



# STATE UNIVERSITY OF NEW YORK AT STONY BROOK

COLLEGE OF  
ENGINEERING

Report No. 27

A NONLINEAR MODEL FOR TEMPLET  
REGULATED PROTEIN SYNTHESIS

by

I. Gerst

and

S. N. Levine

OCTOBER, 1964

*Spec*

*TAI*

*.11532*

*no. 27*

*c. 2*

A NONLINEAR MODEL FOR TEMPLET REGULATED PROTEIN SYNTHESIS

I. Gerst and S. W. Levine

Research performed under Office of Naval Research Contract

NONR-4006-00

OFFICE OF SCIENTIFIC RESEARCH CONTRACT

AF AFCSR-667

## SUMMARY

A mathematical model of biological polymerization reactions is formulated and applied to protein synthesis. Under certain reasonable assumptions, including that of a nondissociated enzyme template complex, the resulting set of first order nonlinear differential equations were solved by integral transform techniques. The case where the enzyme template complex dissociates is also treated by perturbation techniques.

The results indicate that the synthesis can be regarded as consisting of three stages (assuming constant monomer concentration): (a) an initial stage in which the concentration of protein  $M_N$  is given by  $M_N \sim Ct^{2N+1}$  (where  $N$  is the degree of polymerization and  $t$ , time) and therefore representing a lag phase; (b) an intermediate stage in which the protein concentration is determined by an exponential polynomial; (c) a third state of a linear increase in protein concentration with time. Comparisons of these predictions with the in vitro data reported by Nirenberg [5] indicate satisfactory agreement.

Methods are also described for determining the rate constants, associated with the polymerization, from the in vitro kinetic data. This important result should facilitate the kinetic analysis of peptide synthesis.

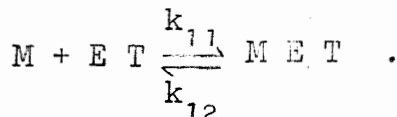
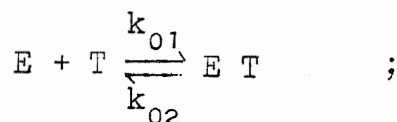
## A NONLINEAR MODEL FOR TEMPLET REGULATED PROTEIN SYNTHESIS

In a previous investigation [1] a linear model was presented for protein synthesis by polyribosomes. Some analytical results for a nonlinear model are given here.

### Relation

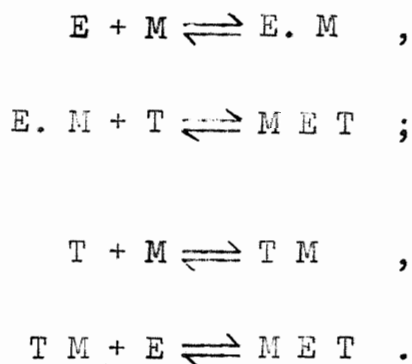
In the following we consider a system of messenger RNA template bound to ribosomes of total concentration  $T_0$  and amino acid acylating enzyme (i.e., transferase) of total concentration  $E_0$ . We assume that  $T_0$  and  $E_0$  remain constant during the course of the polymerization so that denaturation and hydrolysis are assumed not to occur.

The details of complex formation between the polymerase, the messenger RNA and the amino acid-adaptor RNA compound is not known. In the following we assume that the initial complex is formed by the following reactions



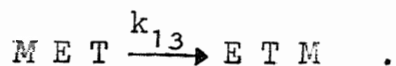
The first expression describes the complex formation between the free polymerase (E) and the messenger RNA (T) (presumed adsorbed onto a ribosome) with a rate constant  $k_{01}$  for the association and  $k_{02}$  for the dissociation. The amino acid initially present as the amino acid-acyl-adaptor RNA compound, designated by M, then forms the complex M E T, as is indicated by reaction (2a). If subsequent studies

ate that instead of (1a) and (2a), the significant reactions

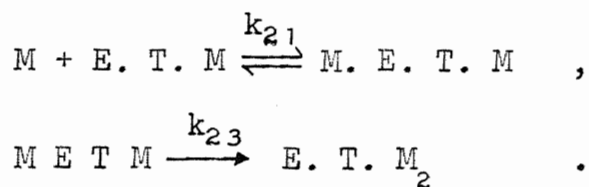


the present formalism will be altered in detail only but the  
 al methods described here are still applicable.

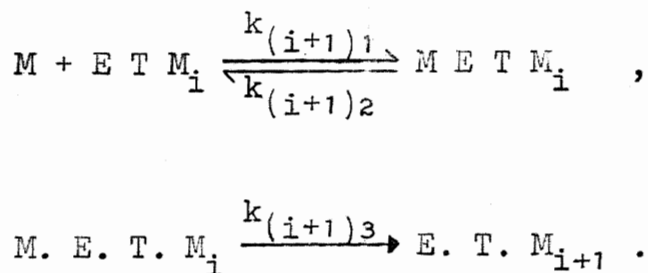
Following the formation of MET, any internal readjustment of  
 required to initiate the polymerization may be represented by



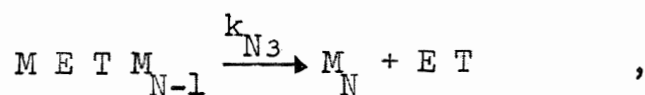
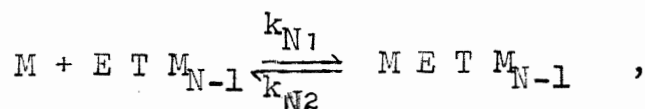
such transformation is necessary, then the present analysis  
 es with  $k_{13}=0$ . The polymerization reaction subsequent to  
 may now be written as



eneral, after (1a) we have a sequence of the following pairs  
 eactions

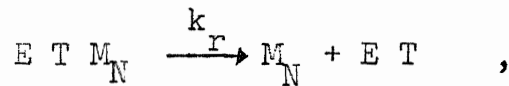
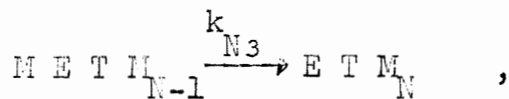


