Adv. Biophys., Vol. 25, pp.1–43 (1989)

SELECTION AND EVOLUTION OF BACTERIOPHAGES IN CELLSTAT*

YUZURU HUSIMI

Department of Environmental Chemistry, Saitama University, Urawa, Saitama 338, Japan

Can biophysics as a physical science treat molecular evolution phenomena? The answer to this question depends on the aspect being addressed. It cannot explain biological evolution as merely a series of historical events which occurred on the earth. If we follow the tradition of experimental physics, however, we should make a mimic nature in the laboratory, making simple experimental system which clearly expresses the principle of the phenomenon. Once such an experimental system is established for molecular evolution, we will be able to observe its elementary steps repeatedly at will. Then it will be possible to grasp the universal principle governing both the experiment and natural evolution, if it exists, as is true in other examples of physical science.

Let us take, for instance, the question of how a biopolymer can be evolved most rapidly? In other words, what process control is optimal for the functional optimization of a biopolymer? The physical theory of molecular evolution should provide the answer. Achievement of molecular evolution in the laboratory based on the theory will not only confirm the latter but also introduce a new field of biotechnology which will reveal aspects deeper than those recognized heretofore.

Factors determining the rate of molecular evolution in a laboratory can be classified: kinetic parameters, size parameters, and

*Dedicated to Professor Akiyoshi Wada on the occasion of his 60th birthday.

initial conditions. The most important kinetic parameters are mutation rate μ, and selective value (or Darwinian fitness) W, and its dependence on base sequence, that is, the selective value landscape in the base sequence space. Other kinetic parameters are characteristic times of fluctuation or drift of W and μ. The most important size parameters are the length ν of replicating nucleic acid, and its population size $\mathcal{N}$.

Beginning in the 1970's Eigen and his colleagues were involved in development of the theory of quasispecies, which is most appropriate for the handling of molecular evolution experiments in the laboratory from the standpoint of physics (*1–6*). According to Eigen's theory, as ν becomes smaller, μ can be set larger; large μ implies the potential of fast evolution. Therefore, we should use a small DNA or RNA as a working replicon in order to observe and analyze laboratory evolution. The above mentioned kinetic parameters depend on environmental conditions. Therefore, the most fundamental experiments in molecular evolution are the measurement of these parameters (μ, W, and its landscape in the base sequence space) under well-controlled environmental conditions, using a small replicon with controllable population size. In 1967 Spiegelman and his colleagues (*7–9*) began experimenting on *in vitro* Darwinian evolution using bacteriophage Qβ RNA or smaller derivatives of it, the replication reaction of which is catalyzed by Qβ replicase. In one of their experiments they found that the original length of Qβ RNA was shortened by population changeovers due to advantageous deletion mutants. Under test tube condition, protein-coding-regions on the RNA did not selectively contribute, but were merely genetic loads. In another experiment these researchers observed that a three-point mutant RNA evolved after 12 transfers in a serial transfer starting from a midi-variant RNA under selection pressure when ethidium bromide, which binds to RNA, was present. These were the first experimental studies of molecular population dynamics at the base sequence level. Sumper and Luce (*10*), and later Biebricher *et al.*(*11–13*) reported *de novo* synthesis and evolution of Qβ-replicase-oriented RNA molecules *in vitro*. Although these *in vitro* systems are a well-defined molecular system, the phenotype which reflects selective value is present on the replicating RNA molecule itself; lacking this, the assignment of phenotype to genotype is impossible.

In 1950 Novick and Szilard (*14*) devised a chemostat, a continuous cluture system for a bacterium. They observed successive population changeovers by mutants of *Escherichia coli* and estimated the mutation

rate(*15*). Continuous culture has greater merit in quantitative studies than the serial transfer of batch culture because the environmental and physiological conditions are easily kept constant and automation can be readily introduced. Although there is much data on mutation rate and some on selective value using chemostats (*16*), results such as those concerning the time-generation paradox(*17*) are not clear. One reason for this may be that bacteria are too complex to study at the DNA base suquence level.

Chemostats have also been used in continuous culture of a plasmid (*16*). A plasmid is small and easily to be handled at the base sequence level, however, it has several drawbacks: segregational instability(*18*), tight coupling to host DNA which causes large mutational background and ambiguity in the assignment of genotype to a phenotype for a multi-copy plasmid.

Small bacteriophages are positioned between bacteria and self-replicating RNA molecules in their level of complexity. The status of DNA technology for phage is nearly the same as for plasmid. A small bacteriophage can now be utilized as a well-defined molecular probe in the complex cell machinery. The advantages of phage over plasmid as a working replicon are: loose coupling to host DNA which provides means of overcoming the large background of host mutations, and the case of assigning genotype to phenotype.

Domingo *et al.* (*19, 20*) observed a dynamic equilibrium between mutation and natural selection in an RNA phage population. They analyzed RNAs from a bacteriophage Qβ stock strain using T1-fingerprints, that is, a two-dimensional separation pattern of RNA fragments digested by ribonuclease T1. The stock strain consisted of a wide variety of mutant phages with the fraction of the major component only 0.17. Selecting a mutant, they performed a serial transfer experiment, and after *ca.* 20 generations obtained the same T1-fingerprint as for the stock strain. They concluded that the natural population of RNA phage forms a dynamic equilibrium distribution of mutants, a quasi-species, due to the high mutation rate. Similar observations were reported for an RNA animal virus population(*21*).

Hobom and Braunitzer(*22*) cultured *E. coli* carrying parasitic bacteriophage fd using a manually operated turbidostat, and studied the effect of long term exposure to weak mutagens. The carrier cell, however, was unstable and serial transfers of phages had to be made into a fresh host.

In this article we review the characteristics of a stable continuous culture system of a bacteriophage called "cellstat", and describe the

local landscape of the selective value in base sequence space. As stated, the latter is some of the most fundamental data on molecular evolution.

I. METHODS OF CONTINUOUS CULTIVATION OF A BACTERIOPHAGE

Kinetic parameters of the molecular evolution of bacteriophages are dependent on environmental conditions. These conditions are neither constant nor controllable in batch cultures or on culture plates. It is desirable to use continuous culture just like a chemostat or a turbidostat for a bacterium. Continuous culture of bacteriophage in a single fermenter is unstable because of the ecological interactions which occur between a host cell and the phage (Fig. 1a); emergence of a phage-resistant host cell is a typical instability (*22–24*).

The first stable continuous culture of a bacteriophage was performed using a cellstat (Fig. 1b) (*25, 26*). A cellstat system consists of a turbidostat and several cellstat vessels. The turbidostat serves as the supplier for a physiologically constant host cell flow. With a "turbidostat" the turbidity of the culture, which is nearly proportional to the population density of the bacterium, is kept constant by a regulated dilution flow(*27*). A well-known relation for the turbidostat(*28*)

$$k_g = D_t = \frac{F_t}{V_t} \tag{1}$$

holds between specific growth rate of the host cell k_g, and the dilution rate $D_t = F_t/V_t$, where F_t is the flow rate and V_t is the volume of the culture. The cellstat is the culture vessel of a bacteriophage in the flow of the host cell. "Cellstat" means, as shown later, that the population density of uninfected host cells is kept constant. Cultivation is done to assure that the dilution rate of a cellstat is much greater than the specific growth rate of a host cell:

$$D_c = \frac{F_c}{V_c} \gg k_g = D_t \tag{2}$$

were F_c and V_c denote the flow rate and the volume of cellstat culture, respectively. Under this condition, a mutant host cell which emerges in the cellstat is washed out rapidly, so that selection of the mutant cell does not occur. Inequality (*2*) is the condition of a fresh host, which means that the mean residence time of the host cell in the cellstat V_c/F_c is much shorter than its generation time $1/k_g$, or that the physiological conditions of the host cell are kept constant by a large

