Sequence and Symmetry in Ribosome Binding Sites of Bacteriophage f1 RNA

GEORGE FIECZKIE, PETER MODEL AND HUGH D. ROBERTSON
The Rockefeller University
New York, N.Y. 10021, U.S.A.

(Received 6 May 1974)

RNA was synthesized in vitro from a32P-labeled monomeric triphosphates with Esherichia coli RNA polymerase from covalently closed, circular, double-stranded DNA isolated from cells infected with bacteriophage f1. This RNA, which serves as an efficient messenger in vivo, was bound to ribosomes and the initiation complex was digested with trypsin to obtain ribosome-protected fragments. Ribosome-protected fragments were isolated, purified and separated by two-dimensional analysis using electrophoresis and chromatography. Sequence analysis, including eluciation of the ability to determine relative brightness, was done by conventional techniques.

The sequence of the ribosome-protected fragments was found to fall into three classes. One sequence corresponds to the amino-terminal region of the protein product of f1 gene F, a D3A binding protein. It is proposed that a second sequence may correspond to the amino-terminal region of a precursor to the major coat protein. No assignment has yet been made for the third sequence.

Comparisons are made between these three sequences and others that are available, both in terms of sequence features that have been pointed out earlier, and in terms of certain considerations of symmetry and syntax prominent in these binding site sequences that have not been discussed before.

I. Introduction

Progress is being made toward an understanding of the biology of bacteriophage f1 and its relatives f and M13. These phages constitute an ideal system because several important areas of their molecular biology are being studied simultaneously. In particular, genetic mapping (Lyon & Zinder, 1972), restriction enzyme mapping (K. Horiiuchi, G. Verin & N. D. Zinder, unpublished results) and studies on the mechanisms of genetic recombination (Boon & Zinder, 1971; Hartman & Zinder, 1974, unpublished) are being carried out. In addition, determination of the mode of DNA synthesis in vivo and in vitro is being pursued in a number of laboratories (Wickner & Kornberg, 1973; Wickner et al., 1974). At the same time, efforts are being made to

The term "syntax" is used here to mean a connected or orderly system, a homocentric arrangement of parts or elements, in many cases involving nucleic acids, syntax would be detectable only following determination of the secondary structures of the nucleotides in question. The examples given are meant only as an example of such an arrangement in the genetic code. Less obvious examples of "syntaxal rules" might be group of nucleotides so arranged as to give rise to the "internal terminators" or "5S palindrome" to be discussed here. Fierczkiewicz (1973 and unpublished results) has more comprehensively considered the implications of the existence of "syntax and splice."
understood the mechanism by which transcription and translation are regulated. 

(1)...and L. D. Robertson...

Preparation and fractionation of f1 in vitro reactions

(b) Materials and Methods

Ultrasound nuclease was purchased from Sigma (Mann). Tris base (Tris) was from Difco (Baltimore). NAD+ was from Boehringer Mannheim. cDNA was purchased from Life Technologies. dNTPs were from Promega. Restriction enzymes were from New England Biolabs. 

RNase protection experiments were performed as described previously. 

Ribozyme binding and protection of RNA

MONOCHROME-1 RIBOSOME BINDING SITES

1. Introduction

Ribozyme-binding sites in vitro reactions were carried out under conditions used by Robertson et al. (1973). Incubation was at 37°C in a volume of 0.1 ml containing

Abbreviations used: f1 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion. 1/2 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion, followed by sonication to a molecular size of 1.5 kbp.

(c) Ribozyme binding and protection of RNA

Binding reactions with isolated f1 DNA were carried out under conditions used by Robertson et al. (1973). Incubation was at 37°C in a volume of 0.1 ml containing

1. Introduction

Ribozyme-binding sites in vitro reactions were carried out under conditions used by Robertson et al. (1973). Incubation was at 37°C in a volume of 0.1 ml containing

Abbreviations used: f1 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion. 1/2 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion, followed by sonication to a molecular size of 1.5 kbp.

(c) Ribozyme binding and protection of RNA

Binding reactions with isolated f1 DNA were carried out under conditions used by Robertson et al. (1973). Incubation was at 37°C in a volume of 0.1 ml containing

Abbreviations used: f1 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion. 1/2 RFL, restriction enzyme digested f1 DNA following BamHI or EcoRI digestion, followed by sonication to a molecular size of 1.5 kbp.
and a 5% solution of yeast RNA that had been hydrolyzed for 10 min in 0.1 M KOH was used to homogenize the preparations were incubated with the shorter oligonucleotides present in these fibers. Two-fifths of the digestion mixture was being carried out on mixed RNA fragments from ribosomes after precipitation with HCl. 15% of the fiber RNA was added to each RNAse T1 or pancreatic RNAase digestion reaction. The RNA was excised from the digestion of separated fragments and pooled after three successive washings in acetone to recover products I, II, or pancreatic RNAase digestion were located and eluted.

