Specific Binding of a Chemically Synthesized Prokaryotic Ribosome Recognition Site

PROSPECT FOR MOLECULAR CLONING AND EXPRESSION OF EUKARYOTIC GENES*

(Received for publication, December 5, 1979)

Ernest Jay,+ Arun K. Seth,+ and Gilbert Jay§

From the †University of New Brunswick, Department of Chemistry, Fredericton, New Brunswick E3B 5A3, Canada and the §National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

An icosadeoxyribonucleotide containing the several features found in prokaryotic mRNA ribosome binding sites has been synthesized. This sequence can stimulate the binding of initiator fMet-tRNA, to the ribosome to form a stable 71 S initiation complex identical with those induced by natural messengers. The binding of this synthetic ribosome binding site is absolutely dependent upon initiation factor IF3, and the bound fMet-tRNA is sensitive to puromycin indicating the formation of a functional initiation complex. A heptadeca- deoxyribonucleotide, identical with the cistron but lacking the terminal A-T-G codon, can also stimulate the stable binding of fMet-tRNA, to the ribosome, suggesting that the selection of the proper A-U-G initiation codon by fMet-tRNA is subsequent to and a result of the recognition and binding of the fMet-tRNA, 30 S ribosome complex to the initiation site. The prospect of ligating a similar synthetic ribosome binding site in front of a eukaryotic gene for cloning in an appropriate prokaryotic vector to assure the expression of the protein is discussed.

The rapid advance in the technology of molecular cloning has generated much interest and excitement in the possibility of the synthesis of biologically important mammalian proteins in bacterial cells carrying the appropriate DNA insert. However, while cloning vehicles have been constructed to assure the efficient transcription of eukaryotic genes that have been translated specifically at the correct initiation codons in vitro, the efficient initiation of protein synthesis at the correct eukaryotic initiation codons in vivo to produce the expected proteins.

RESULTS

Features of the Chemically Synthesized Ribosome Binding Site—Analysis of the nucleotide sequences of a large number of prokaryotic ribosome binding sites has revealed several features which might either jointly or separately serve as recognition elements for the initiation of protein synthesis for reviews see Refs. 1 and 2). In view of this, we have synthesized an icosadeoxyribonucleotide, dA-A-T-T:C-T-A-G-G-A-T-T-T-A-A-T-A-T-C-A-T-G, containing all these features for ribosome binding studies. This hypothetical ribosome binding site contains: (i) an A-T-G initiation codon at the 3' end, which will allow interaction with the initiator fMet-tRNA, (ii) a polypurine sequence, A-G-A-A-G, which has been postulated and shown to interact with the polypyrimidine stretch at the 3' terminus of 16 S rRNA (6, 7), (iii) a heptadecanucleotide sequence, G-A-T-T-T-A-A, corresponding to the R-R-U-U-U-R-R sequence which is found in a large number of ribosome binding sites in phage RNAs (1, 2), (iv) a pyrimidine, C, immediately to the 5' side of the A-T-G initiation codon, since not only does C occur most often at this position but tetranucleotide binding studies have also shown that Y-A-U-G binds much better than R-A-U-G (8-10); (v) two termination codons, T-A-G and T-A-A, both to the 5' side of the initiation codon but in different reading frames, to assure that initiation at this site will not be blocked by translating ribosomes; (vi) an endonuclease Eco RI restriction site, A-A-T-T-C, at the extreme 5' end which can be used as an insertion site for molecular cloning. It has been clearly demonstrated that Escherichia coli ribosomes can bind specifically to single-stranded phage DNA at regions corresponding to the initiation sites (11) and initiate protein synthesis through the formation of peptide bonds (12). So single-stranded DNA, like mRNA, can serve as efficient templates for the specific initiation of protein synthesis in vitro.

We have chosen to synthesize for our ribosome binding studies oligodeoxyribonucleotides instead of oligoribonucleotides for the following reasons. (i) The chemical synthesis of DNA fragments is much easier than that of RNA fragments. The difficulty in synthesizing RNA fragments precludes the rapid synthesis of analogues with specific base substitutions, deletions, or insertions, for comparative ribosome binding studies. (ii) The synthetic DNA fragment, probably in the
form of a duplex, can be ligated directly to eukaryotic genes prior to their insertion into bacterial plasmids or phage vectors. The presence of such a sequence containing the initiation signals immediately in front of the eukaryotic initiation codon would provide a recognition site that would otherwise be lacking for the attachment of bacterial ribosomes.

As shown in Fig. 1, the expected sequence of the icosaedoxyribonucleotide, synthesized using the "triester" approach (13, 14) and referred to as the 20-mer, has been confirmed by the two-dimensional "wandering spot" technique (15-17). In addition, a 17-mer that is identical with the 20-mer except for the omission of the A-T-G initiation codon from the 3' terminus has similarly been synthesized and analyzed.

Specific Binding of the Synthetic 20-mer to E. coli Ribosomes—It has been clearly demonstrated that the binding of natural mRNAs to E. coli ribosomes, in the process of initiation of protein synthesis, does not occur unless the three initiation factors (IF1, IF2, and IF3) are present (18-20). These reactions proceed at optimal rates at Mg\(^{2+}\) concentrations of around 5 to 8.5 mM. In marked contrast, synthetic initiation factor (IF1, IF2, and IF3) are present (18-20). These reactions proceed at optimal rates at Mg\(^{2+}\) concentrations at around 18 mM and this binding reaction occurs maximally in the absence of initiation factors (1).

