AMINO TERMINAL PEPTIDES OF RNA PHAGE PROTEINS SYNTHESIZED IN THE CELL FREE SYSTEM.

Mary Osborn* and Klaus Weber,
Biological Laboratories, Harvard University,
Cambridge, Massachusetts.

Harvey F. Lodish,
Department of Biology, Massachusetts Institute of
Technology, Cambridge, Massachusetts.

Received October 6, 1970

Summary. The amino terminal peptides of the proteins coded by the RNA of several RNA phages have been isolated and sequenced. The peptides, labeled with ^{14}C amino acids, were isolated from an E. coli in vitro system primed with bacteriophage RNA, after enzymatic digestion of the in vitro made proteins. The amino terminal peptides were identified by coelectrophoresis with peptides isolated from an identical in vitro system in which only F-met -tRNA was labeled. The identification of the amino acids present in each peptide and their sequence was elucidated relying solely on the label.

We are interested in sequencing proteins by radioactive techniques. Such methods are of advantage because they can be used on small amounts of material and on proteins which are radiochemically but not chemically pure. In this paper we have used such procedures to identify and sequence the amino terminal peptides of some RNA phage proteins synthesized in a cell free system primed with bacteriophage RNA. Our sequence of the amino terminal region of the QB coat protein has been used to determine that the RNA fragment sequenced by Hindley and Staples (1) corresponds to the coat initiation region.

Isolation of Peptides

The method used to isolate the amino terminal peptides of proteins made in an E. coli cell free system primed with bacteriophage RNA is shown in Figures 1 and 2. It takes advantage of the fact that all E. coli and phage proteins made in vitro begin with formylmethionine (2,3), and that deformylation should only alter the mobility of peptides containing formylmethionine. Two parallel incorporation mixtures were used: the first contained ^{35}S formylmethionyl-tRNA as the only source of radioactivity;

---

* Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England.

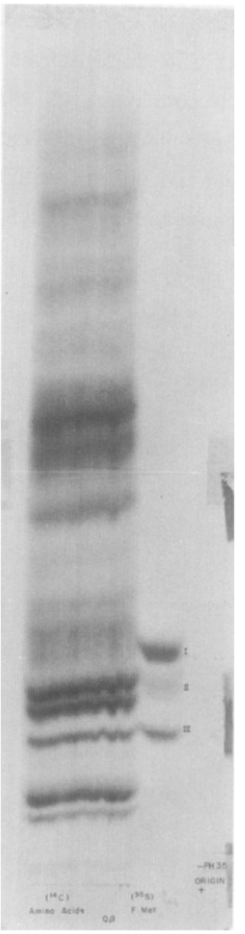

Figure 1: Autoradiograph of radioactive peptides after chymotryptic digestion of cell-free reactions primed with QB RNA. Peptides from the reaction labeled with ^{35}S F-met tRNA are on the right, those from a reaction labeled with a mixture of ^{14}C amino acids are on the left. All reactions contained per ml: 50 μmoles HEPES, pH 7.0, 100 μmoles NH_4Cl, 5 μmoles magnesium acetate, 10 μmoles 2-mercaptoethanol, 2 μmoles ATP, 4.2 μmoles phospheonolpyruvate, 0.3 μmoles GTP, 0.3 ml preincubated S-30 from *E. coli* PR7 (15), and 150 μgm QB RNA. Reactions labeled with ^{35}S F-met tRNA (1.0 ml) also contained per ml: 500 μgm *E. coli* tRNA charged only with ^{35}S N-formylmethionine (5000 mc/mmole) (6) and 0.1 μmoles of each of 20 non-radioactive amino acids. Reactions labeled with ^{14}C amino acids contained per ml: 500 μgm stripped *E. coli* tRNA, 100 μc of a mixture of 15 ^{14}C labeled amino acids (New England Nuclear Corporation NEC-445), 3 μc ^{14}C tryptophan (56 mc/mmole), 4 μc ^{14}C glutamine (95 mc/mmole) and 3 μc ^{14}C asparagine (79 mc/mmole), and 0.1 μmole each of non-radioactive methionine and cysteine. After incubation at 37C for 20 min. the reactions were digested with ribonuclease and precipitated with trichloroacetic acid (16). The precipitates were washed by centrifugation, 3 times with 5% trichloroacetic acid containing 3% casamino acids, twice with ethanol and once with ether. They were resuspended in a volume of 1% NH_4HCO_3 equal to that of the original reaction and digested with α-chymotrypsin (70 μgm/ml reaction) at 37C for 4 hours; occasionally additional enzyme was added after 2 hours digestion. After lyophilization the samples were subjected to paper ionophoresis on Whatman 3MM paper at pH 3.5 (17) for 3 hours at 50 volt/cm. Autoradiography was for 2 days with Kodak Royal Blue X-ray film.