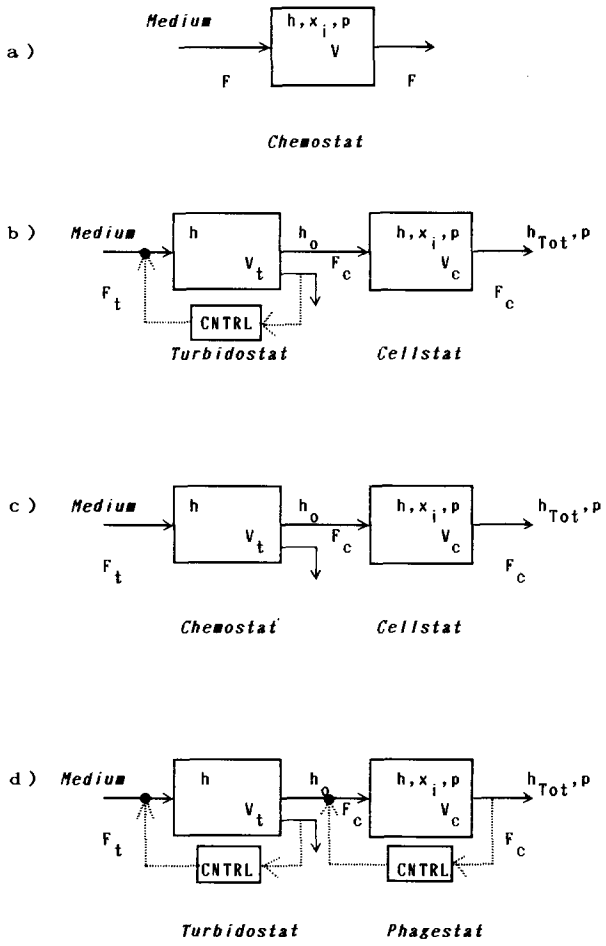

Fig. 1. Continuous culture systems of bacteriophage. Only system(b) is successful. Chemostat or turbidostat is a supplier of the host cell. Turbidostat and phagestat are regulated by an external control device. Chemostat and cellstat are regulated by an endogenous feedback mechanism. F, flow rate; V, volume; $D = F/V$, dilution rate. For cellstat, $D_c \gg D_t$ must be satisfied. h, uninfected host bacterial cell; x_i, infected host bacterial cell; h_{Tot}, total bacteria (uninfected and infected); p, free phage.

influx of these cells produced by the upstream turbidostat. Even under such conditions the bacteriophage is not washed out because the growth rate of the phage is much greater than that of the host cell. Thus we can observe the selection and evolution processes of the phage without interference by host cell mutation. This isolation strategy cannot be applied to plasmid culture, and this is the reason why

phages are more suitable than plasmids as working replicons in laboratory evolutionary studies.

The turbidostat cannot be replaced by a chemostat in this system (Fig. 1c). A chemostat is always operated at the transition phase between the exponential growth phase and the stationary phase. The endogenous regulatory action of population density in the chemostat originates from the buffering action of the large fluctuation between these two phases. This situation is analogous to a 0° C thermostat made of ice-water. In a chemostatic culture, the physiological conditions of the host cell appear to be kept constant macroscopically, but it must fluctuate microscopically. Another feature peculiar to male-specific phage such as fd is that F-pili cease to grow in the phase-transition region used in the chemostat(*29*). An external regulator is therefore required to produce a steady host-cell flow, usually in the exponential growth phase.

The same explanation applies to a phage culture. The cellstat described here is endogeneously regulated by the host-phage interaction. As discussed later, the cellstat becomes unstable as the population density of the phage decreases because its buffering action is weakened. At present, we have no technique to monitor the phage population in real time. Therefore, the completely externally regulated system called "phagestat" (Fig. 1d), which would ideally serve our purpose, is still beyond our capabilities.

II. BACTERIOPHAGES

1. Secretion Type DNA Phages

There is a class of bacterial viruses infecting male *E. coli* known as Ff phage (*30–33*) of which many strains have been isolated from all over the world. M13(34), fd(35), f1(36), and δA(37) were isolated in Munich, Heidelberg, New York, and Tokyo, respectively. Judging from data that indicates that several percent of humans may be carriers of Ff phage (*38*), it is a very popular phage. M13, fd, and f1 are closely related, in fact, the differences in their DNA sequences are less than 3% (*39–42*). A virion or a virus particle is filamentous and contains a single-stranded small circular DNA (length = 6,408 nucleotides). These DNA properties make it possible to sequence the complete genome easily. Thus these bacteriophages are utilized now as a well-defined molecular probe to study population dynamics at the DNA sequence level.

Another characteristic feature of Ff phage is its unique life cycle.

Daughter phages do not kill the host cell, but are secreted through the cell membrane. Ff phage is a parasite rather than a phage. Absence of cell debris keeps a culture vessel clean, making it easy to control environmental conditions.

The initial step of Ff phage infection is the irreversible attachment to the tip of the F pilus (*43*). After this first infection, a pseudo-immunological mechanism inhibits a second infection (*44*). Moreover, *E. coli* cultivated in a synthetic medium has only a few F pili (*45*). For these two reasons, multiple infection is rare in a synthetic medium (*46*). Measurement of the population density of Ff phage is not altered by the multiple adsorption of virions to a host cell.

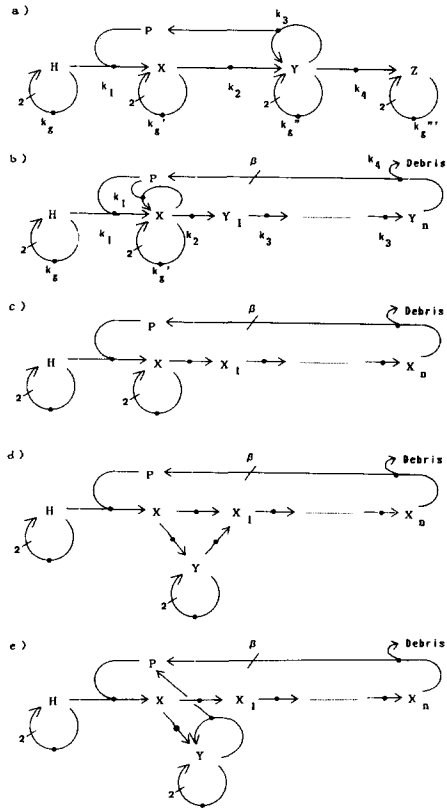

Fig. 2. Chemical reaction models for life cycle of various types of bacteriophage. (a) Ff phage (secretion-type phage). (b) F-specific RNA phage. (c) simple virulent phage. (d) temperate phage. (e) partially virulent phage. H, host cell; P, phage; X, Y, Z, infected cell in various infection stages. Small dots correspond to individual reactions. Short bar crossing an arrow indicates multiplicity of the arrow, for instance, β is burst size.

The life cycle of Ff phage can be expressed by a chemical reaction model as shown in Fig. 2a. P and H denote free phage and uninfected host cell, respectively. The replicative form (RF) DNA is amplified exponentially in a cell of X state. It secretes daughter phages constantly in the Y state. The Z state is an imaginary state to monitor the condition of the fresh host. The population of Z should be negligible in the cellstat. Many other phenomenological kinetic models also describe the host-phage interaction and reproduce the growth curve (*54*).

Filamentousness permits insertion of a foreign gene into Ff phage DNA; in fact, derivatives of M13 are famous cloning vectors in gene engineering (*47*, *48*). This property gives Ff phage an advantage in application to evolutionary molecular engineering (see Chapter IX). Furthermore, filamentousness makes it easy to purify phage particles from a culture solution (*49*). Thus, Ff phage is one of the most applicable working replicons in laboratory evolution experiments.

2. Virulent F-specific RNA Phage

The known RNA phages of *E. coli* are classified into four groups on the basis of the serological and physicochemical properties of their particles, but they are all related (*50*). Qβ phage is one of the most intensively studied organisms in research on the elementary steps of molecular evolution both *in vivo* (*19*, *20*) and *in vitro* (*7–9*).

Contrary to Ff phage, Qβ is a virulent phage. After a 30 min latent period, an infected cell is lysed and thousands of daughter phages burst forth. At first glance, virulent phage seems a poor choice in continuous culture as cell debris will contaminate the vessel. The fraction of debris, however, is very small in cellstat conditions with a high flow rate.

Qβ is a small spherical phage with small single-stranded linear RNA (length = 4,220 nucleotides) (*57*). The initial step of Qβ infection is a reversible attachment to the side of a F pilus (*52*); hundreds of Qβ phages can bind to the pilus. It follows that population measurements of Qβ are much more complicated than of Ff phage. Measurement of plaque forming units should be performed immediately after sampling. According to computer simulation, the difference between free phage and total phage number is small when measured within several minutes after sampling.