ion (6a) represents the reversible adsorption of the amino adaptor RNA molecule while (7a) represents (with the exception of  $i=2$ ) the irreversible formation of the peptide bond. It should be noted that the above formulation assumes that the enzyme-peptide complex remains irreversibly associated during the course of protein synthesis. Evidence supporting a tightly bound complex has been provided by Tissiers et al [3]. These workers report that during protein synthesis the most active ribosome fraction (70 S fraction) from E. Coli forms an undisassociated complex with the newly formed protein. They postulate that the protein is released by dissociation of the 70 s ribosome fraction into 50 s and 30 s fractions which do not irreversibly bind the protein. Kurland and Lingel [4] also found it necessary to postulate the existence of a special release factor in the case of protein synthesis on rabbit liver ribosomes. It appears from such results as these that a release step is required at the termination of the synthesis. If  $N$  represents the total number of amino acids in the complete protein then the terminal reactions may be presented by



where the release occurs concomitant with the addition of the last amino acid.

Alternatively, it may be that the addition of the amino acid occurs first and that this is followed by release:



$k_r$  is the rate constant associated with the release of bound in.

sis

It is convenient to summarize the mechanism discussed above.

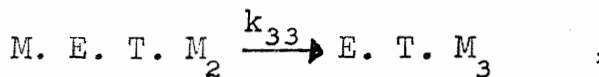
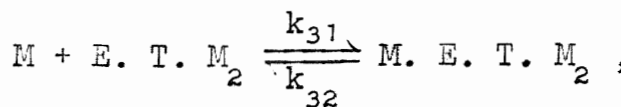
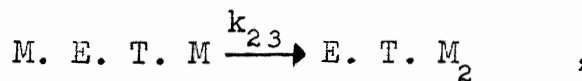
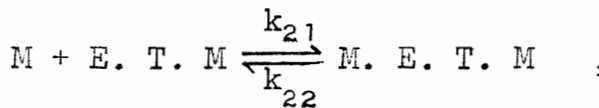
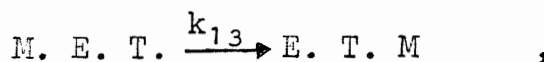
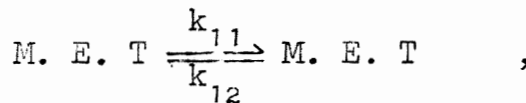
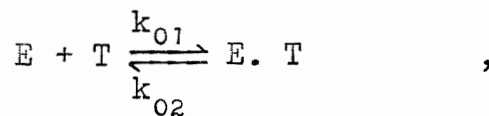
= Concentration of nucleic acid template.

= Concentration of enzyme.

= Concentration of monomer.

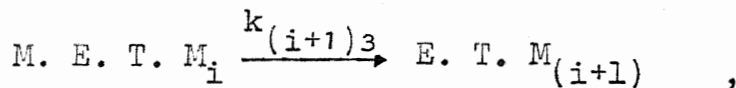
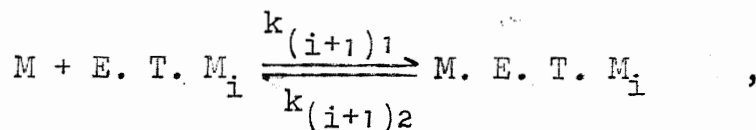
= Concentration of polymer consisting of  $i$  monomers.

Then the system of reactions is given by the set of equations:



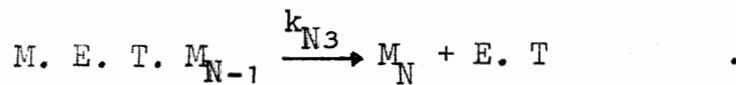
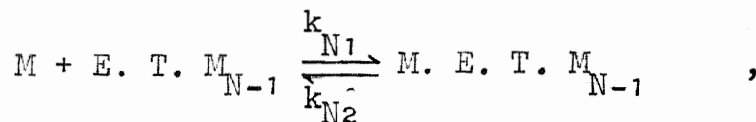
.....

In general, after the first equation in the above we have  
 sequence of the following pairs of reactions:



where  $i = 0, 1, \dots, N - 1$ , and  $M_0$  is to be interpreted as the null  
 member (a chain having no monomers).

If  $N$  is the total number of monomers in the chain, then the last  
 reactions are



Here  $M_N$  is the complete protein.

$$\text{Write } x_i = M. E. T. M_{(i-1)} \quad (i \geq 1) ,$$

$$y_i = E. T. M_i \quad (i \geq 0) .$$

Then the system of O.D.E's corresponding to the reactions (1)

(2) is

$$\frac{dE}{dt} = -k_{01}(E)(T) + k_{02}y_0 ,$$

$$\frac{dT}{dt} = -k_{01}(E)(T) + k_{02}y_0 ,$$



$$\frac{dy_0}{dt} = k_{01}(E)(T) - k_{02}y_0 - k_{11}My_0 + k_{12}x_1 + k_{N3}x_N \quad ,$$

.....

$$\frac{dx_i}{dt} = k_{i1}My_{i-1} - (k_{i2} + k_{i3})x_i \quad (i=1, 2, \dots, N-1),$$

$$\frac{dy_i}{dt} = k_{i3}x_i - k_{(i+1)1}My_i + k_{(i+1)2}x_{i+1} \quad ,$$

.....

$$\frac{dx_N}{dt} = k_{N1}My_{N-1} - (k_{N2} + k_{N3})x_N \quad ,$$

$$\frac{dM_N}{dt} = k_{N3}x_N \quad .$$

This system will be considered under the assumptions that  $M$  is constant and the  $k_{i1}, k_{i2}, k_{i3}$  are all non-negative numbers. Note the last equation may be considered apart from the rest of the system.

The system (3) has certain first integrals which are now defined. These may be used to eliminate some of the variables and thus reduce the order of the system.