the second contained as their ^{14}C derivatives all the amino acids except methionine and cysteine. Both contained QB RNA as primer; other constituents and the procedures for precipitation and chymotryptic digestion of the mixtures are given in the legend to Figure 1. The ^{35}S and ^{14}C peptides were coelectrophoresed at pH 3.5 and identified by autoradiography.

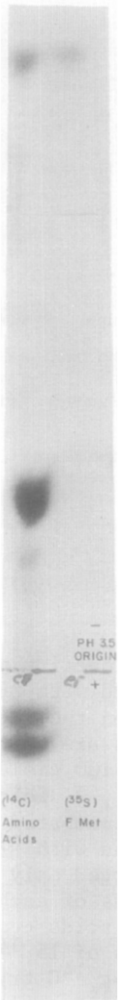

Figure 2: Autoradiograph of peptide III labeled with ^{35}S or ^{14}C, after deformylation. ^{35}S peptide III and the corresponding ^{14}C region (Figure 1) were eluted with 1% triethylamine, lyophilized, dissolved in water and relyophilized. Each was deformylated by treatment with 0.5M HCl (0.3 ml, 20 min at 90C) and then subjected to ionophoresis under conditions identical to those in Figure 1. The deformylated ^{14}C peptide was eluted, lyophilized, and used for the sequencing studies.

With QB RNA as messenger three ^{35}S peptides were detected, (I, II and III in Figure 1). II is from the RNA synthetase since its synthesis was depressed specifically when QB coat protein was added to the reaction (4,5,6). III appeared to contain and to be larger than I, because treatment of III labeled with ^{35}S methionine with excess chymotrypsin, converted all of the ^{35}S radioactivity into material which migrated identically with peptide I. I and III are from the coat protein, which was shown by gel electrophoresis in sodium dodecylsulphate to be the major product of QB RNA in our cell free system. As is the case with f2 RNA (3), QB RNA apparently directed the synthesis of little material which corresponded to A protein. ^{35}S labeled III and the corresponding region of the ^{14}C labeled electropherogram were cut out and each was eluted from the paper, chemically deformylated and rerun at pH 3.5 (Figure 2). After deformylation almost all the ^{35}S radioactivity migrated as a single spot with greatly increased mobility, indicating that it had acquired an additional positive charge. Most of the ^{14}C radioactivity did not increase mobility upon deformylation, presumably because it corresponds to internal peptides of QB proteins. The ^{14}C peptide with mobility identical to that of the ^{35}S deformylated derivative was eluted and used for the sequencing studies.

Sequencing of peptide III from the amino terminus of QB coat protein

The sequence of peptide III was determined using conventional techniques except that amino acids were identified by ^{14}C counts after passage over the columns of a Beckman amino acid analyser. ^{3}H labeled amino acids were added as standards to each sample immediately before the column run. The effluent from the analyser column was collected, each fraction dissolved in Kinard's solution (7) and counted in a dual channel scintillation spectrophotometer. The peptide has the sequence

(F) Met - Ala - Lys - Leu - Glu - (Thr_2 - Val) - Leu

The evidence for this sequence is as follows:

Amino Acid Composition. After acid hydrolysis of peptide III, (6N HCl, 20 hours with added carrier amino acids), the profiles obtained from the long column of the amino acid analyser are shown in Figure 3. The upper profile shows the ^{3}H counts from the amino acid standards. The lower profile shows the ^{14}C counts resulting from hydrolysis of ^{14}C labeled, deformylated, peptide III. Threonine, glutamic acid, alanine, valine and leucine are present. The short column showed only lysine for this peptide. From the known specific activities of the ^{14}C labeled amino acids used for the