The RNA replicase cannot be proofread, so the fidelity of replication of Qβ is poorer than that of DNA replication. A subunit of RNA replicase is coded by the phage's own RNA which may mutate, whereas DNA polymerase is not coded by Ff phage DNA which may also

mutate but without influencing its replicating enzyme. Thus Qβ is able to evolve much faster than Ff phage.

The life cycle of Qβ phage can be expressed by the phenomenological chemical reaction model shown in Fig. 2b. H and P denote an uninfected host cell and free phage, respectively. X denotes an infected host cell in the initial stage. $Y_1 \ldots Y_n$ denote an infected host cell progressively accumulating daughter phages. The number of intermediate stages n determines the latent time and sharpness of burst in the population: $n = 24$ reproduces experimental growth curves(*53*); β denotes burst size; $\beta = 600$–$10{,}000$. On the other hand, for Ff phage with very short latent time, $n = 1$(see Fig. 2a). Reaction of P with X represents the multiple adsorption of the phages to the F pilus.

III. BASIC POPULATION DYNAMICS OF CELLSTAT CULTURE

1. Conditions of Cellstat

We consider hypothetical phage-host interaction which satisfies the following conditions (sufficient conditions of cellstat) (*54*):

1) P (free phage) reacts only with H (uninfected host).
2) Subsequent P-generating reactions consist of one-body reactions.

From condition 1), the dynamics of population density of free phage in a cellstat is given by

$$\frac{dp}{dt} = -D_c p - \sum_i k_i p h + \sum_j \beta_j k'_j x_j, \tag{3}$$

where p, h, x_j are population density of free phage, of an uninfected host cell, and of infected host cell in the j-th infection state, respectively. D_c, k_i, k'_j are dilution rate of cellstat, rate constant of i-th adsorption reaction, and rate constant of j-th burst reaction, respectively. Dynamics of an uninfected host cell is given by

$$\frac{dh}{dt} = D_c(h_0 - h) - \sum_i k_i p h + k_g h, \tag{4}$$

where h_0 and k_g denote input population density of a host cell from a turbidostat and specific growth rate of an uninfected host cell, respectively. From condition 2), dynamics of an infected host cell in the i-th infection state is given by

$$\frac{dx_i}{dt} = \sum_j a_{ij}x_j + k_i ph, \tag{5}$$

where x_i, a_{ij} denote population density and the algebraic sum of various rate constants, respectively. a_{ij} is independent of h_0 and k_g. The steady state solution of Eq. (5), that is, the solution derived from a condition in which the left-hand side equals zero are given by

$$\bar{x}_i = Q_i \bar{p}\bar{h}, \tag{6}$$

where Q_i is a function of rate constants and independent of p, h, h_0, k_g. The steady state solution of Eq. (3) becomes

$$0 = -D_c\bar{p} - \sum_i k_i \bar{p}\bar{h} + \bar{p}\bar{h}\sum_i k_i' Q_i. \tag{7}$$

For $\bar{p} \neq 0$, the steady state solution for $\bar{h}$ is given by

$$\bar{h} = \frac{D_c}{\sum_i k_i' Q_i - \sum_i k_i}. \tag{8}$$

This value is independent of h_0 and k_g. In conclusion, if the life cycle of the phage satisfies conditions 1) and 2), then the population density of an uninfected host cell in the steady state of a cellstat is constant irrespective of fluctuation in the turbidostat. This is the reason the system is called a "cellstat". This denomination is quite parallel to that of "chemostat", in which the concentration of a growth-limiting chemical is constant despite fluctuation in the medium reservoir.

Free phages cannot discriminate the host cell coming from the turbidostat from the host cell generated by cell division(k_g). Therefore, $\bar{h}$ is independent of k_g as well as of h_0.

From Eq. (4), the population density of free phage in the steady state is given by

$$\bar{p} = h_0\left(\frac{\sum k_i' Q_i}{\sum k_i} - 1\right) + \frac{k_g - D_c}{\sum_i k_i}. \tag{9}$$

There is a maximum D_c value above which the steady state $\bar{p}$ becomes negative, corresponding to the washout of the phage. Thus we obtain the condition of existence:

$$D_c \leq D_{c,\max}(h_0). \tag{10}$$

Inequalities (2) and (10) are satisfied simultaneously when the host population density, h_0, is large enough.

2. Secretion Type Phage Ff-family (fd, M13, f1, δA, etc.)
The equations describing the population dynamics of Ff phage in cellstat based on the scheme shown in Fig. 2a are as follows:

$$\begin{aligned}
\frac{dh}{dt} &= D_c h_0 - D_c h - k_1 hp + k_g h, \\
\frac{dx}{dt} &= \quad - D_c x + k_1 hp + k_g' x - k_2 x, \\
\frac{dy}{dt} &= \quad - D_c y \qquad + k_g'' y + k_2 x - k_4 y, \\
\frac{dz}{dt} &= \quad - D_c z \qquad + k_g''' z \qquad + k_4 y, \\
\frac{dp}{dt} &= \quad - D_c p - k_1 hp + k_3 y.
\end{aligned} \tag{11}$$

Here, h, x, y, z, and p are the population density of H, X, Y, Z, and P, respectively; the coefficients have already been defined in Fig. 1b and 2a. Total population density of the host cell defined in Fig. 1 is $h_{Tot} = h + x + y + z$. Equations (11) do not contain the variable corresponding to the growth-limiting factor of *E. coli*, and the specific growth rates take their maximum values, because *E. coli* must be cultured in the exponential growth phase in order to culture Ff phage. In the steady state, we obtain from Eqs. (11),

$$\begin{aligned}
\bar{h} &= \frac{D_c}{k_1 k_2 k_3/(D_c + k_2 - k_g')(D_c + k_4 - k_g'') - k_1}, \\
\bar{p} &= h_0[k_2 k_3/(D_c + k_2 - k_g')(D_c + k_4 - k_g'') - 1] \\
&\qquad - (D_c - k_g)/k_1.
\end{aligned} \tag{12}$$

Equation (12) indicates that the cellstat condition is satisfied. Figure 3a shows the steady state population density as a function of dilution rate at constant h_0. Figure 3b shows it as a function of h_0 at constant D_c. We can see the instability of population near washout condition. Inequalities (2) and (10) are simultaneously satisfied, when $h_0 > 1 \times 10^7$ cells/ml which is calculated for fd phage in our standard experimental conditions. This value is a realistic one; normally $h_0 = 2 \times 10^8$ cells/ml is used.

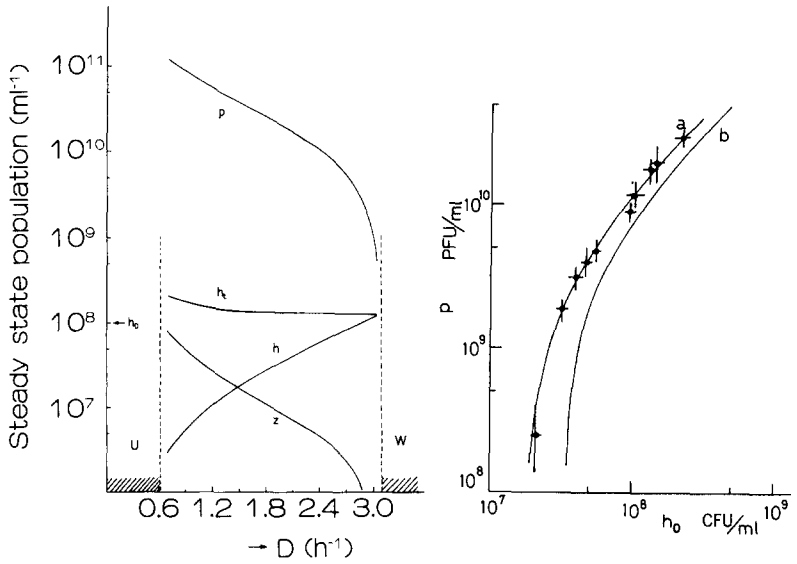

Fig. 3. Steady state population density in a cellstat for Ff phage. (a) Dependence of fd phage culture on dilution rate calculated under the condition of *E. coli*S26, $h_0 = 2 \times 10^8$ CFU/ml, Davis medium at 37°C with aeration. Ordinate: steady state population density of free phage (p), uninfected host cell (h), infected host cell in the last stage of infection (z), and total host cell (h_t). Abscissa: dilution rate of cellstat. U, region of upgrowth instability; W, region of washout instability. (b) dependence on input-population density (h_0) at constant D_c (*25*). (a) $D_c = 2.4\text{h}^{-1}$; (b) $D_c = 3.0\text{h}^{-1}$; ● experimental data.

