Kinetics of the Stages of Transcription Initiation at the *Escherichia coli lac* UV5 Promoter

Susan B. Straney[‡] and Donald M. Crothers*[,‡,§]

Departments of Molecular Biophysics and Biochemistry and of Chemistry, Yale University, New Haven, Connecticut 06511

Received November 18, 1986; Revised Manuscript Received March 11, 1987

ABSTRACT: The kinetics of initiation by *Escherichia coli* RNA polymerase on the *lac* L8UV5 promoter was studied by a gel retardation method that separates protein–DNA complexes from free DNA. The binding constant of the closed complex, the forward and reverse rate constants of isomerization from closed to open complex, and the forward rate constant from the open to initiated complex were measured. Both the forward and reverse isomerization rates were found to be temperature dependent, and the activation energies for these steps were determined. The rates of open complex formation and dissociation were not affected by the addition of ribonucleotide triphosphates; however, the extent of dissociation was greatly reduced if the triphosphates added allowed a short, unstable RNA product to form. The dissociation rate was not affected by heparin, a polyanion competitor that sequesters the polymerase. The rate of initiated complex formation appeared to be dependent on whether the initiating moiety was a mononucleotide triphosphate or dinucleoside monophosphate and on the sequence of the dinucleoside. These results are compared to those found on both the *lac* L8UV5 and other bacterial and phage promoters by less direct measurements. We use the values obtained for the individual rate constants to investigate the predicted steady-state kinetics of initiation-limited transcription, with the conclusion that the rate-limiting step is formation of the open complex in the limit of low polymerase concentration. However, when RNA polymerase is saturating, the rate is determined by the transition from open complex into the stably initiated ternary complex. This ability to switch the nature of the rate-determining step as a function of polymerase levels offers additional flexibility in control of gene expression.

A major transcriptional control point in prokaryotes exists as the initiation step, ensuring that the cell does not expend energy producing unwanted products. Many studies have been done on transcriptional initiation as well as the mechanism of action of RNA polymerase [for recent reviews, see McClure (1985) and Kumar (1981)]. The basic model of RNA polymerase transcription initiation was proposed and developed several years ago (Walter & Zillig, 1967; Chamberlin, 1974) and is shown in eq 1, where R is the RNA polymerase and

$$R + P \overset{K_B}{\rightleftharpoons} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} RP_o \xrightarrow{k_{init}} RNA \qquad (1)$$

P is the promoter site. RP_c is the closed complex, in which the DNA is not unwound. This complex is sensitive to heparin, a polyanion that competes with DNA for the polymerase. RP_o is the heparin-stable open complex, in which base pairs near position 0 (relative to the starting point for RNA synthesis) on the promoter DNA are unwound (Saucier & Wang, 1972; Kirkegaard et al., 1983). This model is undoubtedly a simplified version of the mechanism of RNA polymerase–DNA complex formation, and evidence for intermediate steps both before and after closed complex formation has been found (Roe et al., 1984; Shanblatt & Revzin, 1984; Straney & Crothers, 1985; Buc & McClure, 1985). Stable initiated complexes have also been reported, which contain RNA polymerase lacking the σ subunit, plus the DNA promoter and a 9–11-ribonucleotide oligomer (Carpousis & Gralla, 1985; Straney & Crothers, 1985).

Kinetic studies have been performed by several groups, employing differing techniques. Filter-binding assays have been used to measure the kinetics of initiation on the λP_r promoter (Roe & Record, 1985; Melançon et al., 1982; Roe et al., 1984, 1985). These assays show the presence of polymerase on DNA but are unable to distinguish between multiple conformations that may be stable under the same experimental conditions. Also, factors such as the binding of RNA polymerase to nonpromoter sites or ends may affect the efficiency of specific filter binding.

Abortive initiation assays, developed by McClure and coworkers (1978), have been used on a variety of systems, including the *lac* operon [for example see Hawley and McClure (1980) and Malan et al. (1984)], to measure the kinetic parameters by measuring the rate of formation of a 2- or 3-mer, the abortive product (induced by adding only the necessary ribonucleotides to produce the short oligomer). This assay, however, has several disadvantages. First and foremost, it is not direct, relying on the measurement of later steps in initiation. Second, it requires the assumption that the steps leading to abortive initiation are identical with those leading to productive initiation and that the steps from open complex to abortive product are not rate limiting. Finally, when experiments on the *lac* operon involving abortive initiation were performed in our laboratory (Liu-Johnson, 1986), different products than those found by McClure were obtained. The RNAs were larger and contained runs of U's making them illegitimate transcripts resulting from some sort of slippage of nascent transcript relative to template, a process about which little is known.

A productive initiation or transcription assay has been used on T7 promoters (Nierman & Chamberlin, 1979; Rosenberg et al., 1982) as well as *lac* and λ promoters (Stefano & Gralla, 1979; Gralla et al., 1980). This method, like the previous one, shares the shortcoming of being indirect. Also, as pointed out by McClure (1979), this technique can yield artifacts due to premature termination or pausing.

[‡] Department of Molecular Biophysics and Biochemistry.
[§] Department of Chemistry.

0006-2960/87/0426-5063$01.50/0

Straney and Crothers (1985) have shown in studies on the *lac* operon that various intermediates, both binary and ternary complexes, can be resolved on nondenaturing polyacrylamide gels. The complexes separated included a tentatively identified closed complex, two open complexes, and three initiated complexes. These have been characterized on the basis of DNase I footprinting, protein content, RNA content, and transcriptional ability. Although the two open complexes differ, they are treated as one complex for kinetic purposes here because their DNase footprints are identical and because they are in rapid equilibrium (Straney & Crothers, 1985; Straney & Crothers, 1986). Using this gel separation method to distinguish the open complex from free DNA, as well as the initiated complex I_u from the open complex, we have measured the kinetics of interconversion of these forms on the *lac* L8UV5 promoter, which is CAP independent. Our method has the advantage of being able to examine the kinetic properties of open and initiated complexes directly, through the use of well-separated, extensively characterized complexes in a gel, without having to make assumptions about later steps in initiation or elongation.

Materials and Methods

Materials. The *lac* L8UV5 fragment, a 203-base-pair fragment containing the *lac* L8UV5 promoter, was purified as described previously (Straney & Crothers, 1985). The DNA was labeled with [α-^{32}P]dATP (Amersham) with the Klenow fragment (New England Biolabs). RNA polymerase was purified as described (Straney & Crothers, 1985). Activity, as measured by open complex formation on gels, was found to be approximately 90%.