In order to determine the specificity and efficiency of the synthetic oligodeoxyribonucleotides to serve as templates for the initiation of protein synthesis, we have used the filter binding assay to measure the extent of binding of \[^{35}S\]Met-tRNA to E. coli ribosomes in the presence of added templates (21). Table I shows that the binding of \[^{35}S\]Met-tRNA in the presence of the synthetic 20-mer is absolutely dependent upon added initiation factors IF2 and IF3. In addition, this binding occurs at low Mg\(^{2+}\) concentrations, with optimum at around 8.5 mM, as with natural mRNA. Although the extent of binding of \[^{35}S\]Met-tRNA is higher at 11.5 mM Mg\(^{2+}\) as compared to that at 8.5 mM, it is worth noting that the initiation factor dependence for the binding is much higher at 8.5 mM Mg\(^{2+}\) than at 11.5 mM, thus indicating increased nonspecific binding at elevated Mg\(^{2+}\) concentrations. This binding of \[^{35}S\]Met-tRNA drops off at higher Mg\(^{2+}\) concentrations, conditions that favor nonspecific binding of templates. These results suggest that the synthetic 20-mer is recognized by E. coli ribosomes only in the presence of initiation factors and at physiological Mg\(^{2+}\) concentrations.

Absolute Requirement for IF3 to Induce Binding of the 20-mer—Initiation factor IF3 has been shown to be indispensable for the recognition of natural mRNAs (18, 19, 22). As shown in Table II, IF2 alone is insufficient to induce the binding of \[^{35}S\]Met-tRNA to E. coli ribosomes in the absence of either R17 RNA or single-stranded fd DNA. The inclusion of IF3 is

![Fig. 1. Sequence analysis of the synthetic 20-mer by the "wandering spot" technique.](image-url)
In order to determine whether the A-T-G initiation codon present at the 3' terminus of the 20-mer plays a role either in inducing ribosome binding or in the stabilization of the initiation complex once it is formed, we have compared the binding of the 20-mer to that of the 17-mer. The latter has the terminal A-T-G omitted. As shown in Table II, the extent of binding of [35S]fMet-tRNAf is comparable for either template. This result suggests that the A-T-G initiation codon is probably not required for the initial selection of proper initiation sites for protein synthesis, and that the selection of the correct AUG codon by fMet-tRNAf to initiate protein synthesis takes place subsequent to proper binding of the fMet-tRNAf, 30 S complex to the initiation site on the mRNA.

**Formation of Functional 71 S Initiation Complexes Directed by the 20-mer**—The 1F3-dependent binding of fMet-tRNAf to E. coli ribosomes in the presence of the synthetic 28-mer results in the formation of stable 71 S initiation complexes, as detected by velocity centrifugation in sucrose gradients (Fig. 2). Like initiation complexes formed in the presence of natural mRNAs, these 71 S complexes are fully functional as shown by the ability of puromycin to release the [35S]fMet-tRNAf from the ribosomes.

### DISCUSSION

The use of chemically synthesized oligodeoxyribonucleotides of specific base sequences to study the initiation of protein synthesis represents a novel approach. This would overcome the difficulty in obtaining mutants with specific base changes in the ribosome binding site for functional studies. Although the use of synthetic oligoribonucleotides to study ribosome binding has recently been reported (9, 10), the oligoribonucleotides used were mostly simple A-U-G-containing analogues which lack the necessary ribosome binding site or initiation signals. These studies were also carried out at low temperatures and in the presence of high concentrations of both synthetic templates and Mg2+. Since most triplets can direct the binding of the corresponding tRNA to ribosomes (28, 29), it is difficult to ascertain the functional significance of those studies.

The preliminary results presented here clearly demonstrate that the synthetic 20-mer, representing a hypothetical ribosome binding site, can specifically and efficiently induce the initiation of protein synthesis. It also supports that nucleotide sequences outside the ribosome binding site are not required for the formation of a functional initiation complex. However, it does not exclude the possibility that sequences outside the ribosome binding site play a role in regulation by formation of specific secondary structures. Through the use of analogues with specific base substitutions, deletions, and insertions, it is hoped that several important questions may be answered. They include: (i) does the number of base pairs between the 16 S tRNA and the polyuridine stretch within the ribosome binding site play a functional role in determining the rate of initiation; (ii) does the number of nucleotides lying between the polypurine stretch and the initiation codon affect the rate of initiation; (iii) what determines the selection of the correct initiation codon in mRNAs where there is more than one AUG triplet at the 3' side of the ribosome binding site.

Attempts to determine the ideal sequence in natural mRNAs for efficient recognition by the ribosome and to assess the importance of the different elements in the recognition signal have thus far failed. The reason for this failure, we believe, is because many variables control the efficiency of initiation, and the observed differences in the rate of initiation between cistrons with quite different nucleotide sequences at
their ribosome binding sites may be the result of more than one determinant. The use of synthetic oligodeoxyribonucleotides with specific single base changes, free of interference by secondary structures and adjacent nucleotide sequences, may be the solution to these problems.

It is hoped that from these studies a most efficient ribosome binding site may be found which can be ligated to specific mammalian genes for cloning in bacterial systems. This is of great importance for large scale production of biologically important mammalian proteins in prokaryotes for experimental and clinical use.

REFERENCES