The stability of steady state solution (Eq. (12)) is examined by the Lyapnov method and direct numerical analysis(*54*). In our standard experimental situation, the solution is asymptotically stable in the sense of Lyapnov. Direct numerical solution of differential Eq. (11) for fd confirm the stability, although there is a negligible damped oscillation.

3. Simple Virulent Phages

A chemical reaction model for a simple virulent phage infection is shown in Fig. 2c. This scheme satisfies the conditions of the cellstat. If Z-state of a secretion-type phage is replaced by debris and its k_4 is set equal to k_3, then a secretion type phage becomes equivalent to a virulent phage. Other characteristic features of virulent phages are the cascading infection states and the large burst size. These are required for generating the long latent period.

4. Temperate and Other Type Phages

The life cycle of a temperate phage may be described by the scheme

shown in Fig. 2d; this scheme also satisfies the conditions of the cellstat.

Some virulent phages have a mixed life cycle, *e.g.*, a fraction of an infected cell is not killed, but secretes daughter phages. This kind of phage can be modelled as shown in Fig. 2e; this scheme also satisfies the conditions of the cellstat.

5. Virulent F-specific RNA Phages and a Pseudo-cellstat

The scheme in Fig. 2b for an F-specific RNA phage does not satisfy the conditions of cellstat, because P reacts not only with H, but also with X. The equations describing the population dynamics in the cellstat are:

$$\begin{aligned}
\frac{dh}{dt} &= D_c(h_0 - h) + k_g h - k_1 ph \\
\frac{dx}{dt} &= -D_c x + k_1 ph - k_2 x \\
\frac{dy_1}{dt} &= -D_c y_1 + k_2 x - k_3 y_1 \\
\frac{dy_i}{dt} &= -D_c y_i + k_3 y_{i-1} - k_3 y_i \\
\frac{dy_{24}}{dt} &= -D_c y_{24} + k_3 y_{23} - k_4 y_{24} \\
\frac{dp}{dt} &= -D_c p - k_1 ph - k_1 px + \beta k_4 y_{24}.
\end{aligned} \tag{13}$$

The steady state solution for uninfected host cell is

$$\bar{h} = \frac{D_c/k_1 + D_c h_0/(D_c + k_2)}{\beta \cdot \dfrac{1}{D_c/k_2 + 1} \cdot \left(\dfrac{1}{D_c/k_3 + 1}\right)^{23} \cdot \dfrac{1}{D_c/k_4 + 1} - 1 + \dfrac{D_c - k_g}{D_c + k_2}}. \tag{14}$$

If we omit terms containing h_0 or k_g, the formula coincides with that of secretion type phage. Existence of terms containing h_0 and k_g means that this is not a cellstat in its mathematical sense. If $k_2 \gg D_c$, then

$$\frac{|k_g - D_c|}{D_c + k_2} \approx \frac{|k_g - D_c|}{k_2} \ll 1, \tag{15}$$

because $k_g \ll D_c \ll k_2$. And

$$\frac{D_c h_0/(D_c + k_2)}{D_c/k_1} \approx \frac{D_c h_0/k_2}{D_c/k_1} = \frac{h_0 k_1}{k_2} \ll 1. \qquad (16)$$

Therefore, the contribution of the terms containing h_0 or k_g is very small, especially for small h_0. For our standard conditions and for RNA phage, the culture vessel can be regarded as a cellstat for $h_0 < 2 \times 10^8$ cells/ml (see Fig. 4b). Thus the vessel may be called a pseudo-cellstat.

We would like to comment briefly on other examples. Multiple infection or a super infection makes the culture vessel a pseudo-cellstat instead of a cellstat. The difference, however, is very small for small h_0.

Other population parameters in the steady state for F-specific RNA phage are

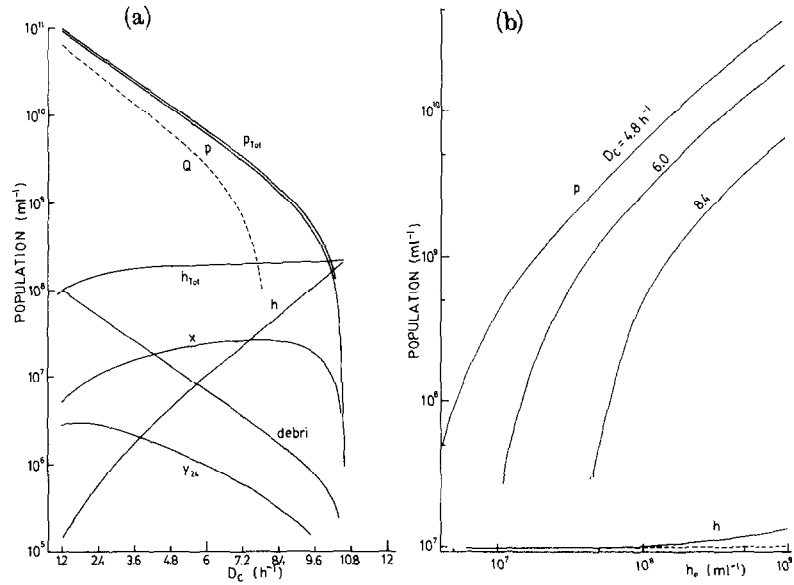

Fig. 4. Steady state population density in a cellstat for F-specific RNA phage. (a) dependence of fr phage culture on dilution rate calculated under the condition of $h_0 = 2 \times 10^8$ CFU/ml in nutrient broth. Kinetic parameters used were obtained from data in ref. *35*. Ordinate: steady state population density of total phage (free and attached to pili) (p_{Tot}), free fr phage (p), uninfected host cell (h), infected host cell (x, y), and total host cell (h_{Tot}). Broken line Q indicates that of Qβ phage culture. (environmental condition is the same as in Fig. 11) Abscissa: dilution rate of cellstat. (b) dependence of fr culture on input-population density (h_0) at constant D_c. Note that cellstat condition (h = const.) is broken at extreme right.

$$\bar{p} = \frac{D_c h_0}{k_1 \bar{h}} - \frac{D_c - k_g}{k_1} \qquad \bar{y}_{24} = \frac{k_3}{D_c + k_4} \bar{y}_{23}$$

$$\bar{x} = \frac{k_1}{D_c + k_2} \bar{p}\bar{h} \qquad \bar{h}_{Tot} = \bar{h} + \bar{x} + \sum_i \bar{y}_i \qquad (17)$$

$$\bar{p}_a = k_1 \bar{p} \bar{x} / D_c$$

$$\bar{y}_1 = \frac{k_2}{D_c + k_3} \bar{x} \qquad \overline{Debris} = k_4 \bar{y}_{24} / D_c,$$

where p_a denotes the population density of the phage adsorbed on F-pili. Total phage number is calculated from $p_{Tot} = p_a + p$.

Steady state populations are plotted against the dilution rate (Fig. 4a) or against the input host population (Fig. 4b). For $D_c = 8\text{h}^{-1}$, the killed host (debris) is within 1% of the host cell population. The difference between p_{Tot} and p is within several percent, so measurement of the infection center (that is, the population of free phage plus infected cells) gives p in the first approximation. A solution of differential Eq. (13) with $D_c = 0$ suggests that we can estimate p in the cellstat if we measure p of the sample solution within several minutes after sampling from the cellstat.