If all the equations in (3) except the second and the last are added together, we get

$$\begin{aligned} \frac{dE}{dt} + \frac{dy_0}{dt} + \sum_{i=1}^{N-1} \left( \frac{dx_i}{dt} + \frac{dy_i}{dt} \right) + \frac{dx_N}{dt} &= -k_{11} My_0 + k_{12} x_1 + k_{N3} x_N \\ &+ \sum_{i=1}^{N-1} (k_{i1} My_{i-1} - k_{(i+1)1} My_i - k_{i2} x_i + k_{(i+1)2} x_{i+1}) \\ &+ k_{N1} My_{N-1} - (k_{N2} + k_{N3}) x_N = 0 \end{aligned}$$

Thus

$$4) \quad E + y_0 + \sum_{i=1}^{N-1} (x_i + y_i) + x_N = C_1, \quad ,$$

where  $c_1$  is a constant of integration.

Similarly, if we add all the equations in (3) except the first and last, then

$$5) \quad \frac{dT}{dt} + \frac{dy_0}{dt} + \sum_{i=1}^{N-1} \left( \frac{dx_i}{dt} + \frac{dy_i}{dt} \right) + \frac{dx_N}{dt} = 0 \quad .$$

Thus we have the integral

$$6) \quad T + y_0 + \sum_{i=1}^{N-1} (x_i + y_i) + x_N = C_2 \quad ,$$

where  $c_2$  is another constant of integration.

Note: If we assume that all the  $x_i$  and  $y_i$  are zero initially, then  $c_1 = E(0) = E_0$  and  $c_2 = T(0) = T_0$ , the initial amounts of enzyme and template, respectively.

The system (3) does not appear to have any other simple first integrals besides (4) and (6) and furthermore, in general, no closed-form solution seems to exist. However, in the special case where  $k_{02} = 0$ , a closed form can be derived and this will be given in the next section.

2. Protein Synthesis in the Case  $k_{o_2} = 0$

---

Let 
$$u = y_0 + \sum_{i=1}^{N-1} (x_i + y_i) + x_N .$$

Then the two integrals (4) and (6) become

) 
$$E + u = c_1 , \quad T + u = c_2 .$$

In (3), with  $k_{o_2} = 0$ , add all the equations except the first two and the last. Then

$$\frac{dy_0}{dt} + \sum_{i=1}^{N-1} \left( \frac{dx_i}{dt} + \frac{dy_i}{dt} \right) + \frac{dx_N}{dt} = k_{o_1}(E)(T) .$$

Writing this in terms of  $u$  by means of (7), we get

$$\frac{du}{dt} = k_{o_1}(c_1 - u)(c_2 - u) ,$$

an equation which is immediately integrable since the variables are separated.

Carrying through the integration, we find that

$$u = \frac{c_1 - c_2 c_3 e^{2\beta k_{o_1} t}}{1 - c_3 e^{2\beta k_{o_1} t}}$$

if  $c_1 \neq c_2$ . Here  $\beta = \frac{c_1 - c_2}{2}$  and  $c_3$  is an arbitrary constant of integration.

Note: When  $c_1 = c_2$ ,

$$u = \frac{(c_1 c_3 - 1) + c_1 k_{01} t}{c_3 + k_{01} t}$$

For the initial conditions  $x_i(0) = y_j(0) = 0$ ,  $E(0) = E_0$ ,  $T(0) = T_0$ , (8) yields

$$u = \frac{E_0 T_0 (1 - e^{2\beta k_{01} t})}{T_0 - E_0 e^{2\beta k_{01} t}}$$

with  $\beta = (E_0 - T_0)/2$ .

Then by (7),

$$E = E_0 (T_0 - E_0) e^{2\beta k_{01} t} / (T_0 - E_0 e^{2\beta k_{01} t})$$

$$T = T_0 (T_0 - E_0) / (T_0 - E_0 e^{2\beta k_{01} t})$$

When the solutions for E and T given by (10) are substituted in the third equation of (3), the system (3), starting with third equation, becomes a non-homogeneous linear system of O.D.E's having constant coefficients. Since the initial conditions are all zero, the solution may be worked out easily using Laplace Transforms.

Let

$$\mathcal{L}[x_i(t)] = X_i(s), \quad \mathcal{L}[y_i(t)] = Y_i(s), \quad \mathcal{L}[M_N(t)] = m(s)$$

Then transforming (3) leads to following system of linear algebraic equations for the transforms:

$$s Y_0 = -k_{11} M Y_0 + k_{12} X_1 + k_{N3} X_N + k_{01} \mathcal{L}[(E)(T)]$$

$$s X_i = k_{i1} M Y_{i-1} - (k_{i2} + k_{i3}) X_i \quad (i = 1, 2, \dots, N-1)$$

$$s Y_i = k_{i3} X_i - k_{(i+1)1} M Y_i + k_{(i+1)2} X_{i+1}$$

$$sX_N = k_{N1} M Y_{N-1} - (k_{N2} + k_{N3}) X_N \quad ,$$

$$s m(s) = k_{N3} X_N \quad ,$$

or

$$(s + k_{11} M) Y_0 = k_{12} X_1 + k_{N3} X_N + k_{01} \mathcal{L}[(E)(T)] \quad ,$$

.....

$$(s + k_{i2} + k_{i3}) X_i = k_{i1} M Y_{i-1} \quad ,$$

12)  $(i = 1, 2, \dots, N-1)$

$$(s + k_{(i+1)1} M) Y_i = k_{i3} X_i + k_{(i+1)2} X_{i+1} \quad ,$$

.....

$$(s + k_{N2} + k_{N3}) X_N = k_{N1} M Y_{N-1} \quad ,$$

$$s m(s) = k_{N3} X_N \quad .$$

Omitting the last equation in (12) for the moment, and pairing other equations successively, we have as a typical pair (except for the first pair)

$$(s + k_{(i+1)1} M) Y_i - k_{(i+1)2} X_{i+1} = k_{i3} X_i \quad ,$$

$$(i = 1, 2, \dots, N-1)$$

$$-k_{(i+1)1} Y_i + (s + k_{(i+1)2} + k_{(i+1)3}) X_{i+1} = 0 \quad .$$

Then solving this set in terms of  $X_i$  ,

13) 
$$Y_i = \frac{k_{i3} (s + k_{(i+1)2} + k_{(i+1)3})}{f_{i+1}(s)} X_i \quad ,$$