A total of four digestion procedures were used with each oligonucleotide obtained RNAse T1, RNAse T2, or pancreatic RNAase digestion. In each case, the conditions of digestion and separation were identical to those described by Barrell (1971), unless noted. Pancreatic RNAase digests of RNAse T2-resistant oligonucleotides were carried out in a volume of 10 μl and exposed to high voltage electrophoresis in one direction on DEAE-paper in pyridine-acetate buffer (pH 6.6) at 450 V/m with a current of 10 μA/m and a power input of 0.5 W. RNAse T2-resistant oligonucleotides were then visualized as described previously (Barrell, 1971), except that the pH 6.6 electrophoresis buffer was made 0.05 M in acetate buffer. Base composition analysis was carried out either by an 18-h digestion of 9×10^7 in a volume of 10 μl and by digestion for 2 h with a mixture of RNAase T2 and pancreatic RNAase (Barrell, 1971), followed by fractionation on paper in pH 3.5 buffer containing 0.003 M EDTA. The uptake of CMC to modify oligonucleotides at residues 5′-G and 5′-U (Gilham, 1962) so that they would be susceptible to pancreatic RNAase digestion only at 3′-residues, was carried out as described from Lumsden, 1973, and Barrell, 1971.

Oligonucleotide obtained following complete pancreatic RNAase digestion of ribosome-protected fragments were exposed to 2 further procedures. RNAse T1 digestion was carried out on 10 μl at a 1 mg/ml solution of RNAse T1 in 0.01 M Tris-HCl (pH 7.6), 0.005 M EDTA, and products were separated by dimensional electrophoresis on DEAE-paper at pH 5.9 as described above. RNAse T2 or alkaline digestion was carried out as described for RNAse T1, oligonucleotides. In certain cases, further fractionation of pancreatic RNAase-resistant oligonucleotides was necessary following 2-dimensional fingerprinting analysis in order to resolve spots that were sequence isomers of each other. Unidimensional electrophoresis in 7% formic acid on DEAE-paper was used to separate bands.

Oligonucleotide products obtained by the above procedures were localized by autoradiography and eluted from DEAE or 3MM paper according to standard techniques (Barrell, 1972) except that 25% of the 3MM paper was used, because the lower viscosity was found to shorten significantly the time required for complete elution. Base composition analysis of all such products were done using alkaline hydrolysis or RNAase T1 digestion followed by electrophoresis at pH 5.9 on 3MM paper as described above.

In one case it was necessary to carry out partial elution phosphatidylsphingosine digestion to characterize oligonucleotides (Barrell, 1971; Ligg, 1972). The enzyme was suspended at a concentration of 1 mg/ml in 0.1 M acetate-acetic acid (pH 5.6), 0.005 M EDTA and 0.001 M (Tev) Tween 80 (Atlas Chemical Industries). Partial digestion of oligonucleotides eluted through the phosphatidylsphingosine digestion plate was achieved by resuspending the dried RNA in 10 μl of enzyme in a capillary tube with a drawn-out tip. This tube was incubated at 37°C and 1 ml portions were removed periodically. Reactions were spotted on 1-mm wide strips of DEAE-cellulose plate, which was exposed to homochromatography, as described for RNAse T1 or pancreatic RNAase digests, using homochromatography mixture c of Brownlee & Sanger (1969). Once an optimum time for digestion had been determined based on results of three unidimensional analyses, digestion was carried out in RNA solutions containing 3 μl enzyme, which were incubated and spotted near one end of a strip of cellulose acetate and exposed to 2-dimensional fingerprinting analysis, as used above for separation of RNAse T1 and pancreatic RNAase digests.

(g) Rationale for sequencing these RNA species

The basic approach, described in previous sections, was used in part for 4 separate preparations of the same RNA fragments, each labeled with 1 of the 1
3. Results

(a) Isolation and characterization of $\Phi$1 RNA fragments protected by ribosomes from pancreatic RNase digestion

Figure 1 shows that about $1\%$ of $\Phi$1 messenger RNA is found associated with ribosomes after treatment of initiation complexes with pancreatic RNase. Furthermore, this association is dependent on the presence of N-formylmethionyl-tRNA. In this respect, $\Phi$1 mRNA behaves like RNA bacteriophage $\Phi$2 RNA (Kochva et al., 1968; Steitz, 1969; Hindley & Staples, 1960), single-stranded $\Phi$1174 and $\Phi$1 RNA (Robertson et al., 1972). RNA synthesized according to the methods described in Materials and Methods, section (c), has been shown to act as template for phage-specific protein synthesis in vivo (Chan, Model & Zinder, unpublished results) and, in particular, a portion of one of the RNA preparations subsequently used for isolation of ribosome-protected fragments was tested and found to direct synthesis of the weak mRNA of $\Phi$1-specific proteins (Model & Zinder, 1974).