The initial time course of p takes a damped oscillation curve (Fig. 5). This is a typical consequence of nonlinear chemistry and the phenomenon is easily understandable as a series of dilution and burst cycles. The amplitude of the oscillation depends on the sharpness of burst, that is, the number of steps n and burst size β. The period depends on the latent time, k_l and n. And from the other characteristic

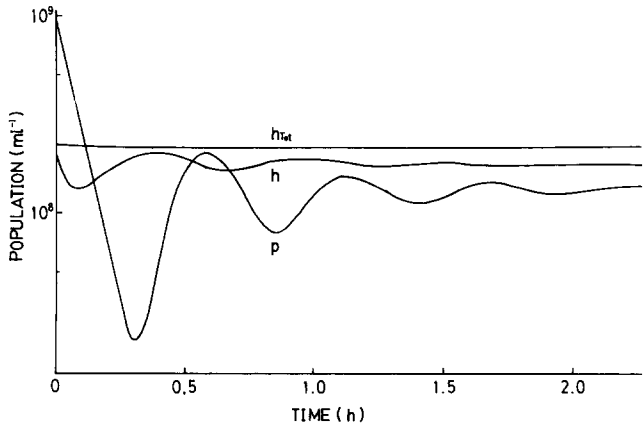

Fig. 5. Initial time course of phage population density in cellstat for F-specific RNA phage (fr phage). A computer simulation using the same kinetic parameters as in Fig. 4.

features of oscillation curves we can get information about the rate constants of the phage-host interaction(*55*).

The stability of the steady state solutions is confirmed by small perturbation-response simulations. An impulse-like change of h_0 causes a tailing response in h_{Tot}, and a delayed broadened impulse response in p. The delay time, of course, reflects the latent time in infection; after the response change, the original values are recovered. Perturbations in D_c cause nearly the same results (*54*). Measurements of aftereffects of these perturbations give the values of the rate constants.

6. Intracellular Process

The above discussions were based on a phenomenological extracellular process. We treated the interaction between two kinds of particles(phage and cell). One (phage) had a single internal state, while the other (cell) had cascadingly changeable internal states. The number density of the particles was observed. We modelled the continuously changing internal state as a discrete one and the transition between discrete states as an elementary chemical reaction. Justification of this treatment is based on the fact that a superposition of many random birth processes becomes a process just like an elementary chemical reaction. For example, the growth of an unsynchronized bacteria population is modelled as a single body autocatalytic reaction, although the division reaction of an individual cell cannot be described as such a simple reaction. Conversely, this is why it is difficult to say anything about the intracellular process from data on the extracellular process. For example, we cannot assign the step number n of infection states to any biochemical reality in an infected cell.

It is necessary to introduce a synchronized culture and analyze intracellular biochemicals in order to study which steps are the most important in the time course of the molecular process in an infected cell. Complexity of such an intracellular process is obvious(*92*). Even if we address the *in vitro* RNA replication process and reduce a flexible macromolecule to a solid particle, a model of the chemical reaction is very complex(*55a, 56–58*).

IV. CELLSTAT APPARATUS

A schematic diagram of the cellstat system is shown in Fig. 6. At

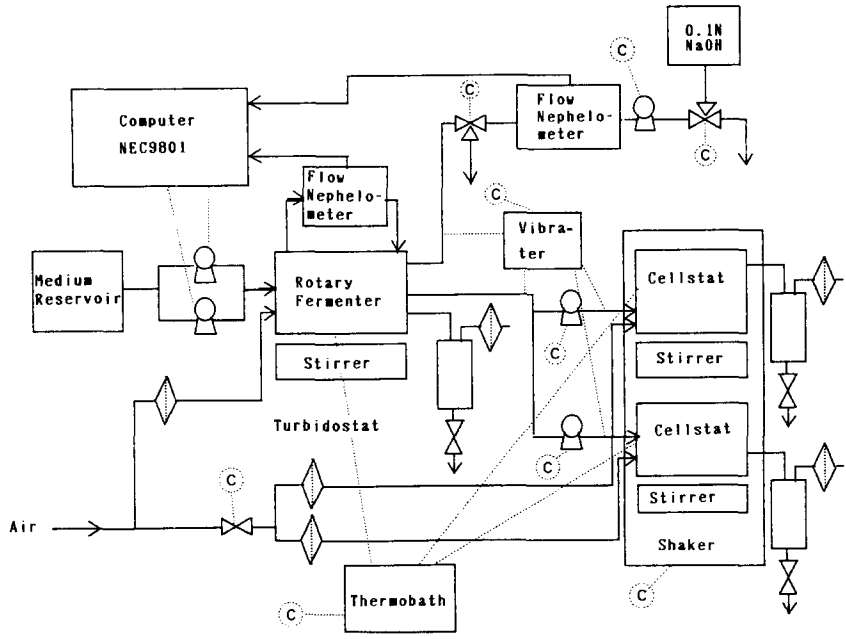

Fig. 6. Schematic diagram of cellstat apparatus. C, computer-controlled.

least two cellstat vessels are connected to a common turbidostat, so that we can determine whether the phenomena in the cellstat are deterministic or stochastic. The main technical problems are how to operate the turbidostat stably and how to prevent wall-growth of the host cell on the surface of cellstat vessels and tubing.

1. Stable Turbidostat

Turbidostat instability originates primarily from wall-growth of the host cell on the surface of the flowcell of the turbidimeter and the deposition of flocks of bacterial cells within this growth. Several techniques have been introduced to prevent this. A time sharing flow nephelometer (*25*) is one example. An alternating cleaning-sampling cycle for the flowcell is introduced in which an 0.1N NaOH solution flows through the flowcell every 4 min for a 2 min period to clean it and to give the offset intensity of scattered light. A computer memorizes this offset value and subtracts it from the value of the scattered light intensity of the sample flow.

Another example is a bubble wall-growth scraper (*26*). We have often observed that air bubbles can remove the wall-growth

when it hits the surface of a vessel; the bubble wall-growth scraper utilizes this phenomenon. Air bubbles are continuously conducted into the sampling port of the turbidostat, which leads to the turbidimeter. The bubbles thus obtained are vigorously circulated. Normally, almost all bubbles are trapped in a bubble separator and returned to the turbidostat through a short circuit path. A solenoid valve cuts the path every 30 min and a large amount of bubbles pass through the flowcell removing the wall-growth on its surface. The computer does not accept turbidimeter data during the passage of the bubbles. The time sharing nephelometer is the more effective of the two techniques in cleaning a flowcell for long-time cultivation, while the bubble scraper is better for short-time cultivation. The latter has the advantage of being a closed system turbidimeter can be used for both batch and continuous culture.

A turbidostat vessel can also be a conventional small stirred tank reactor for short-time cultivation. For long-term cultivation, however, use of a small rotary fermenter is recommended, because the air/liquid interface can scrape off the thick wall-growth and flocks formation is prevented.

The culture medium is fed into the turbidostat by a fine pump and a rough pump. Using the turbidimeter data, the computer calculates the next pump rate every 3 sec, and the specific growth rate of host bacteria every 5 min. The latter is calculated from the following equation:

$$< k_g > = \frac{1}{TV_t} \int_{t-T}^{t} \{F_r(t') + F_f(t')\}\, dt' + \frac{1}{T} \ln \frac{\tau(t)}{\tau(t-T)}, \quad (18)$$

where F_r, F_f, τ, V_t, T denote the flow rate of rough pump, that of fine pump, turbidity, volume of turbidostat and the integration time (normally 1 h), respectively. F_f is proportionally controlled and F_r is set to $V_t \times <k_g>$ minus half maximum of F_f or zero or maximum value, according to the deviation of observed turbidity from the set value.

The constancy of the cell population depends on the potential for wall-growth and flocks formation of the strain. *E. coli*K12 JE3100 strain(*59*) having no 1-pili scarcely grows on the wall. The standard deviation of the turbidity was 0.3% of the set value throughout a 50 h period. The same is true of *E. coli*S26 in the transition phase. *E. coli* S26 in the exponential growth phase has both F-pili and 1-pili and

is apt to adhere to the wall. It requires very careful operation to keep the standard deviation of the turbidity within 1% of the set value.

Experiments on the temperature dependence of the specific growth rate of *E. coli* were carried out, measuring the dilution rate of the turbidostat when the temperature of the thermobath was shifted by an online temperature controller. It was shown that the *E. coli* population immediately follows the change of temperature (up to 1°C/min) between 23–38°C(*60*). This means that the aftereffect of the cooling of *E. coli* during transfer from the turbidostat to cellstats is negligible.