Open Complex Formation. Labeled L8UV5 fragment, $C_{final} = 3.8 \times 10^{-10}$ or 7.8×10^{-10} M, was incubated for at least 2 min in a water bath at the required temperature in standard reaction buffer consisting of 40 mM Hepes, pH 8.0, 10 mM $MgCl_2$, 0.1 M KCl, 0.1 mM EDTA, 0.1 mM DTT, and 1 mg/mL bovine serum albumin. The reaction temperature varied by less than 0.2 °C from the desired temperature in all cases. At time $t = 0$, RNA polymerase was added, to give a final concentration between 5×10^{-9} and 3×10^{-7} M. Samples of 20 μL were withdrawn and added to 3 μL of dye buffer (60% sucrose, with bromphenol blue and xylene cyanol) and 1 μL of either 1 mg/mL poly(dA-dT) or 1.6 mg/mL heparin, to compete off any RNA polymerase not in open complex. The dye buffer and competitor were preincubated at the reaction temperature. For time points 30 s or more apart, the samples were loaded on the gel immediately upon addition to dye buffer and competitor. For time points less than 30 s apart, the samples in dye buffer and competitor were incubated at the reaction temperature for up to 3 min, before loading on the gel. This was not found to affect the yield of open complex or free DNA in any way. The samples were loaded onto a 4% polyacrylamide (40/1 acryl/bis) TBE (89 mM Tris–borate, 2.5 mM EDTA, pH 8.4) gel, run on a constant-temperature gel apparatus (Bio-Rad or Hoefer) to maintain the gel temperature at ±1 °C of the reaction temperature. The gel was loaded while running at 200 V and was run subsequently at 150–200 V, until the xylene cyanol was 8.5–10 cm from the top of the gel.

The gels were left on one glass plate and autoradiographed with Du Pont Cronex X-ray film at 4 °C, typically for 18 h. The bands corresponding to the free DNA and open complex were excised with a razor blade, inserted into 1.5-mL plastic tubes, and counted in a scintillation counter by the Cerenkov method. A gel slice of similar size containing neither complex nor free DNA was used as the background. Counting both the open complex and free DNA bands, and using both to determine the fraction of free DNA remaining (the relevant quantity for calculations), yields values independent of variations in sample loading. Data analysis was performed on a VAX computer, using a least-squares program based on a combination of the Marquardt algorithm and the method of steepest descents.

In determining the effect of ribonucleotide triphosphates on open complex formation, the ribonucleotide mixture containing 200 μM GpA, 10 μM ATP, UTP, and OMeGTP[1] was added with the DNA. The OMeGTP allows a product of 6 bases, at most, to be formed. No stable initiated complex, which requires 9–11 bases of RNA (Straney & Crothers, 1985; Carpousis & Gralla, 1985) can form, however. In this way, the effect of nucleotide triphosphates on open complex formation alone can be determined.

Open Complex Dissociation. Labeled L8UV5 fragment (3.4×10^{-10} M) and RNA polymerase (2×10^{-8} M) were incubated at the reaction temperature in the standard reaction buffer for at least 10 min. Heparin ($C_{final} = 160$ μg/mL) was added at $t = 0$, and 20 mL was withdrawn time points, added to 3 μL of dye buffer, and loaded immediately on the gel as described above. To test the heparin dependence, the concentration of heparin was varied from 40 to 320 μg/mL, at both 31 and 37 °C, as described in Results.

Initiated Complex Formation. Labeled L8UV5 fragment ($C_{final} = 3.4 \times 10^{-10}$ M) and RNA polymerase ($C_{final} = 2 \times 10^{-8}$ M) were incubated at the reaction temperature in the standard reaction buffer for at least 20 min. To determine the extent of open complex formation, a 20-μL sample was removed, added to dye buffer and heparin, and loaded on the gel. Open complex formation appeared to be complete (80–90% of the DNA was in open complex) in virtually all cases. A mixture of ribonucleotides ($C_{final} = 200$ μM GpA, 10 μM ATP, UTP, GTP, 20 μM OMeCTP, for example; see Table II for specific experiments) was added at $t = 0$. The OMeCTP prevented the polymerase from incorporating any further nucleotides after C, allowing production of an 11-*mer*, which with the polymerase and DNA, will form a stable initiated complex [see Straney and Crothers (1985)]. Samples of 20 μL were withdrawn at various time points and added to heparin and KCl, to give final concentrations of 67 μg/mL and 0.36 M, respectively. This mixture destroys any open complex present and prevents any new complex formation. The initiated complex, I_u, however, is stable under these conditions. The sample in heparin/KCl was incubated at the reaction temperature for 4.5 min, to allow all remaining open complex to dissociate. Then, 4 μL of sucrose dye buffer was added to the sample, and it was loaded on the gel. Since the initiated complex is unstable on gels at high temperatures (>25 °C), the gel temperature in this set of experiments was 22 ± 1.3 °C regardless of the reaction temperature (Straney & Crothers, 1985).

Results

Formation of the Open Complex. The equilibrium binding constant of closed complex, K_b, and forward rate constant of the isomerization reaction, k_2, were measured by the relaxation kinetic equations of Eigen and deMaeyer (1963). For a two-step reaction

$$R + P \underset{}{\overset{K_b}{\rightleftharpoons}} RP_c \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} RP_o$$

[1] Abbreviations: OMeGTP, 3′-*O*-methylguanosine triphosphate; OMeCTP, 3′-*O*-methylcytidine triphosphate.

Table I: Equilibrium and Kinetic Parameters for Open Complex Formation

T (°C)	K_b (M^{-1})	k_2 (s^{-1})	k_{-2} (s^{-1})	K_2 (k_2/k_{-2})
20	3.8 (±0.6) × 10^7	9.9 (±1.5) × 10^{-3}	4.3 (±0.5) × 10^{-4}	23 (±4.5)
24	6.9 (±1) × 10^7	9.9 (±1.7) × 10^{-3}	2.9 (±0.3) × 10^{-4}	34 (±6.1)
28	1.7 (±1) × 10^7	2.2 (±0.4) × 10^{-2}	2.1 (±0.2) × 10^{-4}	105 (±19)
31	3.4 (±0.5) × 10^7	2.2 (±0.4) × 10^{-2}	2.1 (±0.2) × 10^{-4}	105 (±17)
34	5.0 (±0.5) × 10^7	3.1 (±0.3) × 10^{-2}	1.5 (±0.4) × 10^{-4}	207 (±57)
37	1.9 (±0.3) × 10^7	6.7 (±0.9) × 10^{-2}	1.2 (±0.1) × 10^{-4}	558 (±93)

in which equilibration for the first step is much faster than for the second step, a condition that holds true in this case (Stefano & Gralla, 1982), an exponential decay is seen with relaxation time τ, such that

$$1/\tau = k_{-2} + k_2[\bar{R} + \bar{P}]/(K_b + [\bar{R} + \bar{P}])$$

where $\bar{R}$ and $\bar{P}$ are the concentrations of free RNA polymerase and promoter, respectively, at equilibrium. Under our conditions, which include a large excess of RNA polymerase, $[\bar{R} + \bar{P}]$ is approximately equal (within 5%) to the concentration of RNA polymerase added and is invariant with time. As a consequence, the kinetic process is pseudo first order, and the observed kinetic curve is exponential even for large perturbations from equilibrium. Theoretically, k_{-2} can also be calculated from this equation, but it can be measured directly and more accurately (see below).