Plate I shows two-dimensional analyses of RNase $T_4$ digests (carried out as described in Materials and Methods) of both intact $\Phi$1 mRNA preparations and the ribosome-protected portion isolated from the $70 S$ region of a sucrose gradient such as that shown in Figure 1. The complexity of the unfractrated $\Phi$1 RNA (Plate I(a)) is too great to be expected from a mRNA or a precursor of this size (Jefferson, 1971; Böhm & Jöppich, 1972). It is evident that ribosome protection results in a great simplification of this pattern, suggesting that only a few specific regions are being recovered from the ribosomes after digestion of the integral complex. A pattern similar to that seen in Plate I is also observed when the protection reaction is carried out either in the presence of 10 mM or 5 mM magnesium acetate; or when the ribosomes used were derived from E. coli infected cells (data not shown).

As will be shown below, each of the three predominant spots in a diagonal array in the lower portion of Plate I(b) represents a $T_4$-resistant oligonucleotide that terminates in $\hat{A}U\hat{G}$. In addition, each of these sequences contains the potential initiator codon from a different protected RNA fragment.

Fractionation of the ribosome-protected RNA fragments by two-dimensional analysis (Plate II) leads to the resolution of the mixture into three principal protected sequences (Plate II). In particular, Plate II shows six to seven prominent oligonucleotide spots that have migrated in the second dimension (thin-layer homochromatography with an unhydrolyzed yeast RNA mixture) in the size class of 30 to 40 bases. Each of the principal oligonucleotides from such separations was eluted, digested with RNase $T_4$, and subjected to fingerprinting analysis, as shown in Plate III. Three major classes of simplified patterns were observed, each containing one of the three large $T_4$ products from Plate I(b) referred to above. Fragments derived from a common ribosome binding site are indicated in Plate II. Sequence category I contains three prominent oligonucleotides that yield closely related RNase $T_4$ fingerprints. As Steitz (1969) has shown, pancreatic RNase $\Phi$1 RNA often produces frayed ends in primosome-rich regions adjacent to ribosome-binding sites, and this effect accounts for the multiple peaks observed in the RNase-protected RNA region observed in Plate II.
### Table 1

**Analysis of RNase T-resistant oligonucleotides from ribosome binding sites**

<table>
<thead>
<tr>
<th>Oligonucleotide number</th>
<th>Proposed sequence</th>
<th>NTP Composition</th>
<th>Pancreatic RNase product</th>
<th>RNase U1 product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G(A)</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>U</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>U/C/A/G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>2</td>
<td>U/C/C/C/U(A)</td>
<td>U</td>
<td>U</td>
<td>(U/C)/U</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>(U/C)/U</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>(U/C)/U</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>(U/C)/U</td>
</tr>
<tr>
<td></td>
<td>U/C/A/G</td>
<td>U/C/C/C/U</td>
<td>U/C/C/C/U</td>
<td>U/C/C/C/U</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>3</td>
<td>U/U/J/A/A-U/O/G</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>U/C/A/G</td>
<td>U/C/C/C/U</td>
<td>U/C/C/C/U</td>
<td>U/C/C/C/U</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>4</td>
<td>A/A/A/A-G(U)</td>
<td>U</td>
<td>A/A/A/A-G(U)</td>
<td>A/A/A/A-G(U)</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>U/C/A/G</td>
<td>A/A/A/A-G(U)</td>
<td>A/A/A/A-G(U)</td>
<td>A/A/A/A-G(U)</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>U/C/A/G</td>
<td>A/A-C-U/O/A/C/A-U-G(A)</td>
<td>A/A-C-U/O/A/C/A-U-G(A)</td>
<td>A/A-C-U/O/A/C/A-U-G(A)</td>
</tr>
</tbody>
</table>

*Notes:*
- NTP Composition refers to the type of nucleotide present in the oligonucleotide sequence.
- Pancreatic RNase and RNase U1 products represent the products generated by digestion with these enzymes, respectively.
PLATE IV. Analysis of partial spleen phosphodiesterase digests of A-rex RNAase T<sub>1</sub>-treated oligonucleotides from 1<sub>1</sub> ribonuclease-susceptible fragments. The spots numbered 4 and 14 in Table I were absent from RNAase T<sub>1</sub> fingerprints of purified bands nos. 3 and 6 such as those shown in Plate II(a) and (b). Except that the 11-mers with which these oligonucleotides were incubated had been synthesized in a reaction containing [32P]ATP as the only radioactive precursor. Partial digestion with spleen phosphodiesterase was carried out as detailed in Materials and Methods, section (f).