14) 
$$X_{i+1} = \frac{k_{(i+1)1} k_{i3} M}{f_{i+1}(s)} X_i \quad ,$$

where  $f_{i+1}(s) = (s + k_{(i+1)1} M)(s + k_{(i+1)2} + k_{(i+1)3}) - k_{(i+1)1} k_{(i+1)2} .$

The first pair of equations in (12) yields

$$15) \quad Y_0 = \frac{\{k_{N3} X_N + k_{01} \mathcal{L}[(E)(T)]\} \{s + k_{12} + k_{13}\}}{f_1(s)},$$

$$16) \quad X_1 = \frac{k_{11} M \{k_{N3} X_N + k_{01} \mathcal{L}[(E)(T)]\}}{f_1(s)}$$

Equation (14) is now iterated starting with  $i = N - 1$ , and the value of  $X_1$  is substituted from (16). This gives

$$X_N = \frac{(k_{N1} k_{(N-1)1} \cdots k_{11}) (k_{(N-1)3} \cdots k_{13}) M^N \{k_{N3} X_N + k_{01} \mathcal{L}[(E)(T)]\}}{\prod_{i=1}^N f_i(s)}$$

which when solved for  $X_N$  results in

$$17) \quad X_N = \frac{k_{01} A B M^N \mathcal{L}[(E)(T)]}{k_{N3} \left[ \prod_{i=1}^N f_i(s) - A B M^N \right]}$$

$$\text{Here } A = \prod_{i=1}^N k_{i1} \quad \text{and} \quad B = \prod_{i=1}^N k_{i3}$$

In view of the last equation of (12), we have, finally

$$18) \quad m(s) = \frac{k_{01} A B M^N \mathcal{L}[(E)(T)]}{s \left[ \prod_{i=1}^N f_i(s) - A B M^N \right]}$$

Eq. (18) may be inverted using the convolution theorem.

If

$$19) \quad F(t) = \mathcal{L}^{-1} \left\{ \frac{k_{01} A B M^N}{s \left[ \prod_{i=1}^N f_i(s) - A B M^N \right]} \right\},$$

then

$$0) \quad M_N(t) = F(t) * (E)(T) \quad ,$$

where E and T are given explicitly by (10). By actual substitution, we can check that the polynomial multiplier of s in the denominator of eq. (19) has  $s = 0$  as a simple zero. It can be shown, using a continuity argument, that all other zeros of this polynomial are in the left half-plane. Thus, assuming these zeros are simple, F(t) is of the form

$$F(t) = at + b + \sum (a_i e^{-\alpha_i t} \cos \beta_i t + b_i e^{-\alpha_i t} \sin \beta_i t), (\alpha_i > 0).$$

The explicit "working out" of the convolution in (20), either numerically or in terms of known functions, is left for a subsequent investigation.

### 3. Perturbation Solution for Small $k_{02}$ (Protein Synthesis)

By assuming that  $k_{02}$  is small in the system of O.D.E's given by (3), that system may be solved by using the standard perturbation procedure, and the solution given as power series in  $k_{02}$  with coefficients which are functions of time. Because of the special structure of the system (3), it turns out that these coefficients may be represented in comparatively simple fashion by means of quadratures.

For simplicity write  $\lambda = k_{02}$ . If we add all the equations in (3) except the first two and the last, we get

$$\frac{dy_0}{dt} + \sum_{i=1}^{N-1} \left( \frac{dx_i}{dt} + \frac{dy_i}{dt} \right) + \frac{dx_N}{dt} = k_{01}(E)(T) + \lambda y_0 \quad ,$$

or in terms of the variable  $u$  introduced in Section 2,

$$21) \quad \frac{du}{dt} = k_{01} (c_1 - u)(c_2 - u) + \lambda y_0 .$$

Denote  $k_{01} (c_1 - u)(c_2 - u)$  by  $v$ . Proceeding with the perturbation technique, write

$$22) \quad u = u_0 + u_1 \lambda + \dots + u_n \lambda^n + \dots ,$$

$$23) \quad y_0 = y_{00} + y_{10} \lambda + \dots + y_{n0} \lambda^n + \dots ,$$

where the  $u_n$ , and  $y_{n0}$  are functions of time. Then we may write

$$24) \quad v = k_{01} (c_1 - u)(c_2 - u) = \sum_{n=0}^{\infty} v_n \lambda^n ,$$

where, by virtue of (22) the  $v_n$  may be expressed in terms of the  $u_n$  as follows:

$$25) \quad \begin{aligned} v_0 &= k_{01} (c_1 - u_0)(c_2 - u_0) \\ v_n &= k_{01} \left\{ -(c_1 + c_2) u_n + (u_0 u_n + u_1 u_{n-1} + \dots + u_n u_0) \right\}, \quad n \geq 1 \end{aligned}$$

If (22), (23) and (24) are substituted in (21) and coefficients of like powers of  $\lambda$  are equated, the following system of D.E's results

$$26) \quad \begin{aligned} u_0' &= v_0 , \\ u_1' &= v_1 + y_{00} , \\ &\dots \dots \dots \\ u_n' &= v_n + y_{(n-1)0} . \\ &\dots \dots \dots \end{aligned}$$



In (26), the  $v_n$  are to be replaced by their expressions in terms of the  $u_n$  given by (25). We now find similar expressions for the  $y_{no}$  in terms of the  $u_n$ . These will also be substituted in (26) so that the latter system will be written completely in terms of the  $u_n$ .

For the case  $\lambda = k_{o2} \neq 0$ , the system of transform equations (11) remains the same except that in the first equation the coefficient of  $Y_o$  becomes  $-k_{11}M - k_{o2}$  in place of just  $-k_{11}M$ . Thus the solution of the system (12) will now yield the following expressions for  $Y_o$  and  $X_N$

$$7) \quad Y_o = \frac{(s + k_{12} + k_{13})(k_{N3} X_N + V)}{g_1(s)}, \quad ,$$

$$8) \quad X_N = \frac{ABM^N V}{k_{N3} \left[ g_1(s) \prod_{i=2}^N f_i(s) - ABM^N \right]} .$$

Here  $V = \mathcal{L}[v(t)] = \mathcal{L}[k_{o1}(E)(T)] ,$

$$g_1(s) = (s + k_{o2} + k_{11}M)(s + k_{12} + k_{13}) - k_{11}k_{12}M ,$$

and A, B,  $f_i(s)$  have the meaning assigned to them in Section 2.