$1/\tau$ is the effective rate constant of open complex formation, and from the integrated rate equation, the concentration [P] of free promoter DNA remaining (i.e., not in open complex) at time t is given by

$$[P]/[P]_{t=0} = Ae^{-t/\tau}$$

where $[P]_{t=0}$ is the free DNA at $t = 0$ and A is an amplitude correction factor that allows for the presence of a fast, unresolved reaction component. $[P]/[P]_{t=0}$ is the fraction of free DNA remaining at time t. Therefore, τ values can be determined by measuring the fraction of free DNA at different times, by means of the gel retardation technique, which allows separation of the open complex from the free DNA, as described in Materials and Methods. A representative gel and plot are shown in Figure 1. Data were analyzed by excising bands from the gels and counting them in a scintillation counter. The τ values were determined at different concentrations of RNA polymerase, at six temperatures, varying from 20 to 37 °C. $1/\tau$ is plotted against $[\bar{R} + \bar{P}]$, as shown in Figure 2 for data taken at 34 °C, yielding values for K_b and k_2 from a nonlinear least-squares fitting to the data. Each curve, such as the one shown, has 12–16 points defining it, each point representing a τ value derived from the 8–11 points on an associative time course curve, such as the one shown in Figure 1.

The values of K_b and k_2 measured at the different temperatures are shown in Table I. K_b appears to have no temperature dependence and is scattered about a mean value of 3.8 × 10^7 M^{-1}. k_2, however, increased almost 7-fold over a 17 °C temperature differential. For one set of curves, a mixture containing primer and NTPs was added (see Materials and Methods); these were found to have no measurable effect on either K_b or k_2 (data not shown). From the temperature dependence of k_2, the activation energy of the forward isomerization ($RP_o \rightarrow RP_c$) can be calculated, according to the Arrhenius equation $k = Ae^{-E_a/RT}$. The activation energy, E_a, was found to be 19.3 (±2.7) kcal/mol.

In contrast to the results of Shanblatt and Revzin (1984), no promoter occlusion was seen. The same total amount of open complex, 70–90%, was seen after sufficient incubation (usually 15 min), regardless of the concentration of RNA

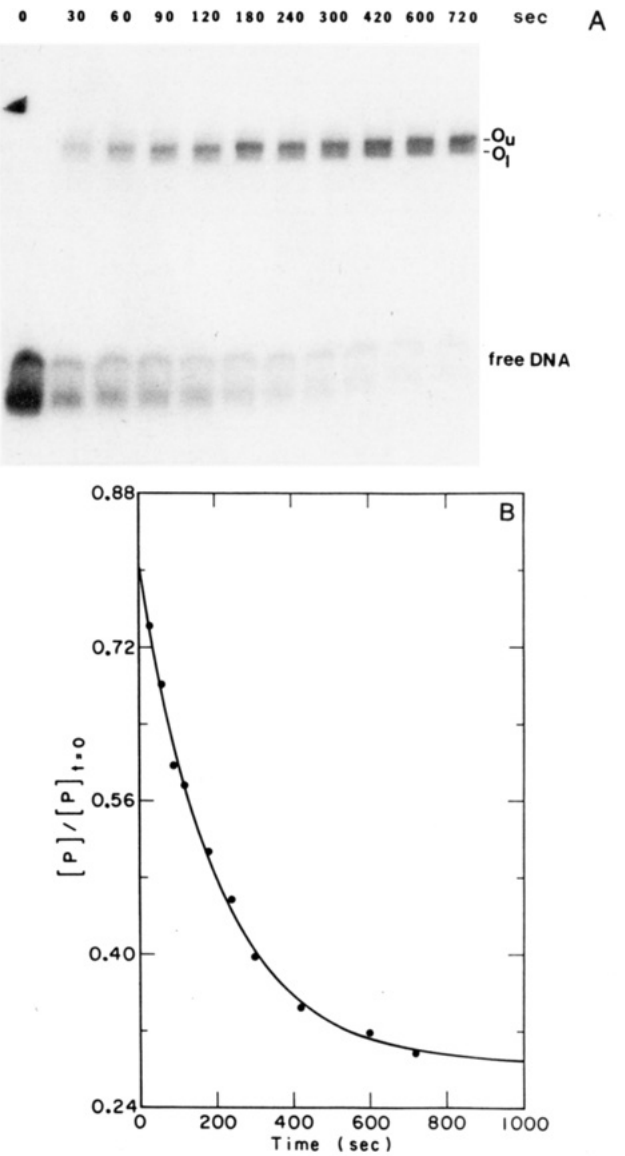

FIGURE 1: (A) Time course of open complex formation. The concentration of RNA polymerase in this reaction was 1.5 × 10^{-8} M; the concentration of DNA was 3.8 × 10^{-10} M. The temperature was 24 °C. The times shown at the top of the lanes indicate seconds after addition of RNA polymerase. O_u and O_l are the upper and lower open complexes, respectively. (B) Fraction of free DNA remaining vs. time, for the gel shown in (A). From this plot, the relaxation time, τ, can be calculated (see Results).

polymerase, which was varied over a wide range (5 × 10^{-9}–3 × 10^{-7} M).

Dissociation of the Open Complex. The dissociative isomerization rate constant, k_{-2}, is an apparent first-order rate constant at fixed conditions of salt concentration and pH. The integrated rate equation is

$$[RP_o]/[RP_o]_{t=0} = Ae^{-k_{-2}t}$$

where $[RP_o]$ is a concentration of open complex at time t and $[RP_o]_{t=0}$ is the concentration of open complex at $t = 0$. The $[RP_o]$ values were sampled by the gel separation method at

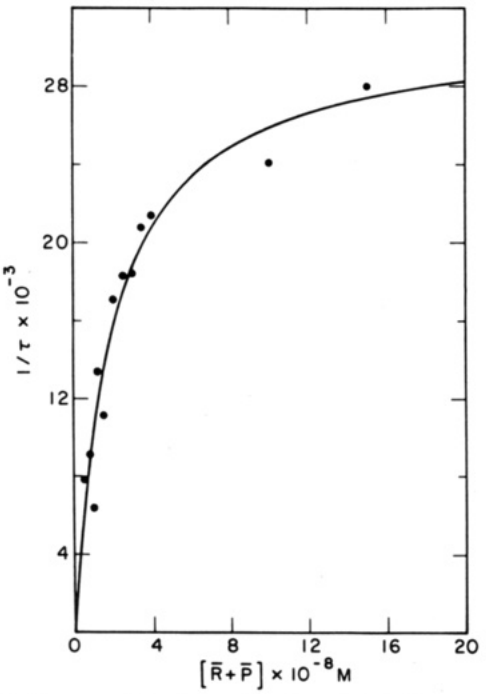

FIGURE 2: Inverse of the relaxation time ($1/\tau$) of open complex formation as a function of the equilibrium concentrations of free DNA and RNA polymerase ([R + P]), at 34 °C. The points on this plot were derived from curves such as the one shown in Figure 1.