(a) Unidimensional analysis of partial digests. Solutions containing the two oligonucleotides were incubated, and the mixtures spotted periodically near the bottom of a DEAE-cellulose thin layer plate. At the end of the incubations, the plate was washed with 8 min in 8% trichloroacetic acid, dried and subjected to conventional autoradiography using the RNA mixture employed in Plate I(b).

Sample lane I, spot 14 (30 x 0.5 mm) from Table I, incubated in 10 µl tritiated trichloroacetic acid in 0.1 M Tris-HCl, pH 7.4, 0.001 M EDTA and 50 µg/ml spleen phosphodiesterase for 1 h at 37°C under nitrogen. Lane II, spot 14 incubated in 10 µl tritiated trichloroacetic acid in 0.1 M Tris-HCl, pH 7.4, 0.001 M EDTA and 50 µg/ml spleen phosphodiesterase for 1 h at 37°C under nitrogen.

(b) A reaction with spot 4 from Table I was incubated for 60 min in a volume of 5 µl and spotted near one end of a cellulose acetate strip. Autoradiographic analysis was carried out as described in Materials and Methods, section (f).

PLATE V. Pancreatic RNAase fingerprints of the three principal ribonuclease-susceptible fragments from 1<sub>1</sub> mRNAs. The same RNA fragments analyzed in Plate III were subjected to pancreatic RNAase digests after isolation from the preparative DEAE-cellulose thin layer plate. Incubations were in drawn-out capillary tubes containing the RNA dissolved in 2 µl of 0.1 M Tris-HCl, pH 7.4, 0.001 M EDTA and were carried out at 37°C for 30 min. Fingerprinting conditions were as in the legend to Plate I(b). (a) Pancreatic RNAase fingerprint of the largest fragment of fragment no. 1 from Plate III: (b) drawing of (a), indicating sequences deduced for each of the pancreatic RNAase-resistant oligonucleotides present (see Table 7); (c) pancreatic RNAase fingerprint of fragment no. 1 from Plate III; (d) drawing of (c); (e) pancreatic RNAase fingerprint of fragment no. 2 from Plate III; (f) drawing of (e). Rhythms have been omitted from the sequences to conserve space.
### Table 1—continued

<table>
<thead>
<tr>
<th>NTP*</th>
<th>Composition*</th>
<th>RNase T1 products*</th>
<th>RNase U2 products*</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>5'-dU</td>
<td>5'-dA</td>
<td>5'-dG</td>
</tr>
<tr>
<td>C</td>
<td>5'-dC</td>
<td>5'-dG</td>
<td>5'-dA</td>
</tr>
<tr>
<td>A</td>
<td>5'-dA</td>
<td>5'-dG</td>
<td>5'-dA</td>
</tr>
<tr>
<td>G</td>
<td>5'-dG</td>
<td>5'-dA</td>
<td>5'-dA</td>
</tr>
</tbody>
</table>

* NTP refers to the ribonucleoside (dNMP) triphosphate used to make the oligonucleotides analyzed in a given line of the Table.

1. *These compositions were determined as described in Materials and Methods. A double dash in this and other columns signifies that the particular oligomer could not be tested as an RNase T1 or U2 product, due to the absence of such banding in ultraviolet shadowed gels made under the conditions described. Heredity bands indicate the proposed nearest-neighbor base pair of the oligonucleotides analyzed (see Materials and Methods, section g). The parentheses within a particular RNase T1 or RNase U2 product indicate that the sequence oligomer is known to be only part of the entire oligonucleotide in question.

2. The other two buffer volumes of binding site no. 1 (Table II) are altered in this 1' terminus sequence in that one is mixing terminal U residues and the other lacks U-U-C-A-A-U.

