*THE INITIATION OF VIRAL PROTEIN SYNTHESIS IN E. COLI EXTRACTS**

By Daniel Kolakofsky† and Tokumasa Nakamoto

DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CHICAGO

Communicated by Dwight J. Engle, September 30, 1966

Early studies on the initiation of polypeptide synthesis were concerned primarily with establishing the events which occurred prior to the formation of the first peptide bond rather than with determining how synthesis was initiated. The binding of phenylalanyl-sRNA and lysyl-sRNA to ribosomes in the presence of poly U and poly A, respectively, prior to incorporation of the amino acids into polypeptides was demonstrated.[1-3] The phenylalanyl-sRNA bound to the ribosomes with poly U was shown to initiate polyphenylalanine synthesis.[4, 5] These as well as other studies with homopolymers and mixed polymers of RNA demonstrated that ribosomes in these systems became attached to various RNA's with totally different nucleotide sequences and initiated synthesis, i.e., the ribosomal attachment was not selective. The possibility that the ribosomes recognized a specific end of the RNA and initiated synthesis was ruled out when it was shown that the overlaps of block polymer were translated in all possible combinations.[6]

The discovery of N-formylmethionyl-sRNA has led to the recent accelerated development in the study of initiation of protein synthesis.[7] N-formyl-methionine was shown to be incorporated into polypeptide and viral proteins.[8-11] Noll suggested that protein synthesis was initiated when N-formylmethionyl-sRNA, because of its resemblance to peptidyl-sRNA, was enzymatically transferred to the peptidyl-sRNA site on the ribosome, i.e., the specificity of the enzyme controlled initiation.[12] We proposed a mechanism in which a single triplet served as the initiating signal for protein synthesis under conditions where ribosomes interacted randomly with mRNA.[13] We suggested that initiation of synthesis was regulated in the alignment of two aminoacyl-sRNA's on the ribosome by a specificity difference of the two ribosomal sites; that is, because the first site accommodated N-formylmethionyl-sRNA but not any other aminoacyl-sRNA, the first peptide bond was formed only when a triplet directed the binding of N-formylmethionyl-sRNA to the first site. We also suggested that N-formylmethionyl-sRNA was able to occupy the first site because of its similarity to peptidyl-sRNA. To account for polypeptide synthesis with RNA polymers not containing the codon for N-formyl-methionine, we postulated that strong ribosome-messenger interaction as well as high magnesium concentrations enabled nonformylated aminoacyl-sRNA's to initiate synthesis.[14] Evidence supporting our proposal was obtained by showing that amino acid incorporation with poly AGU was dependent on N-formylmethionyl-sRNA at low magnesium concentrations but not at high concentrations, and that the poly AGU- and poly AU-directed polymerizations were enhanced by the dipeptidyl-sRNA of phenylalanine at low magnesium concentrations.

We have recently extended our studies to the synthesis of viral proteins in *E. coli* extracts with several viral RNA's, and have been able to demonstrate a re-

quirement for either a formyl donor or N-formylmethionyl-sRNA for maximal synthesis. The results are presented in this communication.

Methods.—Preparation of bacterial extract (S-30): *E. coli* cells, mixed with Macaloid (10 mg/gm of wet cells) were ground with alumina and extracted with 2 vol of a solution containing 0.01 *M* Tris, pH 7.4, 0.01 *M* $MgCl_2$, 0.04 *M* KCl, and 0.001 *M* 2-mercaptoethanol. DNase (about 1 μg/ml of extracting buffer) was added and the extract was allowed to stand in the cold for 10 min. The extract was then centrifuged twice at 30,000 $\times$ *g* for 15 min and the S-30 was dialyzed overnight against the extraction buffer. The extract was then preincubated for 1 hr at 37° with all the components of the reaction mixture described below except for the omission of the viral RNA. The preincubated S-30 was passed through a Sephadex G-25 column and stored at 0°.

Preparation of viral RNA: MS2 phage (original culture provided by Dr. S. B. Weiss) was prepared by infecting *E. coli* C-3000 in tryptone, yeast extract medium. The phage was purified by $(NH_4)_2SO_4$ fractionation, gel filtration, and banding in a CsCl density gradient.