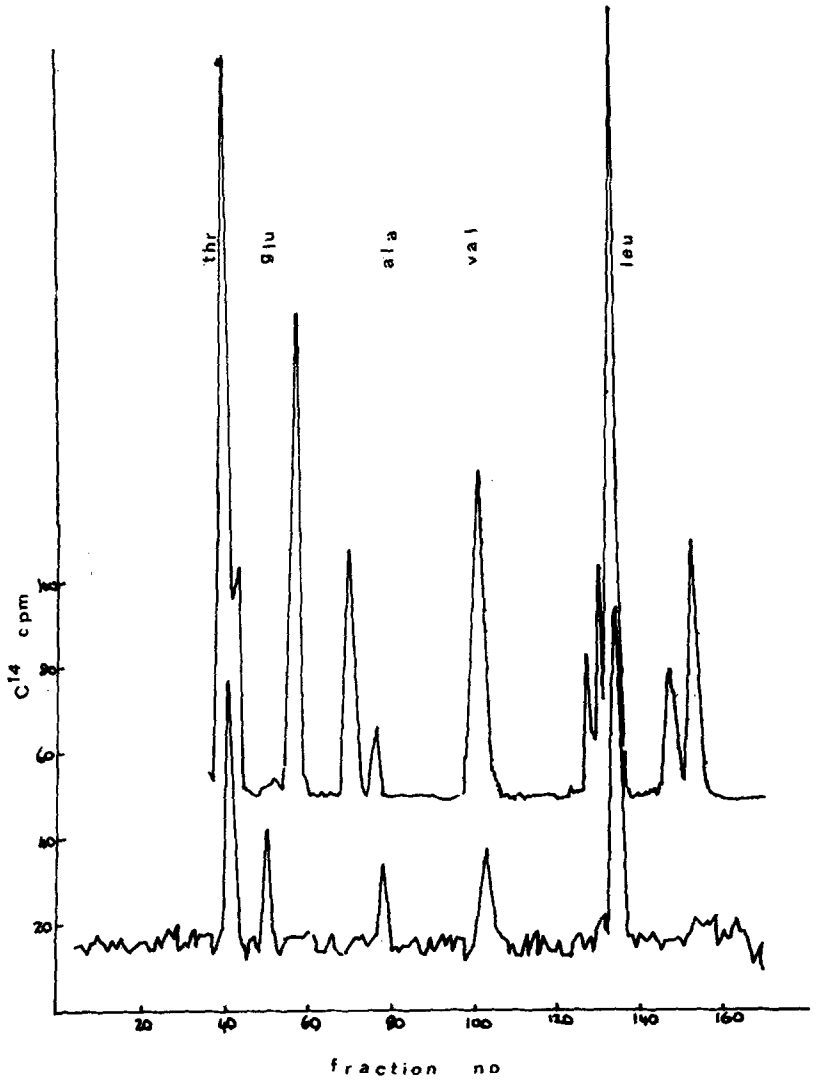

Figure 3: Amino acid composition of peptide III obtained by dual channel counting of the effluent from the long column of the amino acid analyser. The upper profile shows ^{3}H-labeled amino acid standards, and the lower profile the ^{14}C counts resulting from hydrolysis of ^{14}C labeled deformylated peptide III.

---

in vitro reaction (see legend to Figure 1), and from the results in Figure 3, we conclude that the peptide contains two residues of leucine and threonine for one residue of each of the amino acids alanine, valine, glutamic acid and lysine.

^{14}C methionine was not present in the ^{14}C amino acid incorporation mixture. However, methionine is placed at the amino terminus of deformylated peptide III because (a) The purification of this peptide demanded

that it contain an amino terminal formylmethionine residue. (b) When one step of a subtractive Edman degradation was performed on the deformylated ^{14}C peptide no change in the amino acid composition was found. However, a second step yielded amino terminal alanine. (c) When formylated peptide III, labeled with ^{35}S methionine was digested with pronase (as in 6) all the radioactivity migrated with formylmethionylalanine, formylmethionine, or methionine. Hence the amino terminal sequence of III is (F) Met - Ala ...

