Nucleotides Contiguous to AUG Affect Translational Initiation†

M. Clelia Ganoza,* Alan R. Fraser, and Thomas Neilson

ABSTRACT: Genetic and biochemical data suggest that sequences longer than a triplet are required to accurately initiate translation of natural mRNAs. We have examined how bases contiguous with the initiation codon AUG, at either the 5′ or 3′ side of this triplet, affect the formation of a prokaryotic 70S ribosome·oligonucleotide·fMet-tRNA* initiation complex. The reaction was measured both in presence and absence of initiation factors. The following oligoribonucleotides, with bases adjoined to the 3′ side of AUG, formed the initiation complex almost as efficiently as did the AUG triplet itself: AUGUAA, AUGUUA, AUGUUU, AUG$(A)_n$, AUG$(C_1,A_2)_n$, and AUG$(U_4,A_1)_n$, where $n \simeq 18$. These polymers initiated equally well in the presence or absence of purified initiation factors. Bases adjoined to the 5′ end of the AUG codon, in contrast, had considerable influence on complex formation as follows. Tetramers of composition PuAUG formed initiation complexes very poorly, whereas PyAUG molecules were almost as effective as the AUG codon. The pentamers AUAUG and UUAUG were found to be equally effective initiators, suggesting that the purine–pyrimidine specificity in question is restricted to the first base 5′ to AUG. Hexamers bearing different sequences at the 5′ side of AUG varied in their ability to form an initiation complex, although the purine–pyrimidine effect of the first 5′ position was maintained. All of these specificities were maintained whether or not initiation factors were present. This latter result suggests that some specificity resides with the first base at the 5′ side of AUG and that initiation factors do not, at this level of discrimination, perturb the inherent property of 70S ribosomes to recognize base content.

Direct evidence is lacking that any sequence longer than the start codons AUG, GUG, and UUG is required to accurately phase the initiation of protein synthesis. However, because the initiation codons AUG, GUG, and UUG are ambiguous in that they also specify insertion of methionine, valine, or leucine, respectively, into internal positions in the protein, longer sequences are thought to govern discrimination between starter and internal codons. One hypothesis is that purine-rich regions four to ten bases 5′ to the actual initiator codons can base-pair directly with a polypyrimidine region at the 3′ terminus of the 16S rRNA (Steitz & Jakes, 1975; Shine & Dalgarno, 1975).

This hypothesis does not appear sufficient to explain the specificity of initiation. We find by scrutiny of the nucleic acid sequence of MS2 bacteriophage RNA (Fiers et al., 1975, 1976) or ϕ X174 DNA (Sanger et al., 1977) that a number of regions 5′ to internal AUGs or GUGs exhibit equally good complementarity to the 16S rRNA 3′ terminus. Yet, these regions do not initiate synthesis. How, then, does the translation apparatus select the correct initiation codons rather than other AUGs, GUGs, or UUGs along the mRNA?

To answer this question, we have started to systematically examine the influence that bases surrounding the AUG codon have on the formation of an initiation complex. Our approach is designed to extend the work of others who have attempted to define the code that begins protein synthesis (Ochoa & Mazumder, 1974; Steitz & Jakes, 1975; Wahba et al., 1966; Thach et al., 1966). We propose to use oligomers chemically synthesized by the phosphotriester method (Neilson & Werstiuk, 1974) to obtain, in high yield and purity, defined polymers containing numerous base combinations. These polymers can then be tested for activity in model initiation, termination, and propagation reactions.

A striking early finding on initiation we report here is that ribosomes recognize juxtaposed bases at the 5′ rather than the 3′ end of AUG in oligomers of relatively simple composition. Oligomers containing a pyrimidine on the 5′ side of AUG directed the formation of a complex with fMet-tRNA more effectively than did oligomers having a purine in this position. This discrimination between productive and less productive oligomers was manifest whether or not initiation factors were included in the reaction.

Materials and Methods

A. Chemical Synthesis of Oligonucleotides

Synthesis and Proof of Chemical Sequence. The tetramers, pentamers, and hexamers used in this study were synthesized by the phosphotriester method described previously (Neilson & Werstiuk, 1974; England & Neilson, 1976; Werstiuk & Neilson, 1976). Sequence integrity of oligoribonucleotides was checked after each synthetic step. Complete snake venom and bovine spleen phosphodiesterase digestions were carried out at 37 °C on the deblocked material in 0.1 M Tris-HCl (pH 8.0) and 0.1 M sodium succinate (pH 6.5) buffers, respectively. The mobilities of these digestion products were compared with authentic nucleotide and nucleoside markers on Avicel F thin-layer chromatography plates using ethanol–1.0 M NH_4OAc (70:30, v/v) pH 7.3. The ^1H NMR spectral analysis of trinucleotides allowed the identification of individual distinct aromatic proton resonances as sharp signals. Thus, oligomer base ratios could be confirmed.

Materials. Nucleosides were purchased from Terochem

† From the Banting and Best Department of Medical Research, Toronto, Ontario, M5G 1L6, Canada (M.C.G. & A.R.F), and the Department of Biochemistry, McMaster University, Hamilton, Ontario, L8S 4J9, Canada (T.N.). *Received December 19, 1977.* This work was sponsored by a grant from the Medical Research Council of Canada and partly by the J. P. Bickel Foundation.

[1] Abbreviations are: fMet-tRNA, *N*-formylmethionyl-tRNA; AUG, trinucleotide diphosphate ApUpG. Abbreviations for translation factors follow Caskey et al. (1972). In addition to the abbreviations recommended by IUPAC-IUB commission [(1970) *Biochemistry 9*, 4022], the following are also used: trac, triphenylmethoxyacetyl; t, tetrahydropyranyl; *pOθ*, 3′-*O*-(2,2,2-trichloroethyl) phosphate; *p* between two characters, 3′,5′-(2,2,2-trichloroethyl) phosphotriester; MST, mesitylenesulfonyl 1,2,4-triazolide.