R17 *am*11B phage (obtained from Dr. J. D. Watson's laboratory) was grown and purified in a similar manner except that *E. coli* S26, *sup*$^+$ of A. Garen was used as host. The RNA was isolated from the phage by phenol treatment and ethanol precipitation. The MS2 RNA prepared in this fashion was found to be highly infective in an *E. coli* spheroplast system.

Assay procedure: The reaction mixture, unless otherwise specified, contained 0.05 *M* imidazole, pH 7.0, 0.03 *M* KCl, 0.01 *M* 2-mercaptoethanol, 0.004 *M* PEP, 0.003 *M* ATP, 0.0002 *M* GTP, 5 μg pyruvate kinase, 0.025 ml of S-30, and 4×10^{-5} *M* each of the C^{12}-amino acids, all in a final volume of 0.25 ml. Viral RNA, $MgCl_2$, C^{14}-amino acids, and N^5-formyltetrahydrofolate were added as indicated. Incubation was for 30 min unless otherwise specified. Radioactivity was assayed as previously described.[13]

Materials: The *E. coli* and Macaloid were obtained from sources previously mentioned.[13]

Results.—The major problem we encountered in making protein synthesis directed by MS2 RNA dependent on formylation was in depleting the S-30 preparation of endogenous formyl donors without fractionating the *E. coli* extract. Even when the S-30 preparation was dialyzed before preincubation and then either dialyzed or passed through a Sephadex column after preincubation, the dependence of protein synthesis on a formyl donor was often quite small. In Figure 1, the effect of a formyl donor on protein synthesis at different levels of magnesium with a relatively fresh preparation of S-30 is shown. The upper curve represents incorporation of leucine with the addition of N^5-formyltetrahydrofolic acid (N^5-formyl-THFA), and the lower curve represents incorporation without the formyl donor. The stimulation of protein synthesis by the formyl donor is only about twofold. Note that maximal synthesis occurs at approximately 0.009 *M* magnesium.

An increased dependence of amino acid incorporation on the formyl donor was seen when the S-30 preparation was kept at 0° for several days after preincubation. The results shown in Figure 2 were obtained with the same S-30 preparation which had been used 6 days earlier in the experiments of Figure 1. The upper curve again represents leucine incorporation with the formyl donor at different levels of

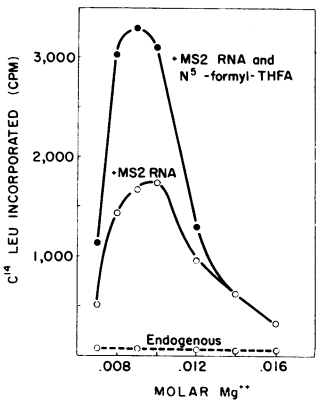

FIG. 1.—Effect of formylation on MS2 protein synthesis with fresh S-30. The general reaction mixture was as described in *Methods* and contained C^{14}-leucine of specific activity 42.3 $\mu c/\mu$mole. Where specified, 20 μg MS2 RNA and 30 μg of N^5-formyl-THFA were added. The S-30 was assayed within 1 hr after preparation.

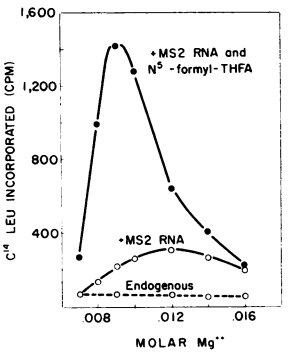

FIG. 2.—Effect of formylation on MS2 protein synthesis with aged S-30. Conditions for incubation were similar to that described in Fig. 1.

magnesium. The incorporation of amino acids has become considerably more dependent on the formyl donor. At 0.007 *M* magnesium the incorporation is completely dependent on formylation. A high, sharp peak of incorporation with N^5-formyl-THFA is again obtained at 0.009 *M* magnesium. However, incorporation without the formyl donor is now quite low. An equally interesting change is the higher magnesium concentration required for maximal synthesis in the absence of the formyl donor. The pattern is somewhat similar to that obtained with poly AGU in that the system is more dependent on a formyl donor at low levels of magnesium and less at high. Although the incorporation observed in the absence of the formyl donor might still be due in part to residual formylating activity of the system, the higher magnesium concentration required for optimal synthesis suggests that incorporation is not all due to residual formylation. When N-formyl-methionyl-sRNA was added to the system in place of N^5-formyl-THFA, identical results were obtained. It thus appears that initiation of synthesis with N-formyl-methionyl-sRNA has a critical magnesium optimum, and at the higher levels of magnesium where synthesis is maximal without the formyl donor, N-formyl-methionyl-sRNA is not an effective initiator.