Replacing  $X_N$  in (27) by using (28), we get

$$9) \quad Y_o = \frac{(s + k_{12} + k_{13}) \prod_{i=2}^N f_i(s) \cdot V}{g_1(s) \prod_{i=2}^N f_i(s) - ABM^N} ;$$

or, since

$$\begin{aligned} g_1(s) &= (s + k_{11}M)(s + k_{12} + k_{13}) - k_{11}k_{12}M + \lambda(s + k_{12} + k_{13}) , \\ &= f_1(s) + \lambda(s + k_{12} + k_{13}) , \end{aligned}$$

we may write  $Y_0$  (using the geometric series expansion) as a power series in  $\lambda$  in the following way:

$$(30) \quad Y_0 = V \sum_{n=0}^{\infty} [(-1)^n (s + k_{12} + k_{13})^{n+1} h_1(s) \lambda^n] / h(s)^{n+1}$$

where  $h(s) = \prod_{i=1}^N f_i(s) - ABM^N$ ,  $h_1(s) = \prod_{i=2}^N f_i(s)$ .

Let  $G(t) = \mathcal{L}^{-1} \left[ \frac{(s + k_{12} + k_{13}) h_1(s)}{h(s)} \right]$

and  $G_n(t) = \underbrace{G(t) * G(t) * \dots * G(t)}_{(n+1)}$  ( $*$  = convolution),

so that  $G_0(t) = G(t)$ . Then inverting eq. (30) yields

$$\begin{aligned} y_0 &= v * \sum_{n=0}^{\infty} (-1)^n G_n(t) \lambda^n \\ &= \sum_{n=0}^{\infty} v_n \lambda^n * \sum_{n=0}^{\infty} (-1)^n G_n(t) \lambda^n \end{aligned}$$

Thus

$$(31) \quad y_0 = \sum_{n=0}^{\infty} (-1)^n [v_0 * G_n(t) - v_1 * G_{n-1}(t) + \dots + (-1)^n v_n G_0(t)] \lambda^n.$$

Equating the coefficients of  $\lambda^n$  in (31) to the like coefficients in (23), we get the relations

$$(32) \quad y_{n0} = (-1)^n [v_0 * G_n(t) - v_1 * G_{n-1}(t) + \dots], \quad n = 0, 1, \dots$$

Since the  $G_i(t)$  are "known functions" and the  $v_i$  are given in terms of the  $u_i$  by (25), eq. (32) represents the  $y_{n0}$  in terms of the  $u_i$ . Note that the functions  $u_0, u_1, \dots, u_n$ , only, are involved in the expressions for  $y_{n0}$ .

Coming back to the system of D.E's (26), the general equation of this system reads

$$u_n' = v_n + y_{n-1,0} \quad ,$$

or, in view of (25) and (32),

$$(33) \quad u_n' = k_{01} [-(c_1 + c_2) + 2u_0] u_n + \text{terms involving } u_0, \dots, u_{n-1} .$$

Thus assuming  $u_0, \dots, u_{n-1}$  are known, (33) determines  $u_n$ . Since (33) is a linear equation of the first order in  $u_n$ , its solution may be given in closed form in terms of quadratures. In particular, note that the coefficient of  $u_n$  is the same for all these equations. Calling this coefficient  $A$  and the terms free of  $u_n$ ,  $B_n$ , and taking account of the initial conditions  $u_n(0) = 0$ , the solution of (33) may be written in the form

$$(34) \quad u_n = e^{\int_0^t A dt} \int_0^t e^{-\int_0^t A dt} \cdot B_n dt .$$

Since the first equation of (33), giving  $u_0$  is immediately integrable, (cf. Section 2), the induction may be carried through and all succeeding terms  $u_n$  represented in closed form by means of quadratures.

For example, the second equation of (33) would be

$$u_1' = k_{01} [-(c_1 + c_2) + 2u_0] u_1 + k_{01} [(c_1 - u_0)(c_2 - u_0)] * G_0(t) \quad ,$$

since  $y_{00} = v_0 * G_0(t)$  from (32) and  $v_0 = k_{01}(c_1 - u_0)(c_2 - u_0)$  from (25). Here  $u_0$  is to be replaced by the  $u_0$  found in Section 2 (the unperturbed solution).

There remains the development of an expression for the amount of protein,  $M_N(t)$ , in the perturbed solution. From (28) and (29), we have

$$X_N = \frac{ABM^N}{k_{N3}} \cdot \frac{1}{(s+k_{12}+k_{13})h_1(s)} \cdot Y_0 ;$$

and since

$$m(s) = \mathcal{L}[M_N(t)] = k_{N3} X_N / s \text{ by (12), this}$$

gives

$$(35) \quad m(s) = ABM^N \cdot \frac{1}{s(s+k_{12}+k_{13})h_1(s)} \cdot Y_0 .$$

If now,

$$(36) \quad H(t) = \mathcal{L}^{-1} \left[ \frac{1}{s(s+k_{12}+k_{13})h_1(s)} \right] ,$$

then

$$(37) \quad M_N(t) = ABM^N \sum_{n=0}^{\infty} [y_{n0} * H(t)] \lambda^n .$$

Since

$$\mathcal{L}[H(t) * G_0(t)] = \frac{1}{s(s+k_{12}+k_{13})h_1(s)} \cdot \frac{(s+k_{12}+k_{13})h_1(s)}{h(s)} = \frac{1}{sh(s)} ,$$

eq. (37) can be rewritten, in view of (32), as

$$(38) \quad M_N(t) = ABM^N \sum_{n=0}^{\infty} m_n(t) \lambda^n ,$$

where

$$m_n(t) = H_1(t) * (-1)^n [v_n - v_{n-1} * G_2(t) + \dots], \quad n \geq 0$$

$$H_1(t) = \mathcal{L}^{-1} \left[ \frac{1}{sh(s)} \right]$$

and  $v_n$  are to be expressed in terms of the  $u_n$  by (25).