2. Wall-growth

The wall-growth causes an error to the population density data, which are normally measured by sampling a bulk homogeneous culture. Prevention of wall-growth is classified into physical, chemical and biological methods. *E. coli*S26 grows on every surface tested: it adheres to glass, stainless steel, Teflon, and silicone-coated glass. A bacteriostatic silan coating showed little preventive effect. We know of no chemical method for strain S26.

As mentioned the air/liquid interface can remove thick wall-growth. An effective physical method is to pass bubbles through the tubing, and then to shake the vessel vigorously every 5 min. The level sensor of cellstat is locked when the shaker is used intermittently. Teflon tubing used to transfer the host cell from the turbidostat to cellstats is vibrated by a solenoid. Small glass tips which sweep the round bottom of the vessel are also effective against wall-growth on the cellstat.

A biological anti-wall-growth method using a pililess strain was proved to be very effective. The strain JE3100 has only F-pili and hardly adheres to the wall. Its F factor, however, is not as stable. The fimbriae 1-pili are mainly responsible for the adhesion to the wall. The expression of 1-pili genes shows phase variation. Using the *lac*-fusion technique with bacteriophage Mu, a phase-locked mutant VL751 (*61*) was obtained. Mating it with male(F′) strain #42 resulted in the X4-4 strain having F-pili but no 1-pili (*62*). The time course of wall-growth on the cellstat for strain X4-4 was estimated by a series of washing experiments (*63*). In one of these after an appropriate period of cultivation the supply-line to the cellstat was switched from the turbidostat to the medium reservoir.

Then, *E. coli* and phage in bulk began to be washed. The *E. coli* adhering to the wall, however, was not washed but continued to grow. After several hours, the population in bulk, which was observable, reached the minimum, indicating a balance between dilution and supply from the wall-growth. The value of 1×10^5CFU/ml in bulk after 24 h cultivation is one or two orders of magnitude smaller than that of strain S26; it is normally $h_{Tot} = 2 \times 10^8$CFU/ml, so 0.05% of the host cell comes not from the turbidostat, but from wall-growth.

3. Measurement of Phage Population Density

In order to measure population density in cellstat accurately, it is necessary to sample directly from the bulk of a culture using a clean pipette. Continuous sampling through a tube gave errors originating from wall-growth on the tube surface. The tube was used as both sampling port and water-level controller. Air passed actively through the tube and culture droplets seemed to move rapidly in it but island-like wall-growth still took place.

Plaque forming unit (PFU) measurement was performed using the double agar layer plating method (*64*) at 37° C and/or 40° C for Ff phage.

Samples of Ff phages can be stored in 50% glycerine at −20° C. Without glycerine, diluted Ff phage solution had a gradually decreasing titer due to adherence to the wall of the stock tubes.

When population density is very high(10^{10}–10^{12}PFU/ml), a high performance liquid chromatograph (HPLC) using a gel permeation column(G4000SW) and a UV(270nm) detector can be utilized to monitor directly the population density of the phage in the culture. Use of a flow-fluorometer or fluorescence microscope may be other physical methods of counting phage particles stained by fluorescent dye, but the sensitivity of apparatus used is critical.

V. PHAGE POPULATIONS IN THE CELLSTAT

Bacteriophage populations in the cellstat can be classified into three categories:

(1) Genetically invariable pure population.
(2) Genetically invariable mixed population.
(3) Genetically variable population.

Class (1) shows a typical nonlinear chemical kinetics such as an oscillating reaction. Results of computer simulation were discussed in page 15 for virulent F-specific RNA phages. Experimental data on damped oscillation have been obtained for phage Qβ (*55*a, *65*).

The cellstat was operated at the critical dilution rate near washout. The wall-growth was minimized by biological and physical method.

Classes (2) and (3) are discussed in the next sections.

VI. COMPETITION EXPERIMENTS

For a genetically invariable mixed population, we observe the selection process. Mixing two phage strains, we can perform a competition experiment and measure the selection coefficient by analyzing the obtained population changeover curve.

The selection coefficient s is defined as the relative difference between the selective value W of one strain and that of the other strain. For an exponentially growing population, the selective value is equal to the specific growth rate k_g, which is equal to the dilution rate D_c in a continuous culture. Therefore, $s(\boldsymbol{B})$ for strain having sequence $\boldsymbol{B}$ is

$$s(\boldsymbol{B}) = \frac{k_g(\boldsymbol{B}) - k_g(\boldsymbol{B}_0)}{k_g(\boldsymbol{B}_0)} \tag{19}$$

were $k_g(\boldsymbol{B}_0)$ is a specific growth rate of some standard strain having base sequence $\boldsymbol{B}_0$, and $k_g(\boldsymbol{B})$ is that of strain $\boldsymbol{B}$. Bacterial population is a typical example (*66*).

Since a virulent phage having a latent time τ and burst size β grows like $\beta^{t/\tau}$ under a condition of sufficient host, W is equal to $ln\beta/\tau$, if we ignore phage death. From scheme b in Fig. 2, $\tau = 1/k_1h + 1/k_2 + n/k_3 + 1/k_4$. Here we introduce new parameters, partial selection coefficients:

$$\begin{aligned} s_i(\boldsymbol{B}) &= \frac{k_i(\boldsymbol{B}) - k_i(\boldsymbol{B}_0)}{k_i(\boldsymbol{B}_0)} \\ s_\beta(\boldsymbol{B}) &= \frac{ln\,\beta(\boldsymbol{B}) - ln\,\beta(\boldsymbol{B}_0)}{ln\,\beta(\boldsymbol{B}_0)}. \end{aligned} \tag{20}$$

If we can postulate $s_i = s_0$ $(i = 1, 2, 3, 4)$, then

$$s(\boldsymbol{B}) = s_0 + s_\beta + s_0 s_\beta \doteqdot s_0 + s_\beta. \tag{21}$$

Secretion-type phage does not grow exponentially because it has the character of a stamping machine in at least some growth phases. Critical dilution rate $D_{c,max}$ (Eq. (10)), above which phage washout occurs, may be regarded as selective value. This is true in the two examples mentioned above. For Ff-phage under our standard experimental conditions, $D_{c,max}$ is nearly equal to $\sqrt[3]{h_0 k_1 k_2 k_3}$. There-

fore, $s \fallingdotseq \sqrt[3]{(1+s_1)(1+s_2)(1+s_3)} - 1$. So if we can postulate $s_i = s_0$ for $i = 1, 2, 3$, then s_0 can be regarded as selection coefficient s of the phage. As the quality of experimental data allows only one adjustable parameter, we take s_0 as the selection coefficient for a secretion-type phage in this paper. When the postulate is not true, the partial selection coefficients differ substantially from the true values, but the overall s-value remains nearly the same (*54*).

A model of the competition experiment between two species of closely related Ff phages is shown in Fig. 7 where $s_4 = 0$ is assumed. This assumption is not important because a fraction of the Z-state is very small for large D_c, that is, under the condition of fresh host. We also assume that there is no mixed infection. This assumption was checked experimentally and found allowable (*46*), and theoretically confirmed by the fact mentioned in Section II. The major artefact came from the wall-growth. Curve fitting was performed using a profile of the initial stage of population changeover, where the influence of wall-growth is negligible.

There are various methods of discriminating two phage species in population measurements. A plaque of f1 phage is clearly distinguishable from that of others on a double agar layer plate at 40° C incubation using *E. coli*S26 as an indicator; we were therefore able to use f1 phage as a reference phage. δA and fd are discriminated in nearly the same way. Chimera phage M13mp11 makes a blue plaque on a plate containing X-gal using *E. coli* JM105 as an indicator;

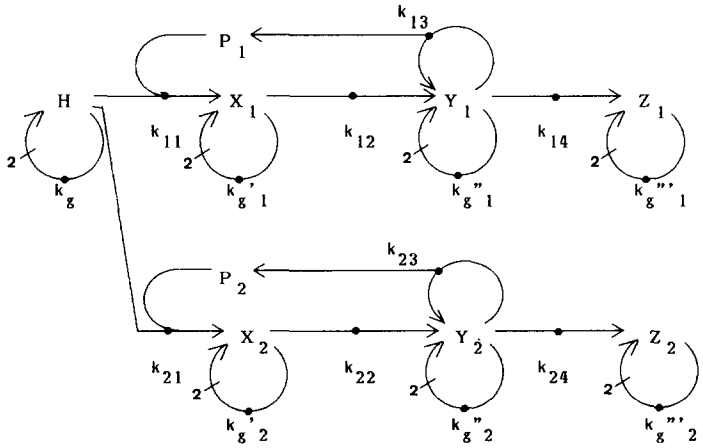

Fig. 7. A chemical reaction model for competition experiment between two kinds of Ff phage. See Fig. 2 legend for detail.