Preliminary analysis of the products synthesized at 0.010 *M* magnesium with and without the formyl donor was carried out using the sucrose gradient sedimentation assay.[15] The results are presented in Figures 3 and 4. The major peak in the optical density profile is presumably due to the 70*S* ribosomes, and the other two are due to the 50*S* and 30*S* particles. The association of a significant amount of radioactivity with the viral RNA seen in the 30*S* peak suggests that viral proteins were probably synthesized in both cases. The technique was not reliable, however, when used to analyze the products synthesized at the higher magnesium concentration optimal for synthesis without the formyl donor. Evidently, the higher level of magnesium interferes with the association of the protein and the RNA. Further analysis is being carried out with other techniques.

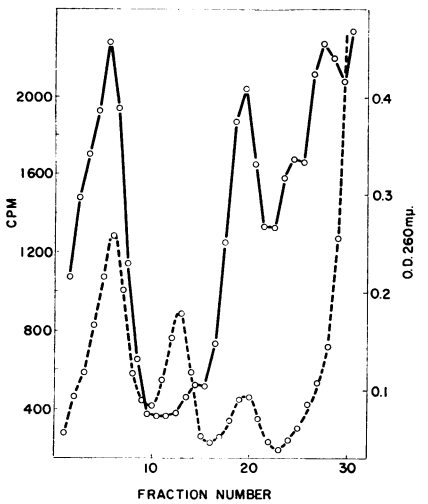

FIG. 3.—Sucrose gradient sedimentation assay of the product synthesized in the presence of N^5-formyl-THFA.

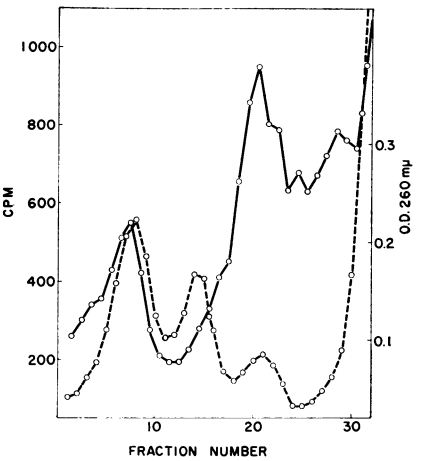

FIG. 4.—Sucrose gradient sedimentation assay of the product synthesized in the absence of formyl donor. Conditions were identical to those given below, except that N^5-formyl-THFA was omitted.

FIGS. 3 AND 4.—The reaction mixture was essentially as described in *Methods*, except that 0.05 ml of S-30, 100 μg of MS2 RNA, and C^{14}-leucine and C^{14}-lysine, specific activity 231 μc/μmole, were added. Of the reaction mixture 0.2 ml was layered onto a 5–20% sucrose gradient and centrifuged for $2^1/_2$ hr at 38,000 rpm in a Spinco SW 39 rotor. Fractions collected were assayed for radioactivity by the usual procedure. Radioactivity is represented by the solid line, and optical density by broken lines.

The incorporation of histidine was of special interest since it is not a constituent of the viral coat protein but is in at least one of the proteins specified by the other cistrons of the viral genome. Ohtaka and Spiegelman showed that histidine was incorporated some time after the onset of valine incorporation.[16] We have carried out a similar kind of experiment at a low magnesium concentration in the presence of a formyl donor. The results are presented in Figure 5. At 0.0075 *M* magnesium. even leucine incorporation is characterized by a long lag, but the incorporation of histidine clearly begins much later.