different times after adding heparin (which stops open complex formation). A representative gel and associated curve are shown in Figure 3. The values of k_{-2} at different temperatures are collected in Table I. Each is the average of three to four separate determinations. A clear temperature dependence is shown, with the rate of dissociation increasing as the temperature decreases. This gives the activation energy of $RP_o \rightarrow RP_c$, which is −12.6 (±2.7) kcal/mol. From the k_2 and k_{-2} values, the equilibrium isomerization constant, K_2, can be calclalted. Since $K = Ae^{-\Delta H/RT}$, the enthalphy, ΔH, can be determined; ΔH was found to be 31.8 (±3.0) kcal/mol.

A dependence on heparin concentration has been reported by some workers [for example, Cech and McClure (1980)]. However, in our studies, no effect on the value of k_{-2} was seen. At both 31 and 37 °C, the concentration of heparin was varied between 40 and 320 μg/mL. No heparin dependence was found (data not shown).

Addition of individual nucleotide triphosphates or primers (10 μM GTP, CTP, UTP, ATP; 200 μM GpA; 200 μM GTP) had no effect on either the rate of dissociation or the fraction of open complex that eventually dissociated (70% on the average). A mix of ribonucleotides (200 μM GpA; 10 μM ATP, UTP, and OMeGTP) that allow formation of a 6-mer RNA but not a stable initiated complex had the same rate of dissociation. However, 75–80% remained in the open complex, about 2.5 times more than in the absence of these ribonucleotides.

Initiation Step. Formation of the stable initiated complex with bound RNA, RP_i(RNA) (Straney & Crothers, 1985), accompanied by loss of α subunit

$$\text{rNTPs} + RP_o \rightarrow RP_i(\text{RNA}) + \sigma$$

is also a pseudo-first-order process, which gives a rate equation similar to that for open complex dissociation:

$$[RP_o]/[RP_o]_{t=0} = Ae^{-k_{init}t}$$

In these experiments, the RNA polymerase and DNA were allowed to incubate until almost all of the latter was in open

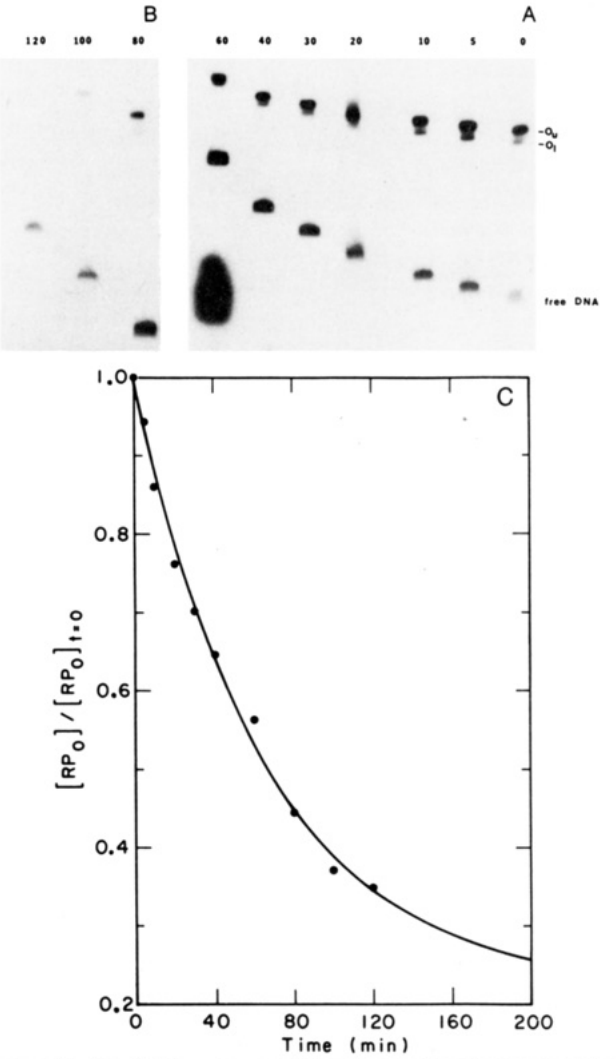

FIGURE 3: (A, B) Time course of the dissociation of open complex. The concentration of RNA polymerase in this reaction was 2.0×10^{-8} M; the concentration of DNA was 3.4×10^{-10}. The temperature was 24 °C. In order to get good separation between the open complexes (O_l, O_u) and free DNA at all times, two gels were used, with the earlier time points shown in (A) and the later time points shown in (B). The times shown above each lane indicate minutes after the addition of heparin. The band under the free DNA in the last lane in (A) is unincorporated label. (C) Fraction of open complex remaining vs. time for the gels in (A) and (B). From this plot, the dissociation rate constant, k_{-2}, can be determined (see Results).

complex. A sample was taken to ensure that this was indeed the case (see Figure 4, lane 1). Then, the rNTPs were added, to allow stable initiated complex formation (see Materials and Methods). At various times, a sample was added to a high-KCl/heparin solution, which destabilizes everything but the initiated complex. This is necessary because the mobility of the initiated complex is very similar to that of the lower open complex (Straney & Crothers, 1985). Therefore, changes in the amount of free DNA on the gel correspond to changes in the amount of open complex left. A gel and its associated curve are shown in Figure 4.

The k_{init} values at six temperatures are collected in Table II. Each represents the average of two or more separate determinations. If a temperature dependence exists, it must be smaller than the scatter in the data ($E_{act} \leq 5$ kcal/mol).

The values reported above are determined with GpA as the initiating dinucleotide. The concentration of GpA was varied between 50 and 200 μM with no great difference in the rate constants. (Indeed, they appear to increase slightly with decreasing initiator, but this is most likely to due to scatter.)