0006-2960/78/0417-2769$01.00/0

Labs Limited, Edmonton, Alberta, Canada; mesitylenesulfonyl chloride, 1,2,4-triazole, and 2,2,2-trichloroethanol were from Aldrich, Milwaukee, Wis.; methylene chloride and silica gel (40–140 mesh) were from J.T. Baker, Philipsburg, N.J.; Avicel F and silica gel G (250 microns) thin-layer chromatography plates were from Analtech Inc., Newark, Del.; bovine spleen (EC 3.1.4.18) and snake venom (EC 3.1.4.1) phosphodiesterase were from Worthington, Freehold, N.J.

Instrumentation. UV absorbance measurements were made on a Cary 118 spectrophotometer. ^1H NMR spectra at 65 °C were recorded on a Bruker WH 90 MHz spectrometer with Fourier transform analysis ability. The solvent used for ^1H NMR analysis was prepared by lyophilizing a solution of 1.0 M NaCl, 0.01 M sodium phosphate (pH 7.0) and redissolving in an equal volume of D_2O (ca. pD 7.2). Oligomers were dissolved in 200–400 μL at a concentration of about ca. 10^{-2} M.

B. Enzymatic Synthesis of Oligonucleotides

Primer-dependent polynucleotide phosphorylase (EC 2.7.7.8), a gift from Dr. D. Logan, York University, Toronto, Canada, was used to synthesize $AUG(A)_n$, $AUG(C_1A_2)_n$, and $AUG(U_4A_1)_n$ as well as certain codons using the method described by Thach (1966). The dinucleotides UpA, ApU, and UpG (purchased from Raylo and Co., Edmonton, Alberta, Canada) were used as primers for polynucleotide phosphorylase to enzymatically synthesize AUG, UAG, UGA, and UAA. Triplets ending in G were made in the presence of ribonuclease T_1 (EC 2.7.7.26) (purchased from Worthington, Freehold, N.J.). After alkaline phosphatase (*E. coli*, EC 3.1.3.1, Worthington, Freehold, N.J.) treatment, each triplet was isolated on DEAE-Cl^--cellulose (Whatman 32) columns, pH 7.6, using a linear gradient of 0–0.5 M NH_4HCO_3.

Oligoribonucleotides of composition $AUG(A)_n$, $AUG(C_1A_2)_n$, and $AUG(U_4A_1)_n$, where $n \simeq 18$, were also synthesized using the codon AUG as primer for polynucleotide phosphorylase and the resulting products were isolated as described above. Base ratios were analyzed on Dowex-1 and -5 after hydrolysis for 18 to 20 h at 37 °C in 0.33 N KOH. Incorporation data, using Q_{13} ribosomes and *Lactobacillus arabinosus* soluble factors (Wahba et al., 1966; Ochoa & Mazumder, 1974), indicated that the oligomers incorporated each of the expected amino acids and verified the average chain length of 18 as assessed by Sephadex chromatography (Stanley et al., 1966).

C. Initiation Complex Formation

Initiation with 70S Ribosomes. Seventy S ribosomes were isolated from Q_{13} cells as described by Ganoza & Barraclough (1975). Electron microscopic studies, a courtesy of Dr. P.Ottensmeyer, Ontario Cancer Institute, Toronto, Canada, showed that the isolated ribosomes existed as 70S particles. This was confirmed by sucrose gradient centrifugation. The f[^{35}S]-Met-tRNA was prepared as previously described (Ganoza et al., 1976). The specific activity of the [^{35}S]methionine (purchased from New England Nuclear Corp.) was about 450 Ci/mM. Unless otherwise specified, the ribosome·AUG·f[^{35}S]Met-tRNA complex was prepared in a solution of 10 mM $MgCl_2$ as described previously (Ganoza & Barraclough, 1975) except that incubation was for 20 min at 24 °C.

Initiation Factor Dependent Formation of f[^{35}S]Met-tRNA·Oligomer·Ribosome Complex. Single colonies of Q13 cells were tested for the RNase I marker as described by Gesteland (1966) prior to growth and after asceptic harvest of the early log cells (2×10^8 cells/mL). Ribosomes were washed two (Ganoza & Barraclough, 1975) and four times (Ganoza et al., 1976). Ribosomes that were washed twice required IF-1 and IF-2 for MS2 RNA directed synthesis (Wahba et al., 1966; Ochoa & Mazumder, 1974; Iwasaki et al., 1968). Ribosomes that were washed four times required IF-3 in addition to IF-1 and IF-2 (Wahba et al., 1966; Ochoa & Mazumder, 1974). Initiation complex formation was carried out in a final volume of 0.045 mL; 100 μg of two or four times washed *E. coli* Q13 ribosomes (Ganoza & Barraclough, 1975), 5.5 mM $MgCl_2$, 160 mM NH_4Cl, 0.28 mM GTP, 1.1 mM DTT in a 55 mM Tris buffer, pH 7.4. To this was added the indicated levels of oligonucleotides (250 pmol of AUG for a standard factor preparation), 5 μg of IF-1, and 2.5 μg of IF-2. Reactions were supplemented, when necessary, with highly purified IF-3 (5 μg).