The synthesis of histidine-containing protein(s) with and without formyl donor at different levels of magnesium shows an interesting difference from the synthesis of proteins containing leucine. The upper curve in Figure 6 represents incorporation of histidine in the presence of N^5-formyl-THFA. The pattern is quite similar to that of leucine incorporation. The peak of maximal incorporation in the absence of the formyl donor, however, is more conspicuously displaced from the peak with the formyl donor. The difference in the leucine- and histidine-incorporation profiles may reflect a difference in the initiating codons of the cistrons. The R17 RNA-directed histidine incorporation, seen in Figure 7, is characterized by a similar displacement of the two peaks. In this particular experiment, RNA from the amber mutant R17 *am*11B with a nonsense mutation in the coat protein region of the RNA, was used as messenger. Although from this experiment it appears that protein synthesis with R17 RNA is more dependent on formylation than is that with MS2 RNA, more experiments will have to be carried out before any conclusions

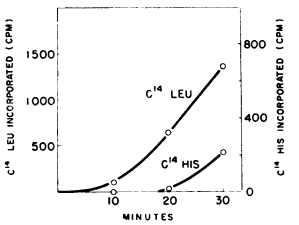

FIG. 5.—Kinetics of leucine and histidine incorporation. Twenty-five μg MS2 RNA, 30 μg N^5-formyl-THFA, and $MgCl_2$ at a final concentration of 0.0075 *M* were added to the basic reaction mixture. C^{14}-leucine of specific activity 231 μc/μmole was added to one set of experiments, and C^{14}-histidine of the same specific activity was added to the second.

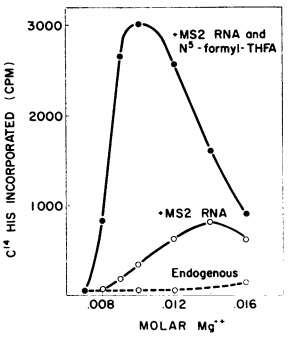

FIG. 6.—Effect of formylation on histidine incorporation with MS2 RNA. The reaction mixture was similar to that given in Fig. 2, except that C^{14}-histidine of specific activity 231 μc/μmole was the labeled amino acid and only 10 μg of MS2 RNA were added where indicated.

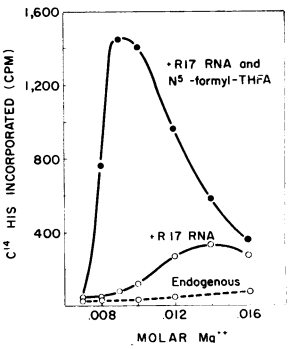

FIG. 7.—Effect of formylation on histidine incorporation with R17 RNA. Conditions were identical to that described in Fig. 6, except that 25 μg of R17 *am*11B RNA were added in place of MS2 RNA.

on the differences can be reached since the results are very much dependent on the S-30 preparation and on the batch of N^5-formyl-THFA.

We have also studied the effect of formylation on protein synthesis directed by TYMV-RNA, a plant viral RNA, and have been able to show a similar dependence of synthesis on formylation.[17] Our efforts to demonstrate either formylation or acetylation of methionyl-sRNA in rat liver extracts have not been successful so far. However, we have found that the *E. coli*-formylating enzyme is capable of formylating methionyl-sRNA from rat liver.

Discussion.—Thus far, the ~~the~~ more essential features of our hypothesis have been supported, for the most part, by the results of studies reported since we first presented our proposal for the mechanism of chain initiation.[18–20] Sundararajan and Thach, after showing that the AUG codon promoted the in-phase reading of adjacent 3′ codons, concluded that the AUG codon fixed the point of initiation by virtue of its innately high activity in directing the binding of N-formylmethionyl-sRNA to ribosomes.[21] However, since the sRNA preparations used by these investigators always contained N-formylmethionyl-sRNA as well as other aminoacyl-sRNA's, a factor which should also be considered is the formation of a more stable complex by the binding of a second aminoacyl-sRNA after N-formylmethionyl-sRNA is bound to the first ribosomal site. Clark and Marcker have reported that the methionine from the methionyl-sRNA is incorporated into the N-terminus of viral protein whether formylated or not, and concluded that a special property of the sRNA was responsible for the ability of the methionyl-sRNA to initiate synthesis.[22] Bretscher and Marcker have shown that the methionyl-sRNA which can be formylated reacts with puromycin in the presence of ribosomes and the AUG triplet, and have proposed a mechanism for chain initiation basically similar to the one we proposed.[23, 24]