3. These sequences were derived by partial phenol/phenolase digestion in addition to the determinations shown here (see Table IV).
(b) Determination of the sequence of the T₄-resistant oligonucleotides from ribosome-protected fragments

Plate III(a), (c) and (e) represents the RNase T₄ fingerprints of the three prominent sequence. The patterns represented here are those obtained after fractionating oligonucleotides prepared from RNA made from all four ribosome-related t₄T₄(1-125) phosphates. When ribosome binding sites synthesized using only one labeled ribonucleoside triphosphate were fractionated in this way, the resulting T₄ fingerprints represented only the spots that either contained (at a position other than at the B' end) the base whose precursor had been labeled and/or that had been immediately adjacent to that B' end. (i.e. as a nearest neighbor.)

RNase T₄-resistant oligonucleotides was cloned from the various fingerprints and subjected to further analysis, as described in Materials and Methods. Portions were digested with pancreatic RNase, RNase U₇ and their base compositions determined by digestion with alkali in RNase T₄. Table I shows such results for oligonucleotides from the RNase T₄ fingerprints represented in Plate III. For each of the spots, except those indicated in Table I, a sequence is proposed and its location indicated in the drawings in Figure 1(b), (g) and (i). T₄-resistant oligonucleotides no. 4 and no. 14 of Table I have the sequence (A₆)₅. From mobility on homochromatography we can estimate that a equals 5 to 8. Also, oligonucleotide no. 4 is one A residue shorter than no. 14. Plate IV(a) shows unidimensional analysis using homochromatography on DEAE-cellulose thin-layer plates of partial spleen phosphodiesterase digests of oligonucleotides no. 4 and no. 14 prepared using [3H]ATP. Since this enzyme works by cleaving mononucleotides 3'-monophosphates from the 5'-terminus of RNA chains (Bernardi & Bernardi, 1969), we expect that the smallest oligonucleotide to be labeled in such a digest will be A₆G. It is evident in Plate IV(b) that adenine 3'-monophosphate (Ap) is released as expected. The shortest oligonucleotide observed in four positions above the location of the undigested spot no 14 and three positions up from undigested spot no. 4. This oligonucleotide co-electrophoresed with authentic A₆-G on several systems, suggesting that no. 14 is (A₆)₅ and no. 4 is (A₆)₅.

These assignments are confirmed in Plate IV(b), where two-dimensional analysis of a partial spleen alkaline digest of the putative (A₆)₅ is shown. The successive changes of mobility in the vertical direction rule out improper cleavages, which might have produced oligonucleotides consisting entirely of A residues with a mobility in the first dimension considerably slower than the initial material (Sanger et al., 1973; Ziff et al., 1972). The oligonucleotide spot indicated by the arrow in Plate IV(b) was cloned and its alignment as A₆-G confirmed by co-electrophoresis with authentic A₆-G.

The largest spot in binding site no. 1 (oligonucleotide no. 6, Table I) contains within its 10-base length a run of pyrimidines more than long. In addition, the 3'-terminal sequence of this fragment consists entirely of pyrimidines (Table IA, oligonucleotide no. 2). In order to assign unique sequences to these polyuridylate tracts, panomic RNase digestion in the presence of CMOT labeled U and O residues was carried out using RNA labeled with either ATP, UTP or CTP. The details of these experiments are shown in Table 2. These data allow only one possible sequence (A₅)₅ and the seven-base tract in the 15-base T₄ product, and for the 9 terminal sequence.
(d) Determination of the sequences of the pancreatic RNAase-resistant oligonucleotides from ribosome-binding fragments

The results summarised in Tables I and II, and Table III and V are nearly sufficient by themselves to allow us to propose complete sequences for these three RNAase-resistant fragments. In fact, binding site no. 2 has a single unique sequence that can be derived from the RNAase T₁ oligonucleotides and those produced by pancreatic RNAase cleavage. Determination of sequence no. 3 also requires partial spleen phosphodiesterase digestion (Plate IV) and CMCT-blocked RNAase digestion (Table III). Figure 2 shows the proposed complete sequences for the three RNAase-resistant oligonucleotides. These overlaps between the various RNAase-resistant oligonucleotides can be made uniquely in the case of binding site no. 1. This proposed sequence predicts the observed pancreatic RNAase digestion products and their specific labeling patterns, as detailed in Table II and Plate V. Binding site sequence no. 2 can also be ordered by knowledge of the nearest-neighbor bases of the four T₁ products observed; again, the pancreatic RNAase products fulfill the expected labeling pattern. In the case of binding site no. 2, we require the pancreatic RNAase digestion data in order to assign an order to the six RNAase T₁ digestion products. Examination of the data on which these sequences are assigned will show that much redundancy exists in the assignment of each base to a position in the sequence, and that there are only a few places where the assignment rests on a single experimental result.