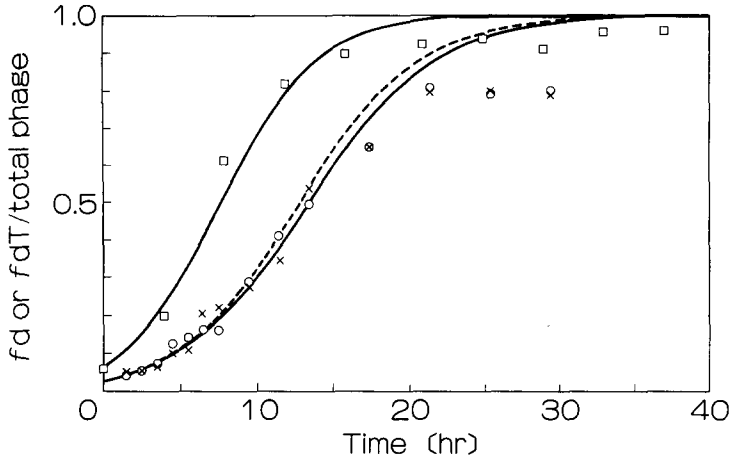

Fig. 8. Competition between fd and f1 in a cellstat. Ordinate: fraction of fd phage population in total population. Abscissa: time after infection. At time 0, a fd:f1 = 1:19 mixed phage population was put into cellstat vessel. □ f1 *vs.* fd wild type, $h_0 = 2.4 \times 10^8$ CFU/ml, $D_c = 2.3h^{-1}$; ○ f1 *vs.* fd mutant (fdT). $h_0 = 2.1 \times 10^8$ CFU/ml, $D_c = 3.0h^{-1}$; ×, same as ○ except $D_c = 3.2h^{-1}$.

comparison between M13mp11 and any wild type phage is thus easily performed. M13mp11 was not used, however, as a universal reference phage, because its selective value was very low and the accuracy of selection coefficient between two wild type phages measured indirectly as the difference of their larger differences from M13mp11 became very poor. Moreover, M13mp11 is unstable, because the inserted portion is easily deleted as described in the next section.

M13 phage is discriminated from other strains by agarose gel electrophoresis, because the charge on the major coat protein is different. This method, however, is not routinely used.

A typical example of a competition experiment is shown in Fig. 8. At time 0, a mixed (f1: fd = 19:1) phage population is put into a cellstat. At every sampling point, clear plaques (for f1) and turbid ring plaques(for fd) are counted on the double layer agar plate incubated at 40° C. These data gave the time course of the fraction of fd phage or fdT phage in the cellstat. Solid curves are fitted theoretical curves which gave the selection coefficient data. Here, fdT is a single point mutant isolated in the stock strain of fd wild type in Takanami's laboratory (*40*). The point mutation G1859 → A1859 is examined by a restriction enzyme digestion. New York phage f1 is defeated by Heidelberg phage fd and its mutant fdT.

The discrepancy between experimental data from the theoretical curves seen at the final stage in the profiles originated from the wall-

growth. Host bacteria attached to the walls of the culture vessels were infected by the phage which was dominant at the initial stage. Therefore, the host on the walls secreted the overcome phage even at the final stage, giving artefact data. Wall-growth does not reproduce, so the final stage of a population changeover curve was not reproducible, but the first half of it was. Roughly speaking, the time interval required for population changeover is proportional to $1/sD_c$. When we perform 100 h continuous culture with $D_c = 3.0\ h^{-1}$, we can measure an s-value as small as 0.003. Accuracy of the data depends on the wall-growth occurring in the experiment.

VII. BIOLOGICAL RELAXATION EXPERIMENTS

We observed the evolution process for a genetically variable population. An artificial mutant or an artificial recombinant DNA may not be stable from the standpoint of the equilibrium between the mutation and the natural selection. The steady state population of the phage in the cellstat increases as the selection coefficient (s) increases. So the time course of increase in the phage population can be regarded as a biological relaxation profile from the artificial nonequilibrium (low s) to a selection equilibrium (high s). Just as in the physical

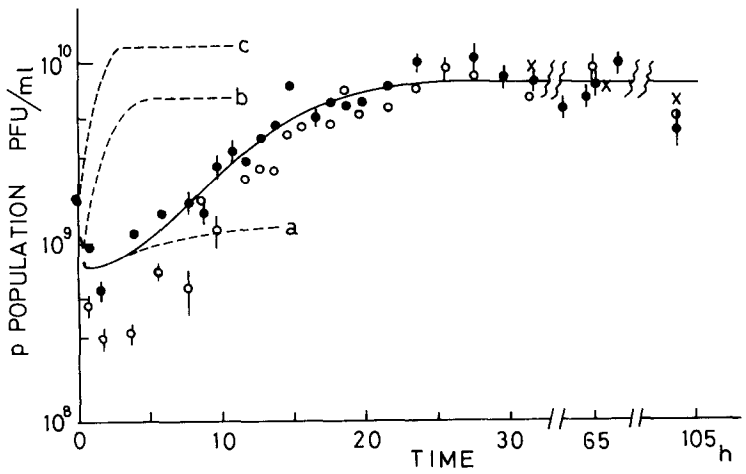

Fig. 9. A biological relaxation starting from fd107 (*25*). Time course of cellstat culture of an elongated fd recombinant fd107. Two independent cultures (○ CI and ● CII) have a common host source ($h_0 = 1.8 \times 10^8$ CFU/ml). $D_{cI} = 3.1h^{-1}$, $D_{cII} = 3.0h^{-1}$. ×, population density data obtained from HPLC. Broken lines are theoretical curves of the invariable pure population having various selection coefficients s. (a) −0.38; (b) −0.18; (c) 0.0, *i.e.*, wild type.

or chemical relaxation case, by analyzing this relaxation profile, we can determine the value of the kinetic parameters for this kinetic process. Thus we can measure the mutation rates and the selective values. Although many genetic engineers have pointed out such a genetic instability, a quantitative treatment of the relaxation process under well-defined environmental conditions has become feasible only with use of the cellstat.

A typical example was shown by the cultivation of a chimera phage fd107(*25*). Filamentous phage fd107(*66*) has a recombinant DNA (10,771 bases) between fd DNA (6,408 bases) and a plasmid pBR322 (4,363 bases). The region from pBR322 is not essential to qualify it as a phage. Therefore, the length of the elongated filamentous phage may decrease gradually as a result of the population changeovers by deletion mutants, just as did $Q\beta$ RNA in Spiegelman's experiment (7). In Fig. 9, two time courses of cellstat cultures CI and CII are shown, into which the same host cell flowed. The difference between them may be partially due to the different dilution rates of the two cellstats. The final steady state was not reached until 25 h after infection. If the population had been genetically invariant and pure, the steady state would have been more rapidly attained as indicated by the broken lines. If the population had been contaminated with the advantageous mutant of a very small population, a quite different step-function-like curve would have been obtained (*54*). These curves should be viewed as the behavior of a highly mutable population. The fact that two independent cellstats gave similar curves means that the phenomena were deterministic, that is, advantageous mutations occurred very frequently.

The following two indications show that these mutations are site specific deletion mutations. Firstly, the length distribution of the filamentous phage determined by electron microscopy shows that the ratio of length of the fd wild type, fd107, and the finally dominant phage (fd107–111) is 1:1.7:1.22. Here, 1.7 corresponds to the degree of polymerization of the elongated phage; 1.22 means that the finally dominant phages were not the wild type fd. This phenomenon is explained by the absence of the specific deletion site (deletion hot spot) at the insertion site of pBR322. Agarose gel electrophoresis of virions gave a single sharp band for fd107–111 between fd and fd107 bands.