Purification of Initiation Factors. IF-1, IF-2, and IF-3 were purified (Suttle et al., 1973) from midlog *E. coli* K12 cells (Miles Laboratories, Inc., Elkhart, Ind.) using either the 0.5 M or the 1.0 M NH_4Cl first ribosomal wash. The latter wash contained considerably more IF-3. The factors from 200 g of cells were then concentrated by precipitation with ammonium sulfate to 70% saturation at 4 °C, dialyzed using sulfide-free spectropore dialysis bags, and chromatographed on DEAE-Cl^--cellulose (Whatman DE-32 or DE-52) and phosphocellulose Na^+ (Whatman P-11). Factors were concentrated on Amicon filters using PM10 (for IF-1 or IF-3) or PM30 filters for IF-2. IF-1 and IF-3 were further concentrated by dialysis against 50% glycerol, 10 mM Tris, pH 7.4, and 1 mM DTT. Low levels of IF-2 were used to score for IF-1 (Wahba et al., 1966; Suttle et al., 1973). IF-3 was assayed at 37 °C with MS2 RNA (Iwasaki et al., 1968) using saturating levels of IF-1 and IF-2 and scoring for the synthesis of acid insoluble proteins directed by MS2 RNA. Analysis of reactions with IF-1, IF-2, and IF-3 revealed that a significant fraction of the ribosomes occurred as 70S particles.

MS2 RNA bacteriophage was grown and harvested and the RNA was extracted and analyzed on formaldehyde–sucrose gradients with [^{3}H]λDNA as marker (Van der Meer & Ganoza, 1975).

Initiation factor dependence was checked routinely using both AUG and MS2 RNA in the binding reaction described above. MS2 bound 30% less fMet-tRNA than the free AUG codon when 40 and 62.5 pmol respectively of each RNA were used. Four times washed ribosomes bound: from 2600 to 4000 dpm of fMet-tRNA with AUG and MS2 RNA and no added IF-1, IF-2, and IF-3; about 24 000 dpm with 2.0 μg of IF-1 and 1.0 μg of IF-2 and about 48 000 dpm with these amounts of IF-1 and IF-2 and 4.5 μg of IF-3. Essentially no binding was observed without IF-2. Analyses of the bound counts with and without MS2 were performed after sucrose density analysis through 8% sucrose discontinuous gradients and electrophoresis at pH 3.5 (Van der Meer & Ganoza, 1975). The results verified that fMet-Ala was synthesized on addition of propagation factors and alanyl-tRNA indicating that the coat site of MS2 was the only detectable initiation site. Addition of other proteins to this system (ribosomes were washed twice) such as the rescue protein (Ganoza et al., 1973) or hydrolases (Ganoza et al., 1976) had no effect. Four times washed ribosomes required IF-3 for complex formation.

Results

Chemical Synthesis and Characterization of Oligoribonucleotides. To examine whether bases adjacent to AUG influenced the initiation reaction, we have begun to construct a series of oligoribonucleotides, (N)AUG(M), where N and M represent any ordered sequence of the nucleotides A, G, C, and U from 0 to 18 bases long. All oligoribonucleotides used

TABLE I: Summary of Preparation of Protected Oligoribonucleotides.[a]

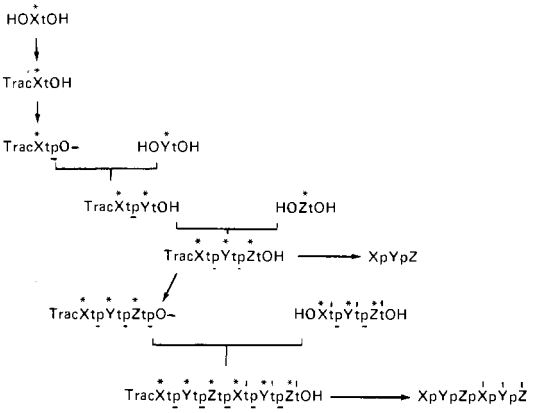
reactants						product		
	quantity			quantity			quantity	
compd	mg	mmol	compd	mg	mmol	compd	mg	%
A	823	1.10	U	420	1.28	AU	970	70
AU	750	0.59	G	400	0.85	AUG	700	61
A	500	0.66	A	360	0.79	AA	500	53
AA	250	0.18	A	97	0.21	AAA	190	53
AA	500	0.36	U	140	0.42	AAU	505	73
AAU	300	0.155	G	88	0.187	AAUG	220	54
U	610	0.97	A	533	1.16	UA	725	59
UA	630	0.49	A	340	0.75	UAA	610	64
UA	320	0.25	U	97	0.28	UAU	265	59
UAU	200	0.110	G	63	0.134	UAUG	156	60
U	300	0.47	U	185	0.56	UU	310	55
UU	300	0.26	A	142	0.31	UUA	275	58
UU	500	0.44	U	170	0.52	UUU	510	68
C	440	0.59	A	328	0.72	CA	400	48
CA	400	0.29	U	114	0.35	CAU	286	52
CAU	150	0.079	G	45	0.096	CAUG	95	47
GA	350	0.25	U	97	0.28	GAU	300	62
GAU	250	0.129	G	73	0.155	GAUG	180	53

[a] Column 1 contains the 5′-trityloxyacetyl reactants, and A stands for TracBz-AtOH; column 4 contains the incoming nucleosides and U stands for HOUtOH: column 7 contains the trityloxyacetyl product and AU stands for TracBz-At-*p*-UtOH. Two equivalents of pyridinium mono-2,2,2-trichloroethyl phosphate activated by 4 equiv of MST in anhydrous pyridine is used in each phosphorylation step. The coupling step to the incoming nucleoside is driven by 1.2 equiv of MST.

TABLE II: Summary of the Specific Deblocking Data of the 5′-Trityloxyacetyl Oligomers.[a]

reactants		product		
	quantity		quantity	
compd	(mg)	compd	mg	%
TracAUG	620	HOAUG	355	68
TracUAA	300	HOUAA	215	86
TracUUU	500	HOUUU	310	77
TracUUA	260	HOUUA	180	83

[a] Ammonia (0.15 M) in methanol was reactant solvent (10 mg in 1 mL).



where X or Y or Z = A^{bz} or G^{bz} or C^{bz} or U

and X or Y or Z = A or G or C or U

FIGURE 1: Chemical synthesis of oligonucleotides of defined sequence. Abbreviations and details of synthesis are described in the text.

in these studies were prepared by the general phosphotriester synthesis developed by Neilson and his associates (Neilson & Werstiuk, 1974; England & Neilson, 1976, Werstiuk & Neilson, 1976). At the time of writing, only this method allowed satisfactory coupling to guanosine residues so as to provide sufficient quantities of oligomers of defined sequence beyond the trimer level.