BACTERIOPHAGE φ1 RIBOSOME BINDING SITES

Table 3

<table>
<thead>
<tr>
<th>Oligonucleotide number</th>
<th>Proposed sequence</th>
<th>NTP</th>
<th>Pancreatic RNAase-resistant oligonucleotides after CMCT modification</th>
</tr>
</thead>
</table>
4. Discussion

(a) Relation of ribosome-binding site sequences to f1

The preceding data have shown that we can isolate some non-ribosome-protected fragments following pancreatic RNase digestion of initiation complexes containing f1 in vitro messenger RNA. Phage f1 has eight genes (Lyszczak & Zabriskie, 1973) and identification of some of these with the three ribosome-binding sites may help to identify features that regulate f1 protein synthesis in infected cells. We should bear in mind that not all ribosome-binding sites isolated by these methods are necessary in the initiation regions for synthesis of an authentic protein (Stea, 1973) and, furthermore, that isolation procedures depend on RNase resistance of the mRNA-ribosome complex rather than features more directly related to initiating protein synthesis; certain additional important sites may be present in low yield.

In Figure 2 we have written hypothetical amino acid sequences beginning at AUG codons present within the three ribosome-protected fragments. In particular, the amino-terminal sequence predicted for binding site no. 2 coincides with the amino-terminal sequence of the fl gene F protein recently determined by Makishima et al. (1972) (Met-Glu-Asp-Glu-Ala-Asp). Gene F protein is one of the primary products whose synthesis is directed by the phage genome both in infected cells and in vitro. The sequence Ala-Glu-Gly-Asp-Asp (Abbeck et al., 1969; Stell \& Offen, 1970), which corresponds to the amino-terminus of the other prominent fl-directed protein, the major coat protein, is not present in Figure 2.

Evidence to be presented elsewhere suggests, however, that binding site no. 1 may correspond to the initiation site for coat protein synthesis. In particular, Model & Robertson (unpublished results) have isolated a small RNA transcript (less than 1000 nucleotides in length) from f1 mRNA, which yields only binding site no. 1 using the procedures described above. This RNA efficiently directs protein synthesis in vitro to yield greater than 90% coat protein. Furthermore, a second RNA species (less than 2000 bases in length) allows protection of binding sites no. 1 and no. 2 and more directly directs synthesis of coat and gene F protein. In addition to confirming that the three RNA sequences sequenced here are widely spaced on the f1 genome, these findings have led us to examine the possibility that there may be a precursor form of the coat protein beginning with the amino acid sequence predicted by binding site no. 1 (Model & Robertson, unpublished results). Analysis of Met-containing peptides eluted from the region of polycrylamide gels at which coat protein migrates suggests that this sequence may be encoded by a long transcript (unpublished data). Such a hypothetical scheme, together with present efforts to extend the ribosome-protected sequences by using RNase T1, rather than pancreatic RNase, to digest initiation complexes (Gupta, et al., 1973), should allow use to determine whether this hypothesis concerning the synthesis of a precursor F coat protein is correct.

(b) Comparison of f1 ribosome-binding sites with previously determined sequences

All three of the sequences shown in Figure 2 contain an AUG codon with the potential to serve as an initiator signal for protein synthesis. No termination codons appear in phase and to the right of these initiation codons within the RNA sequences. In the case of binding site sequence no. 1, there are two AUG codons present. However, the possibility that the AUG near the 5' end of the fragment could serve as an

293

BACTEROIOPHAGE f1 RIBOSOME BINDING SITES

294

Initiator codon can be eliminated, because it has a UAG termination codon in phase 1, after 5 codons.

(c) Common features to the e site of pneumococci

Previous investigators have concentrated on the regions to the left of the initiator AUG codons in sequences such as those shown in Figure 2, in order to discover signals that allow recognition of these sites by ribosomes, and control their expression. In comparisons, we will limit our considerations to those sequences determined for ribosome-protected fragments isolated from initiation complexes, i.e. sequences from R17 (Stea, 1969); f2 (Gupta et al., 1970) and Q2 (Hindley & Staples, 1969; Porter et al., 1974) RNA phages: fX147 gene 0 (Robertson et al., 1973) and T7 early mRNA (Arndt & Hindley, 1973). One feature observed in all of these sequences (the seven from RNA phages, the two RNA-coded sequences mentioned, and the three illustrated in Fig. 2) is the presence of at least one termination codon to the left of AUG. All three binding sites from f1 mRNA have a UAA codon in this region. In two cases (no. 1 and no. 2) this signal is out of phase with the AUG, while in binding site no. 2, it is in phase. Formally, it would be possible for any one of these UAA codons to represent the termination signal for a previous initiation.