Our latest observations on the synthesis of proteins directed by viral RNA's are also in accord with the basic features of the mechanism we have proposed for chain initiation. In contrast to Clark and Marcker, we have found a strong dependence of viral protein synthesis on formylation. At low concentrations of magnesium, amino acid incorporation was completely dependent on formylation, and at a magnesium concentration optimal for synthesis, we have seen as much as tenfold stimulation of leucine and 28-fold stimulation of histidine incorporation by formylation. It is of interest that amino acid incorporation with viral RNA is somewhat similar to that with synthetic RNA polymers in that a higher concentration of magnesium is required for optimal synthesis without a formyl donor. Presumably, in both cases the higher magnesium concentration is required to make nonformylated aminoacyl-sRNA's more effective in initiating synthesis. The data of Clark and Marcker, however, would suggest that with MS2 RNA only the nonformylated methionyl-sRNA which normally can be formylated initiates synthesis.

The dependence of the TYMV-RNA-directed incorporation of amino acids on formylation is interesting. Since TYMV, a plant virus, replicates by utilizing the protein-synthesizing system of the plant host, one would suspect that the virus, like the RNA viruses which infect *E. coli*, would have an initiating codon similar to that of the host. An analysis of the products synthesized with the TYMV-RNA should establish whether the initiating codon for the virus is the triplet for N-formylmethionine or whether a triplet normally not the initiating signal is read, possibly out of phase, as an initiating codon.

Although most of the recent observations have supported the more essential features of our hypothesis, some recent developments have also indicated a need for further qualification and amplification of the hypothesis. We have postulated that initiation of protein synthesis was regulated in the alignment of two aminoacyl-sRNA's on the ribosome since the fulfillment of this condition was necessary for the formation of the first peptide bond. This process of alignment must obviously involve the binding of the aminoacyl-sRNA's onto the ribosome. However, even if two aminoacyl-sRNA's were aligned on the ribosome, formation of the peptide bond would still not be possible if the amino acids were not properly oriented. Clearly then, for the proper alignment of the two aminoacyl-sRNA's, the aminoacyl-sRNA's must be bound to the ribosomes with the reacting groups of the amino acids also properly aligned.

We have previously emphasized the role of formylation in the alignment of the aminoacyl-sRNA's and have implied that formylation was the primary factor responsible for the accommodation of N-formylmethionyl-sRNA in the first ribosomal site. However, the results of Clark and Marcker and of Sundararajan and Thach indicate that other factors are also involved.[12, 21, 22] The former group has stressed the importance of a special property of the sRNA in the effective binding of the N-formylmethionyl-sRNA to the ribosome, and the latter has emphasized the importance of the activity of the AUG codon. Another factor which should not be ignored in the alignment of the aminoacyl-sRNA's is the additional stability provided by the subsequent binding of a second aminoacyl-sRNA. It is conceivable that all of these factors are important in the complex interaction of the ribosome, messenger, and the sRNA. Even in this view, a particular function may be as-

cribed to formylation in the alignment of the carboxyl group of the activated methionine.

The recent studies which suggest that the codon for N-formylmethionine is also read as an internal codon imply that some codons for N-formylmethionine may not initiate synthesis.[12, 19, 20] If this is also true in the living cell, then the data must be reconciled with the proposal that a single triplet signals the initiation of protein synthesis. A reasonable explanation is that secondary structure in many regions of the RNA restricts interaction with ribosomes. A similar explanation has been suggested by Sundararajan and Thach.[21]

Finally, it should be mentioned that although two ribosomal sites of different specificity were postulated in the hypothesis, this is not theoretically a necessary condition. Even with two identical sites, if the activated carboxyl group of the amino acid can be oriented properly only with N-formylmethionyl-sRNA, initiation of synthesis would still be dependent on the formylated methionyl-sRNA.