An outline of the synthetic scheme is presented in Figure 1. Stepwise assembly of a protected oligoribonucleotide proceeds from the 5′ terminus toward the 3′ terminus of any sequence or part of a sequence (block). For sequences containing five or more base residues, stepwise synthesis becomes tedious and a method whereby blocks can be added was used. This was especially true for long sequences containing common short sequences such as AUG.

Tables I–III contain the preparative data for protected oligoribonucleotides. A three-step deblocking procedure was followed (England & Neilson, 1976). Deblocked oligoribonucleotides were isolated using paper chromatography (Table IV). Sequence integrity was checked by complete digestion of the products after each synthetic step as described in Materials and Methods.

All the ribonucleotides used in this work were analyzed by descending liquid paper chromatography in two organic solvents (see Table IV). For every oligoribonucleotide synthesized, only one compound could be detected in each solvent system. We then examined whether these triplets or oligomers were able to stimulate the termination and initiation reactions. As shown in Figure 2, AUG and UAA made by the chemical method gave identical stimulation of initiation and termination as did AUG and UAA made enzymatically. The same successful check was made for UAG, UGA, and UAAAUG in respect to termination (data not shown).

To learn whether the number and type of bases at either the 3′ or 5′ terminus of AUG affected the formation of an initiation complex, a series of oligoribonucleotides was studied as follows.

Effect of Nucleotides Present at the 3′ Side of AUG. We first examined the influence of bases at the 3′-terminal end of the start codon AUG on the formation of an initiation complex with 70S ribosomes. AUGUAA, AUGUUA, and AUGUUU

TABLE III: Summary of Preparation of Protected Oligoribonucleotides.[a]

reactants						product		
	quantity			quantity			quantity	
compd	mg	mmol	compd	mg	mmol	compd	mg	%
AUG	250	0.129	UAA	215	0.133	AUGUAA	140	29
AUG	200	0.103	UUA	156	0.104	AUGUUA	121	31
AUG	250	0.129	UUU	210	0.154	AUGUUU	160	35
UAA	275	0.142	AUG	246	0.150	UAAAUG	190	35
UUA	200	0.112	AUG	240	0.168	UUAAUG	178	44
AAA	170	0.083	AUG	164	0.100	AAAAUG	150	46
CAU	200	0.105	AUG	173	0.105	CAUAUG	170	43
UU	165	0.143	AUG	180	0.109	UUAUG	130	40
AU	183	0.143	AUG	180	0.109	AUAUG	175	51

[a] Three equivalents of reactant is used in the first step of the coupling.

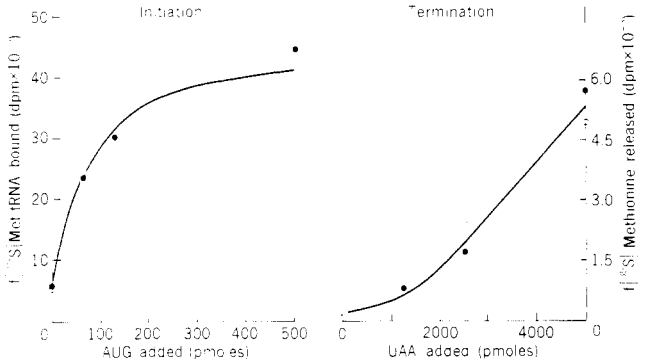


FIGURE 2: Biological activity of chemical and enzymatically synthesized triplets in termination and initiation reactions. In the initiation reaction, binding of fMet-tRNA to ribosomes was performed as described in Materials and Methods using the indicated levels of AUG synthesized chemically or with polynucleotide phosphorylase. Termination reactions using ribosomal-bound f[^{35}S]Met-tRNA and UAA made chemically (see Figure 1 and Materials and Methods) or with polynucleotide phosphorylase were described in Materials and Methods. For both termination and initiation reactions, 450 μg of twice washed ribosomes and 51 800 dpm of fMet-tRNA were used in each incubation mixture. The shaded circles represent reactions containing chemically synthesized triplets and the open circles represent reactions containing enzymatically synthesized triplets.

TABLE IV: Experimental Data of Free Oligoribonucleotides.

sequence	R_{Up}[a] system A	R_{Up}[a] system B	yield (%) for deprotection[b]
AUG	0.68	0.69	40
AAAAUG	0.16	0.21	23
UAAAUG	0.47	0.45	25
UUAAUG	0.48	0.45	27
AUGUUU	0.32	0.49	33
AUGUUA	0.29	0.48	32
AUGUAA	0.27	0.29	29
CAUG	0.55	0.56	29
AAUG	0.44	0.44	37
UAUG	0.58	0.58	35
GAUG	0.40	0.46	31
UUAUG	0.52	0.56	30
AUAUG	0.44	0.49	31
CAUAUG	0.31	0.40	21
Up	1.0	1.0	

[a] System A: 1 M ammonium acetate–ethanol (50/50) on Whatman 40. System B: 1 M ammonium acetate–ethanol (70/30) on Whatman 3MM. [b] Calculated from UV spectrophotometric data assuming a 90% hypochromicity factor.

formed initiation complexes with an efficiency comparable to that of AUG (Figure 3a). The same behavior was observed with the oligoribonucleotides AUG(A)$_n$, AUG(A$_1$C$_2$)$_n$, and AUG(U$_4$A$_1$)$_n$ where $n \simeq 18$ (Figure 3b). These results suggest that the nature and the number of bases linked to the 3′ end of AUG have little or no effect on the formation of an initiation complex.