Table II: Initiation Rate Constants

T (°C)	rNTPs[a]	k_{init} (s^{-1})
20	GpA2 and -3	1.4 (±0.3) × 10^{-3}
24	GpA3	1.5 (±0.2) × 10^{-3}
28	GpA1 and -3	1.6 (±0.3) × 10^{-3}
31	GpA1 and -3	1.6 (±0.4) × 10^{-3}
34	GpA1	1.3 (±0.1) × 10^{-3}
37	GpA1	1.8 (±0.5) × 10^{-3}
37	GpA2	2.0 (±0.6) × 10^{-3}
37	GpA3	2.1 (±0.2) × 10^{-3}
37	ApA	1.2 (±0.3) × 10^{-3}
37	G3	6.2 (±1.9) × 10^{-4}
37	G2	6.3 (±1.9) × 10^{-4}
37	G1	6.0 (±0.7) × 10^{-4}

[a] Reaction mixtures: GpA1 = 200 μM GpA, 10 μM ATP, UTP, GTP, 20 μM OMeCTP; GpA2 = 100 μM GpA, 10 μM ATP, UTP, GTP, 20 μM OMeCTP; GpA3 = 50 μM GpA, 10 μM ATP, UTP, GTP, 20 μM OMeCTP; ApA = 200 μM ApA, 10 μM ATP, GTP, UTP, 20 μM OMeCTP; G3 = 250 μM GTP, 250 μM ATP, 10 μM UTP, 20 μM OMeCTP; G2 = 500 μM GTP, 500 μM ATP, 10 μM UTP, 20 μM OMeCTP; G1 = 1 mM GTP, 500 μM ATP, 10 μM UTP, 20 μM OMeCTP.

When ApA was used at 200 μM, the rate constant decreased slightly. When no dinucleotide was used but the concentrations of the first two nucleotides were raised, the rate constant decreased by a factor of at least 2 (see Table II). The rate constant appears independent of the concentration of GTP between 250 μM and 1 mM.

Discussion

We have presented a gel separation method for the direct kinetic assay of intermediates in initiation. With this assay, values of K_b, k_2, k_{-2}, and k_{init} have been measured. In addition, the activation energy and enthalpy of the isomerization step have been determined. The major advantages of our method are that it is direct, the complexes measured are well-characterized (Straney & Crothers, 1985), and the analysis is not dependent on later possibly rate-determining steps in initiation. In the *lac* operon, particularly, there is evidence (Stefano & Gralla, 1979) that a step after open complex formation is the rate-limiting one. Our results are in agreement, with k_{init} being 1–2 orders of magnitude slower than k_2.

Shanblatt and Revzin (1984) have reported a similar method of gel separation. They use only the linear portion of the association curve, which is fairly short. By employing the relaxation equations of Eigen and deMaeyer (1963) under pseudo-first-order conditions, the entire association curve can be used, which allows a more accurate determination of the parameters. The major difference between our work and that of Shanblatt and Revzin, however, lies in the results. They report a decrease in both the rate and extent of open complex formation as the ratio of RNA polymerase to base pairs increases, which they attribute to promoter occlusion by nonspecifically bound RNA polymerase. After a 30-min incubation at 30 °C, at base pair to polymerase ratios of 20:1 to 2:1, they find the extent of open-complex formation decreases from 76% to 37% (Shanblatt & Revzin, 1984). Our reaction conditions, buffers, and concentrations of substrates were very similar to theirs. At both 28 and 31 °C, at base pair to polymerase ratios of 0.4:1 to 7.7:1, we consistently get 80–90% of the DNA in open complex, after a 15-min incubation. In addition, as is shown in Results, we find an increase in the rate of complex formation with increasing RNA polymerase. The off-rate cannot explain this discrepancy—after 30 min, more than 80% of the DNA is still in open complex at 31 °C and this amount is not concentration dependent. One possibility is that the RNA polymerase preparation used by Shanblatt and Revzin contained either a high proportion of core enzyme, which binds DNA tightly, or some other DNA-binding impurity. Our preparation, on the other hand, included a step to separate holoenzyme from core (Gonzales et al., 1977). Only pure holoenzyme was used.

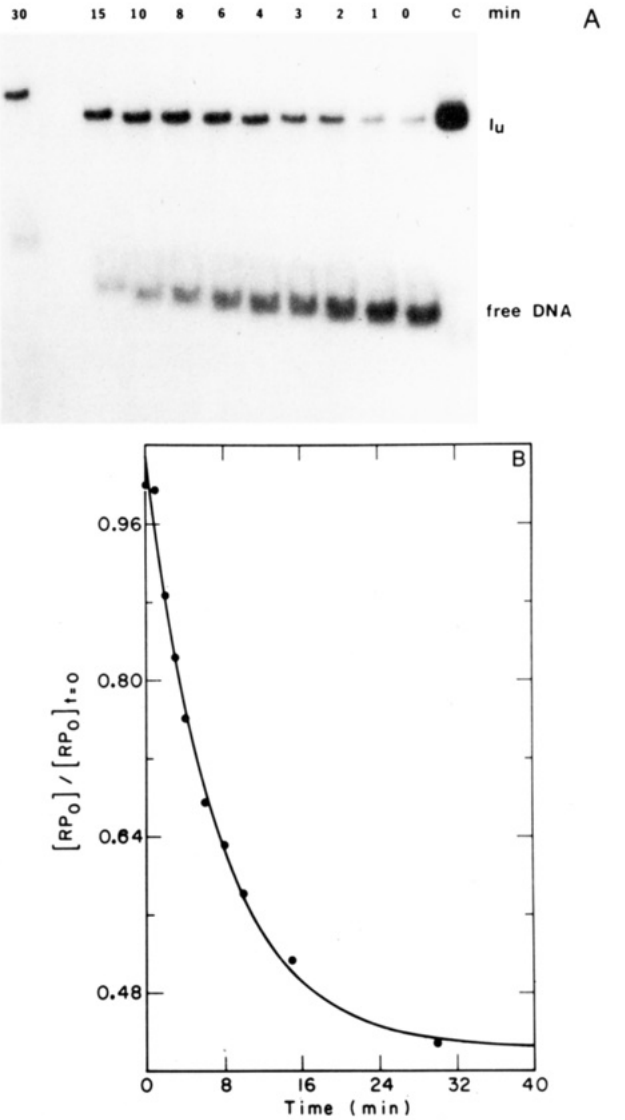

FIGURE 4: (A) Time course of the conversion of open complex to initiated. The concentration of RNA polymerase and DNA are as in Figure 3. The temperature was 37 °C. The final concentration of nucleotides was 50 μM GpA, 10 μM ATP, UTP, and GTP, and 20 μM OMeCTP. I_u is the upper initiated complex, which contains an 11-mer of RNA [see Straney and Crothers (1985)]. Times shown above each lane are minutes after addition of ribonucleotides to preformed open complex. Lane c shows the amount of open complex present before addition of ribonucleotides. Note the similar mobilities of I_u and the open complexes, as mentioned under Results. (B) Fraction of open complex remaining vs. time, for the gel shown in (A). From this plot, the forward rate constant of the conversion of open to initiated complex can be determined (see Results).

Open Complex Formation and Dissociation. We found no temperature dependence of K_b. A similar lack of temperature dependence of the binding constant was seen on the λP_r promoter (Roe et al., 1985) with the filter-binding technique. On the T7A1 promoter, no dependence of the associative rate constant of binding was found (Kadesch et al., 1982). However, a temperature dependence at this step has been reported on the *lac* UV5 operon (Buc & McClure, 1985).