Carboxypeptidase A. After carboxypeptidase A treatment of deformylated peptide III, (10 µg carboxypeptidase A solution in 100 µl sodium barbitol pH 8.0, 1 hour at room temperature), and passage over the long column of the amino acid analyser, the ^{14}C profile shown in Figure 4 was obtained. Leucine, valine and threonine were the only amino acids released. The ratio of ^{14}C counts in each peak in Figures 3 and 4 shows that 47% of the leucine counts, 70% of the valine counts and 65% of the threonine counts were released. Since chymotrypsin can be assumed to split at leucine rather than threonine, or valine, leucine is the carboxyl terminal residue. Considering the amino acid composition of the peptide, the results suggest $(thr_{1.3}\ val_{0.7})\ leu_{0.94}$, and in particular that more than one threonine residue is released.

Trypsin digestion. Peptide III was digested with trypsin, (2 µg trypsin

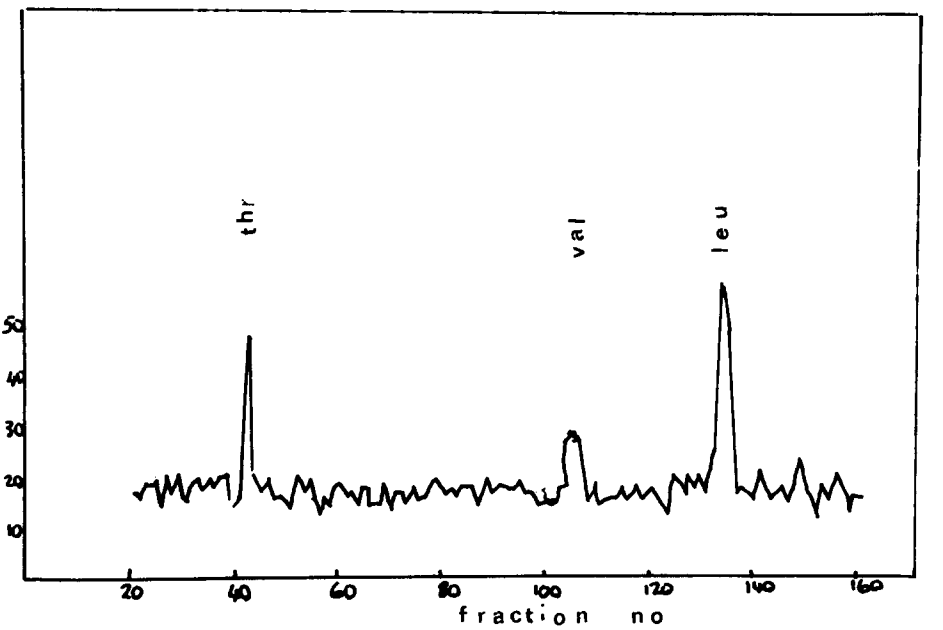

Figure 4: ^{14}C labeled amino acids released after carboxypeptidase A treatment of peptide III. ^{3}H labeled standards are not shown.

in 20 µl 0.5% ammonium bicarbonate, 2 hours at 37C), and the fragments separated by electrophoresis on cellulose thin layer plates, (Brinkman, 2% pyridine in 1% aqueous acetic acid, pH 5.2). Two peaks of radioactivity were detected, T_1 and T_2. The peptides in these peaks were eluted, and their amino acid composition determined. The slower moving acidic fragment, T_1, contained thr_2, glx, val, leu_2. The electrophoretic mobility suggested that it contained a glutamic acid rather than a glutamine residue. The faster moving peptide, T_2, contained only ^{14}C lysine and ^{14}C alanine. Hence, including the results of the previous paragraph, the sequence of III must be:

$\longleftarrow T_2 \longrightarrow$ $\longleftarrow T_1 \longrightarrow$

(F) Met - Ala - Lys.....(Leu Glu) - (Thr_2 Val) - Leu

Edman degradation on the tryptic fragment T_1. Leucine was identified as the amino terminal residue by the subtractive Edman procedure of Konigsberg and Hill (8). Carrier was added, and after the cyclisation step the residual peptide was purified over Dowex 50 H^+. Amino acid analysis of this peptide showed that the number of leucine residues in relation to the other amino acids present was reduced from 2 to 1.1, suggesting that leucine is the amino terminal residue of T_1. The next step of subtractive Edman yielded glutamic acid as the second residue.