Other additional common features have been commented on by several authors. For instance, a number of binding sites have the feature containing three U residues flanked by two purines in either site (Robertson et al., 1973). The f1 binding sites shown in Figure 2 do not possess this feature. A sequence feature involving four or five purines is present in all but two of the RNA phase ribosome-binding-site sequences (Stea, 1969; Hindley & Staples, 1969; Gupta et al., 1970), the sequence from fX147 gene 0 (Robertson et al., 1973) and the T7 early mRNA site (Arndt & Hindley, 1973). The f1 sites each possess such a feature (no. 1, G-O-A-A; no. 2, A-A-G-G; no. 3, A-A-A-A-A-G-G) but lack sequences in common with the previously published binding site purine tracts. In addition, the three f1 binding sites and that from fX147 possess the sequence U-A-A-U, while T7 contains U-A-A-C, to the left of the AUG codon.

(d) Homologies

Homologies between several pairs of the previously published sequences have been pointed out, for example the Q8 coat ribosome-binding site (Hindley & Staples, 1969) and that from the fX147 gene 0 (Robertson et al., 1973) contain a sequence of 14 bases matching with the initiation codons AUG that have 10 of 11 bases in common (T-T-T-A-A-T-C-A-T-G versus U-U-U-U-U-A-U-C-A-U-G). Also, the T7 early mRNA site (Arndt & Hindley, 1973) and that from the R17 replicase (Stea, 1969) contain the common sequence A-A-A-O-U-U-A-A-A-A-G in three states each (unpublished). With regard to the f1 binding site sequences, site no. 1 contains the sequence U-U-U-U-U-A-U homologous to T-T-T-A-A-T in the fX147 gene 0 site (Robertson et al., 1973). In comparing 11 binding sites nos. 2 and no. 3, we observe no homology shown in Figure 3 where 10 of 17 bases to the left of, and including, the AUG codon are the same. Re-examination of the sequencing data for these two regions reveals the only difference, presence of the U residue 12 bases from the 5' end of the homologous region in site no. 2, leads to major and easily reproducible differences in the pancreatic RNaseA and RNaseA2 digestion products of the relevant RNase T1-resistant oligonucleotides.
Features such as the homology shown in Figure 3 raise the possibility that binding site no. 3 is recognized by ribosomes primarily because of its close similarity to binding site no. 2. Therefore, the attempt to assign a gene to the binding site will be of particular interest. It is possible for RNA regions to be protected by ribosomes from RNase A digestion in an (A+U)-RNA-dependent manner without their acting as initiator regions (Steitl, 1973). Features of such sequences that are recognized by ribosomes will be particularly helpful in defining what it is other than an AUG codon that makes an initiator region functional.

(iv) Hairpin loops

Several initiator regions that have been sequenced can be drawn as potentially base-paired structures (hairpin loops) (Steitl, 1973; Robertson et al., 1973). In particular, it has often been possible to construct such loops with the AUG initiator codon contained in the single-stranded region. We have been able to draw such structures for binding sites no. 1 and no. 2 but their potential stabilities as evaluated by the rules proposed by Thomson et al. (1973) are marginal, and only the potential structure of no. 1 has a negative G+C value at 20°C. In addition, the significance of such a structure is uncertain in the light of recent demonstrations that they are common and expected potential structural features that should be found throughout all RNA molecules (Bralla & Deloll, 1974).

(c) True palindromes and internal terminators

Additional common features of the three Fl ribosome binding sites were sought by comparing the sequences to the right of the AUG initiator codon. No obvious regions of homology are observed (Fig. 2), as might have been expected since each such region is thought to encode a different specific peptide sequence. However, four common striking features are observed within these regions. Figure 4 shows these features within the sequences immediately following the AUG codons. The brackets underlying adjacent sequences in sites no. 1 and no. 2 indicate the occurrence of sequences of nine bases (site no. 1) or 10 bases (site no. 2) in length, which read the same backwards and forwards. Such palindromic sequences will be named in the present discussion "true palindromes," since these exact palindromes have commonly been proposed for self-complementary sequences present in double-stranded DNA (Wilson & Thomas, 1974). In site no. 1, the true palindrome U-G-A-A-A-A-A-G-U overlaps the AUG initiator codon and extends seven bases to the right. In site no. 2, this true palindrome A-U-U-A-A-A-A-G-U-G-A-A-U-A-G-U-A is located immediately adjacent to and to the right of the AUG.
BACTERIOPHAGE T4 BIORIBOSOME BINDING SITES

295

We thank Dr. N. D. Zinder for providing the advice, insight, and organizational framework without which this combined effort would not have been possible. We also thank Drs. Carolyn McMillin and Mr. Edward Felix for their expert assistance with the various aspects of this work. We are grateful to Drs. O. York and R. Doorstein, and Dr. Karen Jakas, for gifts of materials. This research was supported in part by grants from the National Science Foundation to N. D. Zinder and to one of the authors (H. D. R.) and by funds from the Andrew Mellon Foundation.