Using the abortive initiation assay, Malan et al. (1984) reported values of K_b and k_2 essentially similar to ours, with a 4-fold lower binding constant and a slightly higher k_2. In

later work (Buc & McClure, 1985) the same group reported activation energies for the formation of open complex (as the sum of two intermediate steps $RP_c \rightarrow RP_i$ and $RP_i \rightarrow RP_o$) as 5.5 + 60 = 65.5 kcal/mol. This is much higher than our value, 19 kcal/mol, which is similar to that found by Roe et al. (1985) on the λP_r promoter (20 kcal/mol).

Our results, overall, are similar to those of Roe et al. (1985) and differ significantly from those of McClure and co-workers (Malan et al., 1984; Buc & McClure, 1985). One explanation for this lies in the type of measurement made. The work of McClure is based on a purely kinetic assay, while both our experiments and those of Roe and co-workers measure the protein–DNA complexes themselves.

As regards the temperature dependence of the dissociative isomerization rate constant, k_{-2}, we again find a value of E_a similar to that of Roe et al. (1985) on λP_r (−13 kcal/mol, as opposed to their −9 kcal/mol) and smaller than that of Buc and McClure (1985), who have determined either −23 or −35 kcal/mol, depending on their method of measurement. In a similar vein, the overall enthalpy of the isomerization reaction that we determined (32 kcal/mol) is close to that of Roe et al. (1985) for λP_r (29 kcal/mol) and smaller than that for *lac* UV5 found by Buc and McClure (about 41 kcal/mol).

The negative temperature dependence of the open → closed reaction rate, as pointed out by Buc and McClure (1985), precludes a simple mechanism of isomerization and makes it likely that there is at least one intermediate step between closed and open complexes. The temperature dependence seen here may reflect that due to the unwinding of the DNA. Wang et al. (1977) and Kirkegaard et al. (1983) have shown that the unwinding of DNA by RNA polymerase decreases as the temperature is lowered.

No heparin dependence of k_{-2} was found, between 40 and 320 μg/mL. Other groups have reported varying results. Pfeffer et al. (1977) found dissociation from the T7C promoter to be heparin independent, and Roe et al. (1984) found polymerase on λP_r to be heparin independent between 5 and 100 μg/mL. However, Majors (1977) reported a dissociation rate constant for heparin of 170 M^{-1} min^{-1} for *lac* UV5, and Cech and McClure (1980) found a k_d(heparin) of 150–340 M^{-1} min^{-1}. Under our conditions, if these constants applied, we should have seen a large heparin dependence. Perhaps the discrepancy can be explained by differences in the sodium content of different heparin preparations (heparin is usually sold as a sodium salt). Sodium ions have been found by Roe and Record (1985) to have a very large effect on the kinetics of polymerase-promoter interactions. It is possible that the increase in dissociation attributed to heparin could be due to the accompanying increase in sodium ions.

As shown in Results, we see an effect of ribonucleotide triphosphates on the dissociation of open complex. This only occurs when all components necessary for production of a short (≤6 base) RNA molecule are present. RNAs this short dissociate from the polymerase–DNA complex, as part of the process known as abortive initiation. The effect is not on the rate of dissociation, but on the extent. This suggests that the abortive synthesis may stabilize the open complex making it temporarily much more resistant to dissociation. This is supported by the work of Straney and Crothers (1986) who found that contacts stabilizing the open complex favor abortive transcription. The addition of these ribonucleotides has therefore divided the open complex into two populations—those that are engaging in abortive transcription and those that are not. The former are very resistant to dissociation, while the latter dissociate at the normal rate.

Initiated Complex Formation. As mentioned earlier, Gralla and co-workers (Stefano & Gralla, 1979; Carpousis et al., 1982) found evidence for a slow step occurring after open complex formation in the *lac* UV5 operon and, by a rifampicin challenge technique, determined a k_{init} of 2.8×10^{-3} s^{-1}. Our results, using GpA as the initiating dinucleotide, are similar, yielding a k_{init} of 1.8×10^{-3} s^{-1} at 37 °C. This is slower by a factor of about 37 than k_2 (6.7×10^{-2} s^{-1} at 37 °C), clearly confirming it as the rate-limiting step of this slow-start promoter. Other promoters appear to have different rate-limiting steps; for example, T7 promoters have faster rates of initiation than of open complex formation (Nierman & Chamberlin, 1979).

Transcription can start at more than one site at this operon. Maizels (1970) found that transcription can begin either at nucleotide −1 on the *lac* UV5 promoter, which is a G, or at +1, which is an A. Carpousis et al. (1982) have identified two additional minor start sites as well. This multiplicity is not surprising, given the findings of Grachev et al. (1984) that even deoxyoligonucleotides, encompassing a range of sizes and sites, primarily from −8 to +2, can serve as primers for transcription.

Different initiating dinucleotides caused a change in the rate of initiation. We found GpA to cause faster initiated complex formation than ApA, which in turn initiated faster than high concentrations of GTP + ATP. Majors (1975), by measuring the extent of productive transcription, found GpA to be a much better initiator than ApA. Other groups have found different results—Carpousis et al. (1982) found that if all nucleotides were present, about twice as much synthesis starting from ATP was seen than from GTP. However, this measurement included both productive and abortive products. Perhaps the two start sites are favored for different types of transcription. The transcript starting with G may be more tightly bound to the RNA polymerase–DNA complex, due to the greater stability of G·C as opposed to A·T base pairs, and this may decrease the fraction that dissociates to form the abortive product. The opposite would be true for the transcript starting with A: more abortive product would be produced. This is consistent with both sets of results.

This does not explain why both GpA and ApA cause faster rates of initiated complex formation than GTP and ATP, however. We initially thought that this might be due to a difference in binding or Michaelis constants between the NTPs and GpA, which would imply a concentration dependence. We found no such dependence within the ranges of 50–200 μM GpA and 250 μM to 1 mM GTP (with high ATP present also), as shown in Table II. We conclude, then, that the difference is not due to differences in the binding constants, because we are working at concentrations above the K_m's of the ribonucleotides. This is in agreement with the findings of Shimamoto and Wu (1980), who found the K_m of the initiating dinucleotide on poly(dA-dT) to be 12 μM. It conflicts, though, with findings of other groups (Gralla et al., 1980; Hawley & McClure, 1982) who have reported K_m's of the first dinucleotide to be in the range of 1–2 mM. We can offer no reason for this discrepancy.

One hypothesis explaining the greater rate of initiated complex formation with a dinucleotide than a mononucleotide is that the formation of the first phosphodiester bond is rate limiting. The dinucleotide has this bond preformed and could therefore form the initiated complex more quickly. Experiments are in progress that may shed more light on this point.