Effect of Nucleotides Present at the 5′ Side of AUG. In contrast to the results above, oligoribonucleotides having bases at the 5′-terminal end of AUG were found to be unequal in promoting initiation. Our results show that UAUG and CAUG can form an initiation complex almost as efficiently as AUG (Figure 4a). On the other hand, when G or A were linked to the 5′ end of AUG, initiation was much less efficient and required higher concentrations of messenger (Figure 4a). A time course experiment (Figure 4b) at one tetramer concentration (64 pmol) clearly shows how a single purine at the 5′ terminus of AUG drastically reduces the rate and yield of the initiation reaction.

For the oligomers that bound poorly to ribosomes such as GAUG and AAUG (above), a simple check was made for traces of inhibitors that could block the initiation reaction. The reference triplet AUG was mixed with each of these less productive oligomers in turn and tested for its ability to promote initiation. In no case was there any inhibition observed (data not shown). In a similar vein, we asked if the low efficiency for initiation of AAUG and GAUG could be due to a chemically mixed population of messenger molecules—some active, most of them inactive. In order to test this possibility, f[^{35}S]Met-tRNA binding was directed by an excess of GAUG, in the presence of 16 pmol of ribosomes until reaction was complete at 20 min. About 40×10^3 dpm of fMet-tRNA was bound at this point. When 33 pmol of ribosomes was used in the reaction, 80×10^3 dpm of fMet-tRNA was complexed. This result suggests that the unbound tetramer molecules are chemically equivalent to the complexed ones since they can enter new initiation complexes.

Studies with Pentamers and Hexamers. The dose–response and time–course curves of the initiation reaction programmed by UUAUG, AUAUG, as well as AUG are shown in Figures 5a and 5b. In contrast to the effect observed when a purine is apposed to the 5′ side of AUG, the inclusion of either a purine (A) or a pyrimidine (U) 5′ terminal to the tetramer UAUG has no significant effect on the efficiency of formation of an initiation complex. It appears, therefore, that the deleterious effect for initiation that a 5′-terminal purine can have on AUG is manifest only if the purine is juxtaposed to the AUG. Alternatively, once AUG within a tetramer is chemically defined as productive by the presence of a 5′-terminal pyrimidine, the inclusion of a further 5′ A does not alter its efficiency as an initiator.

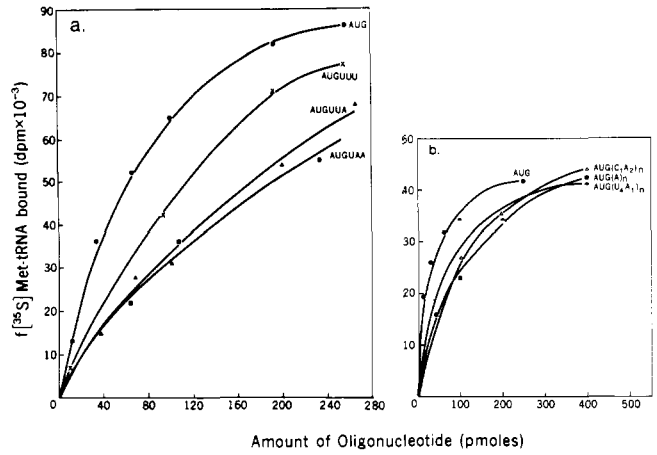


FIGURE 3: Initiation complex formation with AUG, AUGUUU, AUGUUA, AUGUAA, and AUG$(C_1A_2)_n$, AUG$(A)_n$, or AUG$(U_4A_1)_n$. Initiation complex formation with *E. coli* Q_{13} 70S ribosomes was as described in the legend to Figure 2 and in the Materials and Methods section. Synthesis of oligonucleotides was as described by Thach (1966) and in Materials and Methods. The indicated amounts of each polymer were used in all reaction mixtures. In experiment a, incubations contained 128 000 dpm of f[^{35}S]Met-tRNA. Reactions were for 20 min at 24 °C. In experiment b, incubations contained 61 000 dpm of f[^{35}S]Met-tRNA.

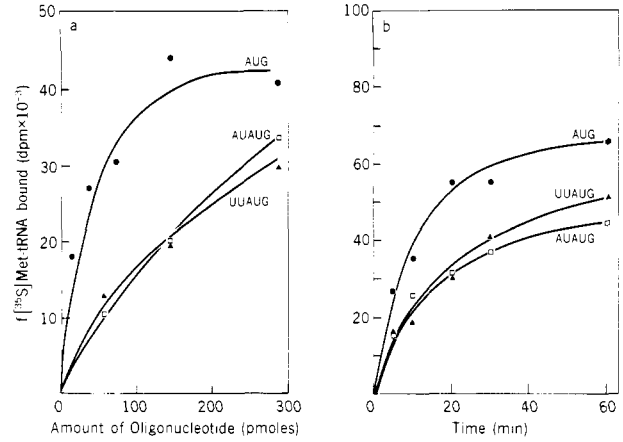


FIGURE 5: Initiation complex formation with AUG and AUAUG or UUAUG. In experiment a, initiation complex formation was measured as described in the legend to Figure 2, using the indicated levels of each tri- or pentanucleotide template. In experiment b, the time course of the complex formation was followed using 64 pmol of each respective tri- or pentanucleotide. Reaction conditions were as described in Figure 2 and in Materials and Methods. Each reaction contained 61 000 dpm of f[^{35}S]Met-tRNA.