REFERENCES


1973). For example, in binding site no. 2 four such internal terminators are indicated reading in the 5' to 3' direction. As might be expected in a sequence containing palindrome symmetry, there are also four such codons that are the palindromic equivalent of the three conventional termination codons (UAU becomes AUA, TAT becomes GUA, and TGA becomes AUG). In fact, the 16-base true palindrome in site no. 2 is composed completely of such forward and backward internal terminators. These features are observed to a lesser extent in sites no. 1 and no. 2 (Fig. 4). Inspection of ribosome binding-site sequences from RNA phage shows that five of the six sites from 817 and 826 possess one such internal terminator (Stiles, 1969; Hindley & Staples, 1969), while they are lacking from the 4X174 and T4 ribosome binding sites (Arnold & Hindley, 1972; Robertson et al., 1973).

1973). The presence of such internal terminator sequences at or near the initiator triplets of mRNA for various proteins would be expected, as was the case for true palindromes, to bind the ribosome among the various synonymous codons that might otherwise be utilized to encode these amino-acid terminal peptide regions. In fact, the presence of internal terminators could influence the choice of amino acids within such amino-acid codons. Such an expectation was suggested by Pien, G. (1973), further strengthened by the severe constraints on observable sequences exercised by other internal terminators or true palindromes (as in the case of 11 binding site no. 2) and the apparent negative complementation of these two elements could have a specific role in evolution. These implications will be considered in a separate communication (Pien, G., unpublished results), which examines whether a beginning can be made toward understanding the evolution of the genetic code and the development of rules for the prediction of allowable nucleic acid sequences.
Membrane-bound and Free Polyosomes in Transformed and Untransformed Fibroblast Cells

MARKUS NOLL and MAX M. BURGER
Biologie der Universität Basel, Klingelbergstr. 79, CH-4056 Basel, Switzerland
(Received 26 March 1974)

A rapid and quantitative fractionation procedure has been used to measure the amounts of free and membrane-bound polyribosomes in growing and stationary PyT3 and 3T3 (mouse) cells. A comparison of growing 3T3 and PyT3 cells does not reveal any significant differences with regard to the ratio of the two polyribosome fractions. The amount of free and membrane-bound polyribosomes decreases in both 3T3 and PyT3 cells as they approach the stationary stage, an effect which is much more pronounced for free polyribosomes. A greatly reduced growth rate in stationary cells, however, the amount of membrane-bound polyribosomes decreases in PyT3 cells while it increases even further in 3T3 cells. By contrast, the amount of free polyribosomes remains at a reduced level in both 3T3 and PyT3 cells when cell multiplication is inhibited.

Based on the hypothesis that membrane protrusion are selectively synthesized on membrane-bound polyribosomes, an attempt is made to relate the results to the accumulated data in the literature and discuss its possible significance with respect to the loss of growth control in PyT3 cells.

1. Introduction

Eukaryotic ribosomes have been shown to exist in two distinct topological states, either bound to the endoplasmic reticulum or free in the cytoplasm (Palade, 1965; Palade & Siskovsk, 1956). A functional difference between the two classes is less well established even though it has been suggested by the work of several groups (Frischfield & Wettstein, 1957; Ragnanu & Iselthecher, 1973; Uno & Oso, 1973; Murry & Sidralmsky, 1973). It seems that secretory proteins are mainly synthesized by membrane-bound polyribosomes while cell sap proteins are primarily synthesized on free polyribosomes (Campbell et al., 1969,1969; Siskovsk & Palade, 1960; Takagi & Ogata, 1968; Redman, 1968,1969; Hicks et al., 1969; Takagi et al., 1969,1970; Pino & Ranker, 1971). Furthermore, Webb et al. (1964,1965) reported that in liver cells considerably higher proportions of polyribosomes are bound to the endoplasmic reticulum than in hepatoma cells. It is conceivable that this difference in the membrane-bound/free polyribosome ratio originates from changes in the membrane system during transformation and alters gene expression at the translational level, as suggested by Sins et al. (1969). Thus, the membrane alteration triggers a change of gene expression which could be part of a primary signal inducing events leading to the next mitosis.