Steady-State Initiation Kinetics. Having evaluated the relevant rate and equilibrium constants in the transcription

initiation mechanism (1), we are in a position to examine the predicted kinetics of initiation at the steady state. The relative magnitudes of the constants are illustrated here with the data obtained at 37 °C with GpA as primer:

$$R + P \underset{}{\overset{K_b = 1.9 \times 10^7\ M^{-1}}{\rightleftharpoons}} RP_c \underset{k_{-2} = 1.2 \times 10^{-4}\ s^{-1}}{\overset{k_2 = 7 \times 10^{-2}\ s^{-1}}{\rightleftharpoons}} RP_o \xrightarrow{k_{init} = 1.8 \times 10^{-3}\ s^{-1}} RP_i(RNA)$$

Applying the steady-state assumption to the intermediates RP_c and RP_o leads to the result that the concentration $[RP_o]$ of the open complex is

$$[RP_o] = \frac{k_1 k_2 [R][P]}{k_{init}(k_2 + k_{-1}) + k_{-1}k_{-2}} \quad (2)$$

The values of the rate constants above show that to a reasonably good approximation we can neglect k_{-2} relative to k_{init} and we can also take $k_{-1} \gg k_2$. With these approximations, eq 2 reduces to

$$[RP_o] = \frac{k_1 k_2}{k_{init} k_{-1}}[R][P] = \frac{K_b k_2}{k_{init}}[R][P]$$

Assuming that all subsequent reactions in transcription are more rapid than the step limited by k_{init}, the rate of initiation-limited transcription is given by $k_{init}[RP_o]$. Expressing the free promoter concentration in terms of the total P_T corrected for $[RP_o]$ ($[RP_c]$ is negligible), we obtain after standard algebra the expression in eq 3 for the rate of production of RNA chains.

$$\text{rate} = \frac{K_b k_2 P_T [R]}{1 + (K_b k_2 / k_{init})[R]} \quad (3)$$

Equation 3 predicts that when [R] is small, corresponding to limiting polymerase, the rate is given by $K_b k_2 P_T[R]$. In this circumstance, the initiation step does not directly influence the rate of transcription, which is set by the rate of forming open complex. On the other hand, when [R] is large, corresponding to excess polymerase, the rate is given by $k_{init}P_T$; as expected, the saturating rate is limited by the rate of the initiation step. The rate rises to half its limiting value when [R] is equal to $k_{init}/K_b k_2$, or about 10^{-9} M with the 37 °C numbers given above. There may be a substantial regulatory advantage in the ability to switch the nature of the rate-determining factors depending upon the level of polymerase produced by the cell. In the present case, the switch would be between rate limitation at the level of open complex formation at low levels of polymerase and at the initiation step when polymerase is saturating. It is plausible that different promoter sequences regulate these two steps, allowing substantial independence in variation of their intrinsic kinetics, a property that may be shared with other slow-start promoters.

In Vivo Transcription Kinetics. As has been pointed out by McClure (1983), the relative rates of in vitro transcription correlate satisfactorily with relative in vivo promoter strength, but there is significant disareement when the rate of initiation is examined quantitatively. For example, the rate of initiation from the *lac* wild-type promoter was estimated by Kennel and Riezman to be about 0.3 s^{-1}, compared to the limiting rate set by the opening rate constant 1.9×10^{-2} s^{-1} (McClure, 1983). Since, as we have verified, k_{init} is significantly smaller than k_2, the discrepancy is even greater when the kinetic barrier posed by the initiation step is included. It seems inescapable that one or more activating factors present in the cell must be missing in the purified systems used to study in vitro initiation.

Registry No. GpA, 6554-00-3; ATP, 56-65-5; UTP, 63-39-8; OMeCTP, 69113-64-0; GTP, 86-01-1; ApA, 2391-46-0; RNA polymerase, 9014-24-8.

References

Buc, H., & McClure, W. R. (1985) *Biochemistry 24*, 2712–2723.

Carpousis, A., & Gralla, J. D. (1985) *J. Mol. Biol. 183*, 165–177.

Carpousis, A., Stefano, J. E., & Gralla, J. D. (1982) *J. Mol. Biol. 157*, 619–633.

Cech, C. L., & McClure, W. R. (1980) *Biochemistry 19*, 2440–2447.

Chamberlin, M. E. (1974) *Annu. Rev. Biochem. 43*, 721–775.

Eigen, M., & de Maeyer, L. (1963) in *Technique of Organic Chemistry* (Freiss, S. L., Ed.) pp 895–1005, Interscience, New York.

Gonzales, N., Wiggs, J., & Chamberlin, M. J. (1977) *Arch. Biochem. Biophys. 82*, 404–408.

Grachev, M. A., Zaychikov, E. F., Ivanova, E. M., Komarova, N. I., Kutyavin, I. V., & Sidelkova, N. P. (1984) *Nucleic Acids Res. 12*, 8509–8524.

Gralla, J. D., Carpousis, A. J., & Stefano, J. E. (1980) *Biochemistry 19*, 5864–5869.

Hawley, D. K., & McClure, W. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6381–6385.

Hawley, D. K., & McClure, W. M. (1982) *J. Mol. Biol. 157*, 493–525.

Kadesch, T. R., Rosenberg, S., & Chamberlin, M. J. (1982) *J. Mol. Biol. 155*, 1–29.

Kennell, D., & Riezman, H. (1977) *J. Mol. Biol. 114*, 1–21.

Kirkegaard, K., Buc, H., Spassky, A., & Wang, J. C. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2544–2548.

Kumar, S. (1981) *Prog. Biophys. Mol. Biol. 38*, 165–210.

Liu-Johnson, H. N. (1986) Ph.D. Thesis, Yale University.

Maizels, N. (1970) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3585–3589.

Majors, J. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4394–4398.

Majors, J. (1977) Ph.D. Thesis, Harvard University.

Malan, P. T., Kolb, A., Buc, H., & McClure, W. (1984) *J. Mol. Biol. 180*, 881–909.

McClure, W. R. (1983) in *Biochemistry of Metabolic Processes* (Lennon, D. L. F., Stratman, F. W., & Zahlten, R. N., Eds.) pp 207–217, Elsevier, New York.

McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171–204.

McClure, W. R., Cech, C. L., & Johnston, D. E. (1978) *J. Biol. Chem. 253*, 8941–8948.

Melançon, P., Burgess, R. R., & Record, M.T. (1982) *Biochemistry 21*, 4318–4331.

Nierman, W. C., & Chamberlin, M. J. (1979) *J. Biol. Chem. 254*, 7921–7926.

Pfeffer, S. R., Stahl, S. J., & Chamberlin, M. J. (1977) *J. Biol. Chem. 252*, 5403–5407.

Roe, J. H., & Record, M. T. (1985) *Biochemistry 24*, 4721–4726.