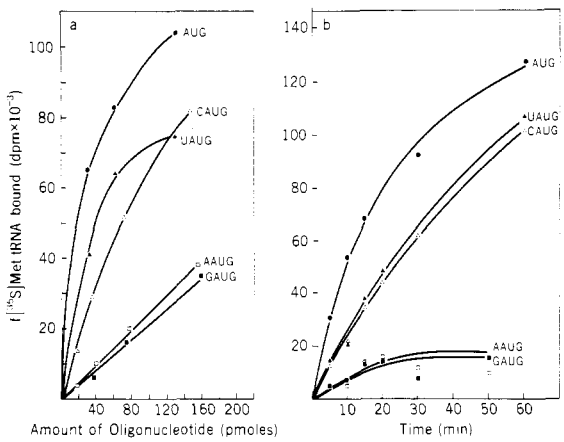


FIGURE 4: Initiation complex formation with AUG, PyAUG, or PuAUG. In experiment a, initiation complex formation was measured for 20 min at 24 °C using the indicated oligonucleotide concentrations. Preparation of ribosomes, f[^{35}S]Met-tRNA (185 000 dpm in each reaction mixture), and assay procedures are described in Materials and Methods. In experiment b, the time course of the reaction was examined as in experiment a using 64 pmol of each oligonucleotide and 185 000 dpm of f[^{35}S]Met-tRNA.

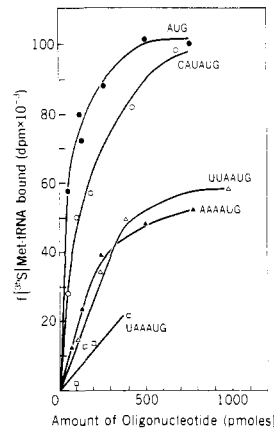


FIGURE 6: Initiation complex formation with AUG and various hexamers containing a 3′-terminal AUG. The indicated amounts of oligonucleotides were used to form a complex with four times washed 70S ribosomes and no added initiation factors at 10 mM $MgCl_2$. Reaction conditions are described under the legend to Figure 2 and in Materials and Methods. Reaction mixtures contained 160 400 dpm of f[^{35}S]Met-tRNA.

Since some naturally occurring initiator AUGs are immediately preceded by a purine (Steitz et al., 1977), it will be of interest to determine if the addition of one or more specific bases on the 5′ side of AAUG (or GAUG) can convert these oligomers into more effective initiators. Such studies are in progress.

A more complex picture emerged when the initiation reaction was programmed by hexamers bearing bases of defined composition on the 5′ side of AUG.

The experiments summarized in Figure 6 show that highly purified 70S ribosomes can still be influenced by nucleotides at the 5′ end of AUG. For example, a complex between CAUAUG, ribosomes, and fMet-tRNA was formed with almost the same efficiency as the AUG·ribosome·fMet-tRNA reference complex. The CAUAUG sequence occurs at the A protein initiation site of phage Qβ (Fiers et al., 1975). The interference to initiation by a juxtaposed 5′-sided A was maintained in the hexamers UUAAUG and AAAAUG. The hexamer UAAAUG was an even poorer initiator than AAUG. As oligomers become larger and more varied in structure, a number of factors can come into play and affect their competence in the initiation reaction. Such factors may include intra- and intermolecular complementarity, complementarity with the 16S rRNA 3′ terminus, the effect of in- or out-of-phase nonsense codons (Ganoza, unpublished data), and so forth.

Lack of Effect of Initiation Factors on Specificity. The effect of initiation factors on the formation of the initiation complex was also examined. In order to obtain as complete a dependence as possible on the purified initiation factors used, ribosomal particles were purified by repeated elution and centrifugation. Such ribosomes were suspended and centrifuged four times (Ganoza et al., 1976). This treatment removes most of the IF-3 from the ribosomes and certain proteins required to reconstruct translation (Ganoza, 1977; Ganoza et al., 1973; Glick & Ganoza, 1975; Ganoza & Fox, 1974), as well as the fMet-tRNA hydrolases (Ganoza & Barraclough, 1975; Ganoza et al., 1976). At 10 mM Mg^{2+} (Figure 7), purified initiation factors IF-1, IF-2, and IF-3 had no discernible effect on the degree of binding of fMet-tRNA by these oli-

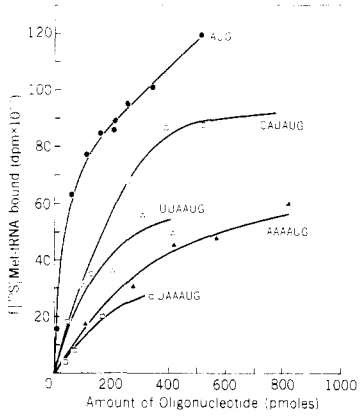


FIGURE 7: Initiation complex formation in the presence of initiation factors and various hexamers containing a 3′-terminal AUG. The indicated amounts of each hexamer or AUG were used in the "initiation-dependent" assay described in Materials and Methods. Each reaction was for 20 min at 24 °C and included *E. coli* Q_{13} ribosomes, washed four times, the described levels of initiation factors IF-1, IF-2, and IF-3, and 290 000 dpm of f[^{35}S]Met-tRNA. Approximately 2000 dpm was incorporated with AUG or each of these hexamers in the absence of initiation factors. This value was subtracted from all the above numbers. About 2800 dpm was incorporated with 2 μg of either IF-1 or IF-2. Two micrograms of each of IF-1 and IF-2 increased incorporation with each polymer about twofold. Addition of 4 μg of IF-3 gave the values reported in this figure. Addition of tenfold more IF-2 obliterated the requirement for IF-1 and IF-3. The preparation and assay of these factors are described in Materials and Methods.

gomers. Although the initiation factors were absolutely required for this reaction at 5 mM Mg^{2+}, the inherent 5′-purine/pyrimidine effect in question was maintained in the initiation factor dependent system when all tetramers, pentamers, and longer oligoribonucleotides were used. All these results suggest that some base specificity for initiation complex formation exists and that it resides in the bases at the 5′ side of AUG. The known initiation factors do not appear to perturb the inherent property of ribosomes to recognize base content.