#### 4. Specialization and Evaluation of Results (Protein Synthesis, $k_{02} = 0$ )

---

In this section, the results obtained previously will be worked out in some detail, and certain end formulas will be developed which may be used to connect theory with experiment. In order not to complicate the end relations unnecessarily, certain simplifying assumptions will be made at this point. We suppose that all reaction constants of the same kind are equal to each other. Thus, let

$$k_{11} = k_{21} = \dots = k_{N1} = k_1$$

$$k_{12} = k_{22} = \dots = k_{N2} = k_2$$

$$k_{13} = k_{23} = \dots = k_{N-1,3} = k_3$$

For the present  $k_{N3}$  is not taken equal to  $k_3$  in order to allow for the possibility of a slower reaction time in the last reaction when the protein is separated from the enzyme-template complex. These assumptions seem reasonable in the light of existing knowledge concerning the synthesis process. As a result of this specialization, the polynomials  $f_i(s)$ , ( $i = 1, 2, \dots, N-1$ ) become identical and we can drop the subscript and denote all of them by  $f(s)$  where

$$f(s) = (s + k_1 M) (s + k_2 + k_3) - k_1 k_2 M = s^2 + \delta s + \sigma,$$

and  $\delta = k_1 M + k_2 + k_3$ ,  $\sigma = k_1 k_3 M$ .

For  $f_N(s)$  we have

$$f_N(s) = (s + k_1 M) (s + k_2 + k_{N3}) - k_1 k_2 M = s^2 + \delta' s + \gamma',$$

where  $\delta' = k_1 M + k_2 + k_{N3}$  and  $\gamma' = k_1 k_{N3} M$ .

In this notation, equation (18) becomes

$$18') \quad m(s) = \frac{k_{01} k_{N3} \gamma^N}{k_3 s [f_N(s) [f(s)]^{N-1} - k_{N3} \gamma^N / k_3]} \cdot \mathcal{L}[(E)(T)].$$

The solution for  $M_N(t)$  may then be obtained as usual, by expanding the rational function in (18') into partial fractions, inverting each term and convoluting the inverse transforms with (E). (T).

(a) Asymptotic Behavior for Large  $t$ .

As noted in Section 2, the rational function of  $s$  represented by the fraction in eq. (18') has a double pole at  $s = 0$ ; all other poles are in the left-hand plane. More explicit information regarding the location of the poles can be given if it is assumed that  $k_{N3} = k_3$ . This will be done later to get an indication of the transient behavior of the system for intermediate values of time. For the present, we are interested only in obtaining the asymptotic behavior of the system for large time and we can ignore the poles in the left-half plane since it can be shown that the convolution of terms corresponding to these with (E). (T) approaches zero as  $t \rightarrow \infty$ .

Thus, we are concerned solely with the principal part of

$$k(s) = k_{01} k_{N3} \gamma^N / s (f(s) [f_N(s)]^{N-1} - \frac{k_{N3} \gamma^N}{k_3}) \text{ at } s = 0.$$

We may write

$$k(s) = \frac{A_{-2}}{s^2} + \frac{A_{-1}}{s} + \dots ,$$

and find by standard techniques that

$$A_{-2} = \frac{k_{01} k_{N3} \gamma}{k_3 \delta' + (N-1) k_{N3} \delta} ,$$

$$A_{-1} = - \frac{k_{01} k_{N3} [2\gamma(k_3 + (N-1)k_{N3}) + (N-1)\delta(2k_3\delta' + (N-2)k_{N3}\delta)]}{2[k_3\delta' + (N-1)k_{N3}\delta]^2} .$$

Thus the dominant terms in  $M_N(t)$  for large time are given by

$$M_N(t) = A_{-2} \cdot t * (E)(T) + A_{-1} \cdot 1 * (E)(T) + o(1) .$$

Evaluating the convolutions, we find

$$t * (E)(T) = \frac{1}{k_{01}} \left\{ \frac{(E_0 + T_0)}{2} t + \frac{1}{k_{01}} \log \left( \frac{E_0 - T_0}{E_0 e^{k_{01}\beta t} - T_0 e^{-k_{01}\beta t}} \right) \right\}$$

$$1 * (E)(T) = \frac{E_0 T_0}{k_{01}} \left( \frac{e^{k_{01}\beta t} - e^{-k_{01}\beta t}}{E_0 e^{k_{01}\beta t} - T_0 e^{-k_{01}\beta t}} \right) .$$

Here, as before,  $\beta = (E_0 - T_0)/2$  .

These expressions may be simplified further if we subdivide the discussion into two cases depending upon whether  $E_0 > T_0$  or  $E_0 < T_0$  .

Thus if  $E_0 > T_0$  then  $\beta > 0$  and we find for large  $t$  that

$$t * (E)(T) = \frac{1}{k_{01}} \left\{ T_0 t + \frac{1}{k_{01}} \log \left( \frac{E_0 - T_0}{E_0} \right) \right\} + o(1) ,$$

$$1 * (E)(T) = \frac{T_0}{k_{01}} + o(1) .$$

In this case, then, we have

$$(39) \quad M_N(t) = \frac{A_{-2}}{k_{01}} \left\{ T_0 t + \frac{1}{k_{01}} \log \left( \frac{E_0 - T_0}{E_0} \right) \right\} + \frac{A_{-1} T_0}{k_{01}} + o(1) .$$

The result for the case  $E_0 < T_0$  may be obtained from this one by interchanging  $E_0$  and  $T_0$  in eq. (39).