Summary.—The synthesis of viral proteins in *E. coli* extracts has been shown to be dependent on N-formylmethionyl-sRNA or a formyl donor. At low magnesium concentrations, synthesis was completely dependent on formylation, and at about 0.009 *M* magnesium where synthesis was maximal, stimulations ranging from 10- to 28-fold were obtained by the addition of a formyl donor. The results support the basic features of the mechanism we previously proposed for the initiation of protein synthesis.

The authors gratefully acknowledge helpful suggestions from their colleagues, Drs. H. S. Anker, F. J. Kézdy, and J. Westley in the preparation of this report.

* The project was supported in part by the Otho S. A. Sprague Memorial Institute grant, the American Cancer Society institutional grant 41-F, and by a grant from the Division of General Medical Sciences of the U.S. Public Health Service.

† University of Chicago predoctoral fellow. This material is taken from a thesis being submitted to the Division of the Biological Sciences, University of Chicago, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

[1] Arlinghaus, R., G. Favelukes, and R. Schweet, *Biochem. Biophys. Res. Commun.*, **11**, 92 (1963).

[2] Nakamoto, T., T. W. Conway, J. E. Allende, G. J. Spyrides, and F. Lipmann, in *Synthesis and Structure of Macromolecules*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 28 (1963), p. 227.

[3] Spyrides, G. J., these PROCEEDINGS, **51**, 1120 (1964).

[4] Nakamoto, T., and F. Lipmann, *Federation Proc.*, **23**, 219 (1964).

[5] Arlinghaus, R., J. Shaeffer, and R. Schweet, these PROCEEDINGS, **51**, 1291 (1964).

[6] Thach, R. E., and T. A. Sundararajan, these PROCEEDINGS, **53**, 1021 (1965).

[7] Marcker, K. A., and F. Sanger, *J. Mol. Biol.*, **8**, 835 (1964).

[8] Clark, B. F. C., and K. A. Marcker, *Nature*, **207**, 1038 (1965).

[9] Adams, J. M., and M. R. Capecchi, these PROCEEDINGS, **55**, 147 (1966).

[10] Webster, R. E., D. L. Engelhardt, and N. D. Zinder, these PROCEEDINGS, **55**, 155 (1966).

[11] Clark, B. F. C., and K. A. Marcker, *J. Mol. Biol.*, **17**, 394 (1966).

[12] Noll, H., *Science*, **151**, 1241 (1966).

[13] Nakamoto, T., and D. Kolakofsky, these PROCEEDINGS, **55**, 606 (1966).

[14] It has since come to our attention that a higher magnesium concentration had been shown to be required for the initial interaction of ribosome with poly U and sRNA than is required in the subsequent polymerization of the phenylalanine [Revel, M., and H. H. Hiatt, *J. Mol. Biol.*, **11**, 467 (1965)].

[15] Capecchi, M. R., and G. Gussin, *Science*, **149**, 417 (1965).

[16] Ohtaka, Y., and S. Spiegelman, *Science*, **142**, 493 (1963).

[17] Kolakofsky, D., R. Haselkorn, and T. Nakamoto, unpublished results.

[18] Capecchi, M., these PROCEEDINGS, **55**, 1517 (1966).

[19] Thach, R. E., K. F. Dewey, J. C. Brown, and P. Doty, *Science*, **153**, 416 (1966).

[20] Khorana, H. G., H. Buchi, H. Ghosh, N. Gupta, T. M. Jacob, H. Kossel, R. Morgan, S. A. Narang, E. Ohtsuka, and R. D. Wells, in *Cold Spring Harbor Symposia on Quantitative Biology* (1966), in press.

[21] Sundararajan, T. A., and R. E. Thach, *J. Mol. Biol.*, **19**, 74 (1966).

[22] Clark, B. F. C., and K. A. Marcker, *Nature*, **211**, 378 (1966).

[23] Bretscher, M. S., and K. A. Marcker, *Nature*, **211**, 380 (1966).

[24] The observation that two factors isolated from *E. coli* ribosomes stimulate amino acid incorporation with viral RNA and AUG-containing RNA polymers is not readily interpretable in terms of our hypothesis (Stanley, W. M., Jr., M. Salas, A. J. Wahba, and S. Ochoa, these PROCEEDINGS, **56**, 290 (1966)).