These data allows the sequence to be written as above. Further evidence that this is indeed the amino terminal sequence of QB coat protein comes from a series of Edman degradations on intact QB coat protein, which yielded the sequence ala-lys-leu, (R. Weil, personal communication), and also from the sequence of QB coat protein recently determined by Konigsberg, Maita, Katze and Weber (9). Comparison of the *in vivo* and *in vitro* amino terminal sequences suggest that as it is the case for other RNA phages, the amino terminal formylmethionine is removed *in vivo*.

Other peptides isolated from reactions primed with QB, f2 and R17 RNA.

The sequences of the other two QB peptides has not been determined unambiguously because of the low number of ^{14}C counts isolated. Peptide I is not pure, but it does contain as a major component the fragment F-met - ala - lys - leu from the coat protein.

^{35}S and ^{14}C amino terminal peptides were also isolated from *in vitro* reactions primed with f2 and R17 RNA. For f2, after tryptic and chymotryptic digestion of the *in vitro* products, we obtained and sequenced the peptide F-met - ala - ser - asn - phe, which is from the coat protein (10). From R17 reactions, we obtained the same peptide, and also the peptide F-met - ser - lys. The first of these peptides, which was obtained in high yield, is known to correspond to the amino terminal region of the R17 coat

protein (11). The other peptide is known from data of Lodish (6), and the RNA sequence of Steitz (12), to correspond to the amino terminal region of the synthetase protein.

Discussion

The peptides described in this paper are from proteins made in an E. coli in vitro system primed with bacteriophage RNA, and have been sequenced relying only on the radioactive label. The method of isolation demands that they are amino terminal peptides. Such peptides are of interest because of the recent work of Hindley and Staples (1) and of Steitz (12), in determining the RNA sequence of the ribosomal binding sites for QB and R17 respectively. Thus our data on peptide III from the cell free system primed with QB RNA confirms that the ribosomal binding site sequenced by Hindley and Staples (1) corresponds to the coat initiation site. A similar conclusion can be drawn from the QB coat sequence recently determined by Konigsberg et al. (9). For R17 Steitz was able to assign each of her three binding sites to a particular phage protein from the known coat protein sequence (11), and from other less direct evidence. The F-met - ser - lys sequence for the synthetase protein described in this paper lends support to her RNA sequence, and is important because as yet the synthetase protein cannot be purified from R17 infected cells, and thus the amino terminal sequence cannot be determined directly. A longer peptide might be isolated and sequenced by using an enzyme other than trypsin for the initial digestion, or by modification of the lysine residues before digestion. For the A-protein, the in vitro approach is difficult because of the very low level of A-protein made in the in vitro system. Since there are now methods to purify this protein from the phage (13) in large quantities (14) it is probably better to determine its amino terminal sequence directly.

Acknowledgement. This research was supported by grants from the National Institutes of Health, and the American Cancer Society, and by the award of a Damon Runyon Cancer Research Fellowship to M.O. We thank Jim Watson for encouragement, and Charlotte Hering and Elaine Holgate for expert assistance.

References.

1. Hindley, J., and Staples, D.H., (1969) Nature 224, 964
2. Lengyel, P., and Soll, D., (1969) Bact. Rev. 33, 264
3. Lodish, H.F., and Robertson, H.D., (1969) J. Mol. Biol. 45, 9
4. Sugiyama, T., and Nakada, D., (1968) J. Mol. Biol. 31, 431
5. Eggen, K., and Nathans, D., (1969) J. Mol. Biol. 39, 279
6. Lodish, H.F., (1968) Nature 220, 345
7. Kinard, F.E., (1957) Rev. Sci. Instr. 28, 293

8. Konigsberg, W., and Hill, R.J., (1962) J. Biol. Chem. 237, 2547
9. Konigsberg, W., Maita, T., Katze, J., and Weber, K., (1970) Nature, 227, 271
10. Weber, K., and Konigsberg, W., (1967) J. Biol. Chem. 242, 3563
11. Weber, K., (1967) Biochemistry 6, 3144
12. Steitz, J.A., (1969) Nature 224, 957
13. Steitz, J.A., (1968) J. Mol. Biol. 33, 937
14. Osborn, M., Weiner, A.M., and Weber, K., European J. Biochem. In press.
15. Lodish, H.F., (1970) J. Mol. Biol. 50, 689
16. Lodish, H.F., (1968) J. Mol. Biol. 32, 681
17. Sanger, F., Brownlee, G.G., and Barrell, B.G., (1965) J. Mol. Biol. 13, 373