We may perform a similar asymptotic analysis for each of the  $x_i$  ( $i = 1, \dots, N$ ) and  $y_i$  ( $i = 0, \dots, N-1$ ). Thus, for our specialization, eq. (17) becomes

$$(17') \quad X_N = \frac{k_{01} \gamma^N}{k_3 [f_N(s) [f(s)]^{N-1} - k_{N3} \gamma^N / k_3]} \cdot \mathcal{L}[(E)(T)] .$$

Using eq. (14) to express  $X_{N-1}$  in terms of  $X_N$ , then  $X_{N-2}$  in terms of  $X_{N-1}$ , etc. we find by induction that

$$(40) \quad X_i = \frac{k_{01} \gamma^i f_N(s) [f(s)]^{N-i-1}}{k_3 [f_N(s) [f(s)]^{N-1} - k_{N3} \gamma^N / k_3]} \cdot \mathcal{L}[(E)(T)] ,$$

( $i = 1, 2, \dots, N$ ).

Equation (13) then gives

$$(41) \quad Y_{N-1} = \frac{k_{01} \gamma^{N-1} (s + k_2 + k_{N3})}{[f_N(s) [f(s)]^{N-1} - k_{N3} \gamma^N / k_3]} \cdot \mathcal{L}[(E)(T)] ;$$

$$(42) \quad Y_i = \frac{k_{01} \gamma^i (s + k_2 + k_3) f_N(s) [f(s)]^{N-i-2}}{[f_N(s) [f(s)]^{N-1} - k_{N3} \gamma^N / k_3]} \cdot \mathcal{L}[(E)(T)] .$$



The dominant term in  $x_i$  and  $y_i$  corresponds to the simple pole of  $X_i$  and  $Y_i$  at  $s = 0$ . Writing

$$(43) \quad X_i = \left[ \frac{A_{-1i}}{s} + \dots \right] \cdot \mathcal{L}[(E)(T)] \quad , \quad (i = 1, \dots, N) \quad ,$$

$$(44) \quad Y_i = \left[ \frac{B_{-1i}}{s} + \dots \right] \cdot \mathcal{L}[(E)(T)] \quad , \quad (i = 0, \dots, N-1) \quad ,$$

we find  $A_{-1N} = k_{01} \delta / \Delta \quad ,$

$$A_{-1i} = k_{01} \delta / \Delta \quad , \quad (i = 1, 2, \dots, N-1) \quad ,$$

$$(45) \quad B_{-1(N-1)} = k_{01} (k_2 + k_{N3}) / \Delta \quad ,$$

$$B_{-1i} = k_{01} \delta' (k_2 + k_3) / \delta \Delta \quad , \quad (i = 0, 1, \dots, N-2) \quad ,$$

where  $\Delta = [k_3 \delta' + (N-1) k_{N3} \delta] \quad .$

Note that both  $A_{-1i}$ , ( $i = 1, 2, \dots, N-1$ ) and  $B_{-1i}$ , ( $i = 0, 1, \dots, N-2$ ) are independent of  $i$ .

Equations (43) and (44) in conjunction with the previously evaluated convolution  $1 * (E)(T)$  imply that as  $t \rightarrow \infty$ ,

$$(46) \quad x_i(t) = A_{-1i} \frac{T_0}{k_{01}} + o(1) \quad , \quad (i = 1, \dots, N) \quad ,$$

$$y_i(t) = B_{-1i} \frac{T_0}{k_{01}} + o(1) \quad , \quad (i = 0, \dots, N-1) \quad ,$$

if  $E_0 > T_0$ . Replace  $T_0$  by  $E_0$  in these formulas if  $E_0 < T_0$ .

(b) Asymptotic Behavior for Small Time.

Equation (18'), written as  $m(s) = k(s) \cdot \mathcal{L}[(E)(T)]$ , leads to the time domain equation

$$(47) \quad M_N(t) = \int_0^t K(t - \tau) \cdot (E)(T) d\tau$$

where  $K(t)$  is the inverse of  $k(s)$ . Since  $\lim_{s \rightarrow \infty} [s^{2N} k(s)] = 0$

$$\text{and } \lim_{s \rightarrow \infty} [s^{2N+1} k(s)] = k_{01} k_{N3} \sigma^N / k_3,$$

it follows from the Initial Value Theorem that  $D^i K(0) = 0$  for  $i = 0, 1, \dots, 2N - 1$  while  $D^{2N} K(0) = k_{01} k_{N3} \sigma^N / k_3$ , ( $D = d/dt$ ).

Then differentiating (47) with respect to  $t$  yields  $D^i M_N(0) = 0$ ,

$$i = 0, 1, \dots, 2N, D^{2N-1} M_N(0) = E_0 T_0 \cdot k_{01} k_{N3} \sigma^N / k_3.$$

Thus, for  $t$  small,

$$(48) \quad M_N(t) = \frac{E_0 T_0 k_{01} k_{N3} \sigma^N}{k_3 (2N + 1)!} t^{2N + 1} + o(1).$$

A similar analysis of eqs. (17'), (40), (41) and (42) yields asymptotic relations for each of the  $x_i$  and  $y_i$ , viz

$$(49) \quad y_i(t) = \frac{E_0 T_0 k_{01} \sigma^i}{(2i + 1)!} t^{2i+1} + o(1), \quad (i = 0, 1, \dots, N - 1),$$

$$(50) \quad x_i(t) = \frac{E_0 T_0 k_{01} \sigma^i}{k_3 \cdot (2i)!} t^{2i} + o(1), \quad (i = 1, \dots, N).$$

(c) Intermediate Values of Time.

The question as to whether the asymptotic results obtained in (a) and (b) above will suffice to delineate the behavior of  $M_N(t)$  or whether additional formulas are necessary for intermediate values of

time is still open. It depends upon the actual values of the characteristic roots, i.e. the poles of  $k(s)$ , and these in turn depend upon the numerical values of the reaction constants. In order to get some indication of the location of the characteristic roots in the complex plane, the case in which  $k_{N3} = k_3$  will now be considered. Here we can solve the characteristic equation explicitly in terms of the reaction constants.