Comparison of fMet-tRNA and Oligomers Binding to Ribosomes. In the studies described so far, initiation was monitored by the binding of radiolabeled fMet-tRNA to ribosomes. It was of interest to learn if the differences in initiation determined in this way were correlated and, perhaps, governed by the extent to which the various oligomers bound to ribosomes. Accordingly, the binding of [^{32}P]AUGUAA, [^{32}P]CAUAUG, and [^{32}P]UAAAUG were also measured during initiation. We found that about 8.8 and 8.1 pmol of either CAUAUG and AUGUAA—both effective initiators—complexed with ribosomes during the reaction. In contrast, only 1.6 pmol or less of [^{32}P]UAAAUG bound to ribosomes. Approximately 1.0 pmol of fMet-tRNA bound when CAUAUG and AUGUAA were messengers, and 0.18 pmol when UAAAUG was used. These results suggest that it is the binding of messenger to ribosomes that determines the level of fMet-tRNA that enters the complex.

Discussion

One of the main unsolved problems in the translation of the genetic code is the mechanism whereby ribosomes recognize mRNA in the correct context, starting at certain internal AUG, GUG, and UUG triplets to initiate protein synthesis. These codons are ambiguous: a fraction of them direct initiation; others specify the insertion of methionine, valine, or leucine into internal positions of proteins.

There is evidence that the secondary and/or tertiary structure of MS2 bacteriophage RNA plays some role in starter-sequence selection. The start codons for MS2 replicase, as well as those for the maturation and A proteins, can be buried by folding (Fiers et al., 1975, 1976). After disruption of the secondary structure by heating and formaldehyde treatment (Lodish & Robertson, 1969) or by fragmentation of the RNA (Kozak & Nathans, 1972), the entire molecule may become exposed and yet the starting sites of the replicase and the A protein are still preferentially utilized. With formaldehyde treatment, only two other dipeptides, not found in vivo, are made (Lodish & Robertson, 1969). Furthermore, the isolated ribosome-protected regions of R17 mRNA rebind to ribosomes in preference to other random sequences containing AUG (Steitz, 1973). All these results suggest that the primary sequence in the region of a starter codon also plays an essential role in protein chain initiation.

The preponderance of alanine and serine at the amino termini of *E. coli* proteins (Waller & Harris, 1961) must raise the question of whether the codons for these amino acids, by virtue of their proximity to start triplets, play any role in the initiation reaction. Golini & Thach (1972) have argued that such sequences 3′ to the initiator codons are not decoded by aminoacyl-tRNAs in order to form an initiation complex. Steitz (1966) first showed that untranslated regions occur 5′-terminal to the mRNA sequences corresponding to the starting points of bacteriophage R17 peptides. It was thus implied that these regions, however diverse their composition appeared at first glance, carried the information for designating some AUG triplets as starters. Further inspection of a number of these untranslated regions led to the hypothesis that certain purines, 4 to 10 bases on the 5′ side of the starter codons, were complementary to the 3′ end of the 16S rRNA (Shine & Dalgarno, 1975). Thus, base pairing between messenger and a ribosomal element became a viable mechanism for site-specific initiation. However, a comparison of the sequences surrounding the initiation codons of translation with the sequences around the same codons where they specify methionine or valine in internal positions of proteins often reveals few distinguishing features (our unpublished observations). A statistically significant number of regions complementary to the 16S rRNA 3′ end were noted in internal in-phase or out-of-phase AUG or GUG codons of the MS2 genome. For example, regions 1038–1041 in the A protein, 1652–1657 in the coat, and 2381–2387 in the replicase exhibit as good if not better complementarity. This complementarity was observed whether the regions were expected to occur hydrogen bonded to other bases or as single-strand portions of the molecule (Fiers et al., 1975, 1976). Similar observations were also made by examining the sequence of the DNA of bacteriophage φX174 (Sanger et al., 1977). In a search for alternative start-determining mechanisms (Ganoza, 1977), the frequent occurrence of in- or out-of-phase nonsense codons, 5′ to initiator AUGs, was noted. Finally, it appears that ribosomes may well be able to initiate protein synthesis in vivo at the pppAUG . . . ends of certain messenger molecules (Ptashne et al., 1976), an alternative pathway certainly in keeping with the in vitro results presented here. The issue of how initiation codons are selected, therefore, remains an open question.

In this communication we report on studies designed to examine systematically in vitro the effect of bases that surround the starter codon AUG. The study is greatly facilitated by the use of chemical synthesis methods described here. Our results may be summarized as follows:

(1) The isolated triplet AUG is highly efficient as messenger to direct the entry of fMet-tRNA into a ribosome-bound initiation complex.

(2) The following oligoribonucleotides initiate with virtually

the same high efficiency: AUG; oligonucleotides bearing a 5′-terminal AUG followed by 18 3′-terminal bases containing A, AU, or AC; and the hexamers AUGUUA, AUGUAA, and AUGUUU. These results suggest that neither the nucleotide composition nor the chain length at the 3′ side of AUG has much effect on the efficiency of formation of the initiation complex in vitro. For the oligoribonucleotides of composition $AUG(X)_{18}$, kinetics of fMet-polypeptide synthesis sustain the conclusion. A more extensive analysis of the role of 3′-sided bases remains to be performed.

(3) Measure of initiation complexes formed with tetramers of composition PyAUG and PuAUG reveals that 70S ribosomes differ in their ability to initiate depending on the nature of the first base 5′ to AUG; UAUG and CAUG initiate well; AAUG and GAUG do not. The preference for pyrimidines (67%) over purines (33%) at the 5′ side of starter AUGs in native mRNAs is noted (Steitz et al., 1977).