Roe, J. H., Burgess, R. R., & Record, M. T. (1984) *J. Mol. Biol. 176*, 495–521.

Roe, J. H., Burgess, R. R., & Record, M. T. (1985) *J. Mol. Biol. 184*, 441–453.

Rosenberg, S., Kadesch, T. R., & Chamberlin, M. J. (1982) *J. Mol. Biol. 155*, 31–51.

Saucier, J. M., & Wang, J. C. (1972) *Nature (London), New Biol. 239*, 167.

Shanblatt, S., & Revzin, A. (1984) *Nucleic Acids Res. 13*, 5287–5306.
Shimamoto, N., & Wu, C. W. (1980) *Biochemistry 19*, 842–848.
Stefano, J. E., & Gralla, J. D. (1979) *Biochemistry 18*, 1063–1067.
Stefano, J. E., & Gralla, J. D. (1982) *J. Mol. Biol. 257*, 13924–13929.
Straney, D. C., & Crothers, D. M. (1985) *Cell (Cambridge, Mass.) 43*, 449–459.
Straney, D. C., & Crothers, D. M. (1986) *J. Mol. Biol.* (in press).
Walter, G., Zillig, W., Palm, P., & Fuchs, E: (1967) *Eur. J. Biochem. 3*, 194–201.
Wang, J. C., Jacobson, J. H., & Saucier, J. M. (1977) *Nucleic Acids Res. 4*, 1225–1241.

The Protein Synthesis Initiation Factor 2 G-Domain. Study of a Functionally Active C-Terminal 65-Kilodalton Fragment of IF2 from *Escherichia coli*†

Yves Cenatiempo,‡ Francoise Deville,‡ Jacques Dondon,§ Marianne Grunberg-Manago,§ Christine Sacerdot,§ John W. B. Hershey,∥ Hans Frede Hansen,⊥ Hans Uffe Petersen,*,# Brian F. C. Clark,# Morten Kjeldgaard,# Troels F. M. la Cour,# Kim Kusk Mortensen,# and Jens Nyborg#

Laboratory of Molecular Biology, University of Lyon, 69622 Villeurbanne, France, Institut de Biologie Physico-Chimique, 75005 Paris, France, Department of Biological Chemistry, School of Medicine, University of California, Davis, California 95616, University Department of Clinical Chemistry, State University Hospital, DK-2100 Copenhagen, Denmark, and Division of Biostructural Chemistry, Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark

Received October 24, 1986; Revised Manuscript Received February 20, 1987

ABSTRACT: Protein synthesis initiation factor 2 (IF2) is present in *Escherichia coli* cells as two forms which are expressed from the same gene: IF2α [97.3 kilodaltons (kDa)] and IF2β (79.7 kDa). During isolation, a smaller form, IF2γ, is generated, presumably by partial proteolysis. It has been purified to homogeneity and has an apparent mass of 70 kDa, as determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Immunoelectrophoresis of IF2α and IF2γ shows that IF2γ is immunologically partially identical with IF2α. The sequence of the 15 N-terminal amino acid residues of IF2γ was determined and compared with that of IF2α. The N-terminal amino acid of IF2γ corresponds to Arg-290 of IF2α, suggesting that IF2γ is generated by proteolytic cleavage of the Lys-289–Arg-290 bond of IF2. Assuming a C terminus identical with IF2α, we calculate that IF2γ comprises 601 amino acid residues and has a mass of 64.8 kDa. The truncated protein was tested for activities characteristic of IF2 in three in vitro assays: fMet-tRNA$_f^{Met}$ binding to 70S ribosomes, N-terminal dipeptide synthesis in a DNA-dependent transcription/translation system, and ribosome-dependent GTP hydroly97–7 The specific activities of IF2γ were comparable with, or only slightly less than, thoşe for IF2α, indicating that IF2γ contains the active centers for interaction with fMet-tRNA$_f^{Met}$, ribosomes, and GTP. A central region in the primary structure of IF2 shows extensive sequence homology with a number of GDP-binding proteins and especially with the G-domain of elongation factor Tu (EF-Tu). A predicted secondary structure of this region of IF2 (present in IF2α, IF2β, and IF2γ) is shown. On the basis of our detailed knowledge about the structure of EF-Tu [la Cour, T. F. M., Nyborg, J., Thirups, S., & Clark, B. F. C. (1985) *EMBO J. 4*, 2385–2388; Jurnak, F. (1985) *Science (Washington, D.C.) 230*, 32–36], and the regions of extensive sequence homology between EF-Tu and IF2, a model for the tertiary structure of the IF2 G-domain was built with the aid of a graphic display system. The model obtained is a 17-kDa globular domain approximately 35 Å in diameter.

Initiation of protein synthesis in *Escherichia coli* is promoted by three proteins, initiation factors 1, 2, and 3 (IF1, IF2, and IF3) [for extensive reviews, see Grunberg-Manago (1980) and Maitra et al. (1982)]. IF2 is the largest of the initiation factors and is present in bacterial cells in two size classes, IF2α [97.3 kilodaltons (kDa)] and IF2β (79.7 kDa) (Howe & Hershey, 1982). It interacts with at least three components of the initiation pathway: GTP, fMet-tRNA$_f^{Met}$, and ribosomes. Through these interactions it promotes the binding of fMet-tRNA$_f^{Met}$ to the 30S ribosomal subunit and catalyzes the hydrolysis of GTP after 70S initiation complex formation (Petersen, 1985). The gene for IF2, *infB*, has been cloned and mapped (Plumbridge et al., 1982), and the primary sequence of IF2α has been deduced from the sequence of the gene (Sacerdot et al., 1984). A second translational initiation site in the mRNA accounts for the expression of IF2β from the same gene (Plumbridge et al., 1985). Examination of the amino acid sequence of IF2 revealed an unusual, highly charged repeating structure in the N-terminal half of the molecule and regions in the C-terminal half with homology

† This work was supported by a grant from the CNRS (UA 1176) to Y.C. and F.D.; grants from CNRS (UA 585), MRI (82V1289), INSERM (831.003), and E. I. du Pont de Nemours and Co. to M.G.-M.; grants from the American Cancer Society (NP-70) and from the National Science Foundation (INT83-12982) to J.H.; grants from the Danish Natural Science Research Council (5-17-4.1.04 to B.F.C.C. and H.U.P., 11-5992 to H.U.P., 11-4008 to T.l.C. and J.N., and 5.10.17.05 to J.N.); the Thomas B. Thriges Fund and Direktør Ib Henriksens Fund to T.l.C. and J.N.; and the Niels Bohr Fellowship Committee and Aarhus University Research Fund to H.U.P.

‡ University of Lyon.
§ Institut de Biologie Physico-Chimique.
∥ University of California.
⊥ State University Hospital.
Aarhus University.

0006-2960/87/0426-5070$01.50/0