With  $k_{N3} = k_3$ , and  $f_N(s) \equiv f(s)$  the denominator in (18') becomes  $k_3 s [f(s)^N - \gamma^N]$ . Thus, we must solve the equation

$$(51) \quad [f(s)]^N - \gamma^N = 0$$

to get the characteristic roots other than the one provided by the factor  $s$ , i.e.,  $s = 0$ . First, we note that (51) can have multiple roots only if  $N[f(s)]^{N-1} f'(s) = 0$  for  $s$  a root of (51). This leads to the conclusion that the only possible multiple root corresponds to  $s = -\delta/2$ , and this value of  $s$  will be a double root if and only if  $\delta^2 = 8\gamma$  and  $N$  is even. Thus for all practical purposes, the roots of (51) will be simple.

Now let  $\rho = \exp \frac{2\pi i}{N}$  so that  $\rho$  is a primitive  $N$ th root of unity.

Then

$$f(s) = \gamma \rho^j \quad (j = 0, 1, \dots, N-1),$$

and solving this quadratic equation we find the  $2N$  roots of (51) given by

$$(52) \quad s_j = -\frac{\delta}{2} \pm \frac{1}{2} \sqrt{\delta^2 - 4(1 - \rho^j)} \quad (j = 0, 1, \dots, N-1)$$

Note that for  $j = 0$ , we get the roots  $0$  and  $-\delta$ . Thus  $0$  is a double pole of  $k(s)$ . If  $N$  is even and  $j = N/2$ , then  $\rho^{N/2} = -1$  and it can be shown that we again have two real roots. Except for these cases

the roots  $s_j$  are all complex.

Equation (52) expresses the  $s_k$  in terms of the square root of a complex number. As is well-known, this square root may be written in the form  $a + bi$  where  $a$  and  $b$  involve only real square roots. For possible future computational use, we list this alternate form of (52).

$$s_j = -\frac{\delta}{2} \pm \frac{1}{2} (a + bi) ,$$

$$53) \quad a = \sqrt{\frac{a_1^2 + \sqrt{a_1^2 + b_1^2}}{2}} , \quad b = \operatorname{sgn} b_1 \cdot \sqrt{\frac{\sqrt{a_1^2 + b_1^2} - a_1}{2}} ,$$

$$a_1 = \delta^2 + \gamma \left( \cos \frac{2\pi j}{N} - 1 \right) \quad b_1 = 4 \gamma \sin \frac{2\pi j}{N}$$

$$(j = 0, 1, \dots, N - 1) .$$

## 5. Conclusions

The results in Section 4 indicate that the synthesis of protein ( $M_N$ ) in the model under consideration can be regarded as consisting of three stages: (a) an initial stage in which the concentration of protein is given by  $M_N \sim C_1 t^{2N+1}$ , and therefore representing an exceedingly slow buildup of the protein complex quite indistinguishable for awhile from a zero protein concentration; (b) an intermediate state in which  $M_N$  is determined by an exponential polynomial; (c) a final state in which  $M_N \sim C_2 t + C_3$ , i.e. linear growth but with an offset  $C_3$ . The constants  $C_1$ ,  $C_2$  and  $C_3$  are given in Section 4 in terms of the various reaction constants.

Nirenberg [5] has reported the results of certain experimental studies of protein reaction kinetics in vitro. Comparing his actual findings with our theoretical prediction, we see that there is substantial agreement between the two. Nirenberg also shows a fourth stage following the linear one in which the protein concentration approaches an asymptotic value. The latter can be interpreted in the light of our model to reflect the situation resulting from a decay in messenger RNA.

As a first consequence of this agreement between Nirenberg's experimental results and our formulas, it becomes possible to use the former (or the results of experiments similar to them) to determine values for certain of the fundamental reaction constants which underlies the protein synthesis process. For example, the slope ( $C_2$ ) of the straight line in the third stage of synthesis can easily be determined from the experimental data. On the other hand, the formulas in Section 4 (cf. eq. (39)) show that

$$C_2 = C_2(E_0, T_0, M, N, k_{i1}, k_{i2}, k_{i3}).$$

If we conduct the experiment several times with different concentrations of monomer  $M^{(j)}$ , but fixed  $E_0$  and  $T_0$ , and denote the measured slopes by  $C_2 M^{(j)}$ , we get a set of algebraic equations

$$C_2(E_0, T_0, M^{(j)}, k_{i1}, \dots, k_{i3}) = C_2 M^{(j)}$$

which serve to determine some of the reaction constants.

To illustrate this procedure, let us assume, for simplicity, that the specialization of the constants assumed in Section 4 applies

and that, furthermore,  $k_{N_3} = k_3$  and  $E_0 < T_0$ . Then we have

$\gamma = \gamma' = k_1 k_3 M$ ,  $\delta = \delta' = k_1 M + k_2 + k_3$ , and eq. (59), with these values inserted, yields

$$54) \quad C_2 = \frac{k_1 k_3 T_0 M}{N(k_1 M + k_2 + k_3)}.$$

Using two different values of  $M$ , designated as  $M^{(1)}$  and  $M^{(2)}$ , in this relation, we readily find that

$$55) \quad k_3 = \frac{M^{(1)} - M^{(2)}}{T_0} \frac{1}{N} \cdot \left[ \frac{M^{(1)}}{C_2 M^{(1)}} - \frac{M^{(2)}}{C_2 M^{(2)}} \right]^{-1}.$$

Thus it becomes possible to determine the important polymerization constant  $k_3$ . By making assumptions regarding the relative magnitude of the other constants and also by using the  $M_N$  intercept, represented by  $C_3$  in the linear asymptotic form, further reaction constants may be evaluated.

## References

1. Gerst, I. and Levine, S.N.: "Kinetics of Protein Synthesis by Polyribosomes". Engineering Report No. 25. State University of New York, Stony Brook, L.I., N. Y., 1964.
2. Simha, R., Zimmerman, J.M. and Moacanin, J. J. Chem. Phys., 39:1239, 1963.
3. Tissiers, A., Schlessinger, D. and Gros, F. Proc. Nat. Acad. Sci., 46:1456, 1960.
4. Webster, G. and Lingel, J.B.: Protein Biosynthesis. Academic Press, New York, N. Y., 1961.
5. Nirenberg et al. Proc. Natl. Acad. Sci., U.S., 48:669, 1962.