mrs Anne-Marie von Essen, Miss Marianne Wedin and Miss Lena Ahlström for expert technical assistance. This investigation was supported by a grant to Ulf Lagerkvist from the Swedish Cancer Society.

References

[1] Crick, F. H. C. (1966) J. Mol. Biol. 19, 548–555.
[2] Min Jou, W., Haegeman, G., Ysebaert, M. and Fiers, W. (1972) Nature 237, 82–88.

[3] Mitra, S. K., Lustig, F., Åkesson, B., Lagerkvist, U. and Strid, L. (1977) J. Biol. Chem. 252, 471–478.
[4] Lagerkvist, U. (1978) Proc. Natl. Acad. Sci. USA 75, 1759–1762.
[5] Mitra, S. K. (1978) FEBS Lett. 91, 78–80.
[6] Waldenström, J. (1968) Eur. J. Biochem. 3, 483–487.
[7] Åkesson, B. and Lundvik, L. (1978) Eur. J. Biochem. 83, 29–36.
[8] Roufa, D. J. and Leder, P. (1971) J. Biol. Chem. 246, 3160–3167.
[9] Gillam, I., Millward, S., Blew, D., Von Tigerstrom, M., Wimmer, E. and Tener, G. M. (1967) Biochemistry 6, 3043–3056.
[10] Gillam, I., Blew, D., Warrington, R. C., Von Tigerstrom, M. and Tener, G. M. (1968) Biochemistry 7, 3459–3468.
[11] Rymo, L., Lagerkvist, U. and Wonacott, A. (1970) J. Biol. Chem. 245, 4308–4316.
[12] Nirenberg, M. and Leder, P. (1964) Science 145, 1399–1407.
[13] Smith, C. J., Teh, H. S., Ley, A. N. and D'Obrenan, P. (1973) J. Biol. Chem. 248, 4475–4485.
[14] Madison, J. T. and Boguslawski, S. J. (1974) Biochemistry 13, 524–527.
[15] Nirenberg, M. W. (1963) Methods Enzymol. 6, 17–23.
[16] Nathans, D. (1965) J. Mol. Biol. 13, 521–531.
[17] Desnuelle, P. (1960) in: The Enzymes (Boyer, P. D., Lardy, H. and Myrbäck, K. eds) 2nd edn. vol. 4, 119–132.
[18] Ohashi, Z., Saneyoshi, M., Harada, F., Hara, H. and Nishimura, S. (1970) Biochem. Biophys. Res. Commun. 40, 866–872.
[19] Yoshida, M., Takeishi, K. and Ukita, T. (1971) Biochim. Biophys. Acta 228, 153–166.