The introduction of an A, or a U at the 5′ end of the efficient initiator UAUG, however, has no further effect on the competence of this molecule to direct the initiation reaction. It can be observed that a random proportion of purines and pyrimidines occurs in this penultimate 5′ situation in various natural mRNAs (Steitz et al., 1977).

A study of hexamers in which various triplets occur at the 5′ side of AUG shows them to vary in their ability to direct initiation. In a qualitative way, the preference of a pyrimidine over a purine in the juxtaposed 5′ position is maintained. Thus, UUAAUG and AAAAUG, whereas still not as effective as CAUAUG (nor the reference AUG), program initiation better than the hexamer UAAAUG.

(4) The purified initiation factors IF-1, IF-2, and IF-3 have no effect upon these specificities. It appears that 70S ribosomes have the inherent ability to recognize the above features in short oligoribonucleotides in the absence of initiation factors.

Although relatively few sequences are reported, the sequences we examined around the starter codon AUG suggest that the code to initiate synthesis may be highly flexible. In this circumstance, it may be profitable to search for forbidden configurations. We may have discovered one such forbidden configuration, UAAAUG. Somewhat unfavorable sequences such as PuAUG may restrict the frequency of initiation to a certain degree.

It is tempting to speculate that flexibility in the start code was required to preserve cell viability against mutation during the evolution of the genetic code. If the initiation code is indeed highly flexible, then a simple mutation in the vicinity of internal in- or out-of-phase AUGs could convert latent starters into sites which could readily generate proteins of new function for evolutionary selection.

Acknowledgments

We are grateful to Mrs. Nada Barraclough and Mr. Rene J. Gregoire for excellent technical assistance. We are also grateful to Dr. Thach for advice on preparing polynucleotides with polynucleotide phosphorylase and Drs. J. Ravel, H. Weissbach, B. R. Glick and A. J. Wahba for advice on preparation as well as gifts of IF_1 and IF_3 used to standardize the purification of these proteins. We thank Dr. A. J. Becker for constructive criticism of this work.

References

Caskey, T., et al. (1972) *Science 176,* 195–197.

England, T. E., & Neilson, T. (1976) *Can. J. Chem. 54,* 1714.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Merregaert, J., Min Jou, W., Raeymaekers, A., Volckaert, G., Ysebaert, M., Van de Kerckhove, J., Nolf, F., & Van Montagu, M. (1975) *Nature (London) 256,* 273.

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G., & Ysebaert, M. (1976) *Nature (London) 260,* 500.

Ganoza, M. C. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31,* 273.

Ganoza, M. C. (1977) *Can. J. Biochem. 55,* 267.

Ganoza, M. C., & Barraclough, N. (1975) *FEBS Lett. 53,* 159.

Ganoza, M. C., & Fox, J. L. (1974) *J. Biol. Chem. 249,* 1037.

Ganoza, M. C., Van der Meer, J., Debreceni, N., & Phillips, S. L. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70,* 31.

Ganoza, M. C., Barraclough, N., & Wong, J. T.-F. (1976) *Eur. J. Biochem. 65,* 613.

Gesteland, R. F. (1966) *J. Mol. Biol. 16,* 67.

Glick, B. R., & Ganoza, M. C. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72,* 4257.

Golini, F., & Thach, R. E. (1972) *Biochem. Biophys. Res. Commun. 47,* 1314.

Iwasaki, K., Sabol, S., Wahba, A. J., & Ochoa, S. (1968) *Arch. Biochem. Biophys. 125,* 542.

Kozak, M., & Nathans, D. (1972) *Bacteriol. Rev. 36,* 109.

Lodish, H. F., & Robertson, H. D. (1969) *Cold Spring Harbor Symp. Quant. Biol. 34,* 655.

Min Jou, W., Haegeman, G., Ysebaert, M., & Fiers, W. (1972) *Nature (London) 237,* 82.

Neilson, T., & Werstiuk, E. S. (1974) *J. Am. Chem. Soc. 96,* 2295.

Ochoa, S., & Mazumder, R. (1974) *Enzymes, 3rd Ed. 10,* 1.

Ptashne, M., Backman, K., Humayun, M. Z., Jeffrey, A., Maurer, R., Meyer, B., & Sauer, R. T. (1976) *Science 194,* 156.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison C. A. III, Slocombe, P. M., & Smith, M. (1977) *Nature (London) 265,* 687.

Shine, J., & Dalgarno, L. (1975) *Nature (London) 254,* 34.

Stanley, W.M., Jr., Smith, M. A., Hille, M. B., & Last, J. A. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31,* 99.

Steitz, J. A. (1966) *Cold Spring Harbor Symp. Quant. Biol. 34,* 621.

Steitz, J. A. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70,* 2605.

Steitz, J. A., & Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72,* 4734.

Steitz, J. A., Sprague, K. V., Steege, D. A., Yuan, R. C., Laughrea, M., Moore, P. B., & Wahba, A. J. (1977) in *Nucleic Acid-Protein Recognition* (Vogel, H. J., Ed.) p 491, Academic Press, New York, N.Y.

Suttle, D. P., Haralson, M. A., & Ravel, J. M. (1973) *Biochem. Biophys. Res. Commun. 51,* 376.

Thach, R. E. (1966) *Proced. Nucleic Acid Res.,* 520.

Thach, R. E., Sundararajan, T. A., Dewey, K. F., Brown, J. C., & Doty, P. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31,* 85.

Van der Meer, J. P., & Ganoza, M. C. (1975) *Eur. J. Biochem. 54,* 229.

Wahba, A. J., Salas, M., & Stanley, W. M., Jr. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31,* 103.

Waller, J. P., & Harris, J. I. (1961) *Proc. Natl. Acad. Sci. U.S.A. 47,* 18.

Werstiuk, E. S., & Neilson, T. (1976) *Can. J. Chem. 54,* 2689.