Secondly, restriction fragments by *Hae*III and *TthHB*8I of single-stranded DNAs of the finally dominant phage in independent culture vessels showed the same patterns in polyacrylamide gel elec-

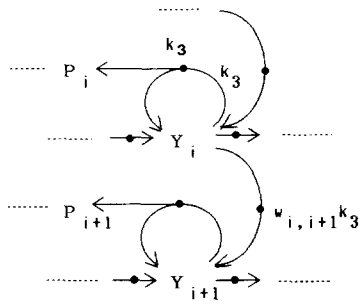

Fig. 10. A chemical reaction model of genetically variable Ff-phage population. Part of the scheme. See also Fig. 2a. P_i, i-th mutant; $w_{i,i+1}$, mutation rate. $i = 1, 2, 3, 4$.

trophoresis. Comparing the pattern of fd107–111 with that of fd107, one can see that many bands are lost, but three new bands have emerged. This is compatible with the postulate that three successive deletions occur in independent vessels.

The solid line in Fig. 9 is a theoretical curve for CII based on the model illustrated in Fig. 10.

The selection coefficients and mutation rates are determined by curve fitting. Obtained are:

$$\begin{aligned} &s(\text{fd107}) = -0.38,\ s(\text{fd107–111}) = -0.13, \\ &w_{ij} = 10^{-4} \sim 10^{-5} \text{ deletion/replication.} \end{aligned} \tag{22}$$

This mutation rate is very high and corresponds to deletion hot spots.

Following single plaque isolation we cultured fd107 several times. Various finally dominant deleted chimera phages have been obtained, but almost every time two independent cellstats, into which the same host cell flowed, have given the same deleted chimera phage judging from length of phage filament and the DNA restriction pattern. This means that this type of the pathway is quite deterministic if the environmental conditions are strictly controlled. Another possibility is the existence of the same initial contamination in both cellstats. Then the phage population would have been an invariable mixed population. This possibility, however, is ruled out for two reasons. Firstly, as mentioned above, a population changeover curve by a very strong advantageous mutant of a very small population in time zero may be quite different from the observed curves. Secondly, generations passed in

preculture were much fewer in number than generations passed in continuous culture. If the mutant emerged in preculture every time, then it also should have occurred in continuous culture. In other words, the phage population was a genetically variable population, which justifies the present treatment.

The engineered cloning vector M13mp11, a chimera between phage M13 and a part of the *lac* operon of *E. coli* plus a synthesized oligo-DNA(*67*) gave similar results. Biological relaxation curves originating from population changeovers by deleted chimera phages reached the final state within 25 h after infection.

A Qβ amber mutant (A-8-7-1B10) was cultured using *E. coli* suppressor strain S26Rle (SuI) as a host(26); the reversion rate of this strain is known to be very high. After sampling 0.1–0.2ml from the cellstat culture (total volume ~25ml) and dilution, one part was plated on *E. coli* Q13(*sup*$^-$) as an indicator bacterium and the other part on *E. coli* QT(*sup*$^+$) as an indicator bacterium. Plaques on the

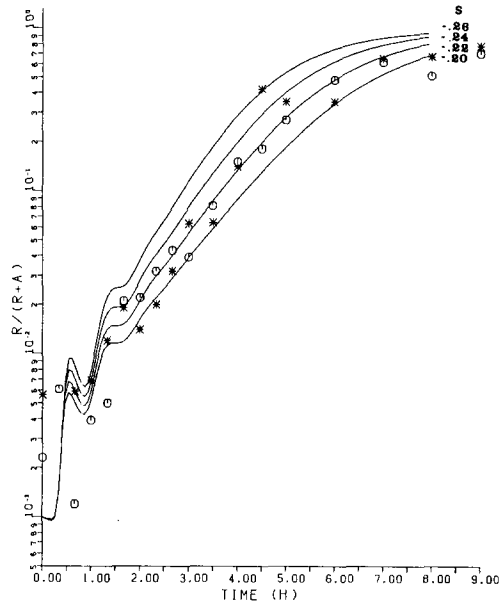

Fig. 11. A biological relaxation starting from a Qβ amber mutant (*26*). Abscissa: time after infection. Ordinate: fraction of revertant in the population. ○, ● data obtained from cellstats #1 and #2, respectively. Multiplicity of infection: 0.3 for cellstat #1, 2.4 for #2. Dilution rate: 2.6h^{-1} for #1, 2.4h^{-1} for #2. Host: *E. coli*S26Rle (*sup*$^+$), $h_0 = 1.9 \times 10^8$ CFU/ml. Solid curves are theoretical curves for cellstat #2 with mutation rate of 0.0005/amber-site/burst and with various selection coefficients s as indicated.

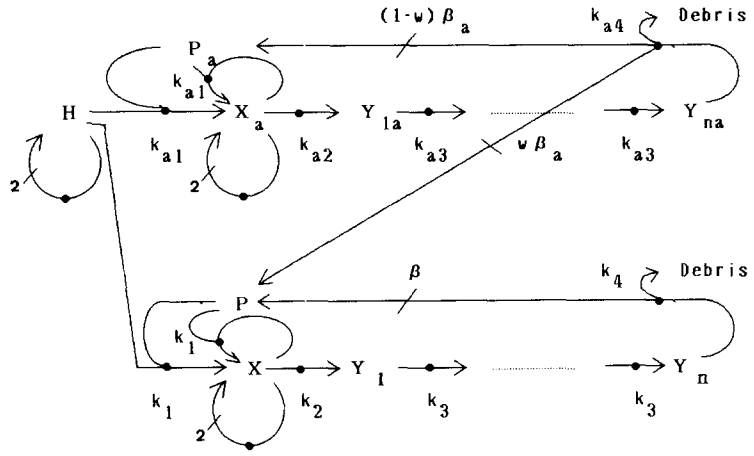

Fig. 12. A chemical reaction model for Qβ reversion process. Subscript "a" denotes amber mutant. See also Fig. 2b. w, mutation rate.

former plate formed the revertant population, while they comprised the total population on the latter. Figure 11 shows the biological relaxation profile from the disadvantageous amber mutant dominant state to advantageous revertant dominant state. Qβ is not totally virulent to *E. coli*S26Rle (37° C the diluted nutrient broth) and some hosts become carrier cells. For simplicity, the obtained data were analyzed according to the totally virulent phage model shown in Fig. 12.

The selection coefficient of the amber mutant of -0.22 was obtained by curve fitting. Mutation rates were not determined by this relaxation curve only, because the initial fraction of revertant was high enough. The relaxation curve does not mainly depend on the mutation rate except in the initial phase.

VIII. LOCAL LANDSCAPE OF SELECTION COEFFICIENT ON THE BASE SEQUENCE SPACE

1. The Base Sequence Space

The dependence of selection coefficient of Ff phages on the base sequence is called the local landscape of selection coefficient near fd phage in the base sequence space. The landscape is analogous to Wright's fitness contours in the field of gene combinations(*68*). The concept of sequence space, however, is based on the DNA base sequence and involves the concept of connectivity. Each base

sequence of length ν corresponds to a point in ν-dimensional sequence space. Each pair of points converted to each other by a single point mutation is connected by a segment in the space. If we reduce the number of bases from four(A, T, G, C) to two(R, Y), the sequence space is a ν-dimensional hypercube the edge of which represents a point mutation.

Hamming distance d_{Hij} between the i-th and j-th sequences in the ν dimensional sequence space is defined by the number of different bases in the two aligned sequences, so $Max(d_{Hij}) = \nu$. The sequence space is very compact. The number of nearest neighbors is also ν. In the sequence space, one point is surrounded by many neighbors. Hamming distances among Ff phages in 6408-dimensional sequence space are: d_H(fd, M13) = 193, d_H(M13, f1) = 60, d_H(f1, fd) = 180.

2. Local Landscape and Slightly Deleterious Mutants

A summary of selection-coefficient data obtained in both competition experiments and biological relaxation experiments is shown in Table I (*69*, *70*). Strain fdT is a mutant with only one base substitution. Strain fdU1 is an advantageous mutant which emerges during long cellstat culture of fd wild type under standard experimental conditions. We can see a distribution of the selection coefficient even among "wild type" strains under our standard experimental conditions. We can also get a hint of a quasi-species distribution around fd. Artificial chimera phages have very low selection coefficient in comparison with the wild type group.

Our data suggests that there are many slightly deleterious mutants in nature under a particular environmental condition. What is selected

TABLE I
Selection Coefficient Landscape near fd Phage

Phage	Hamming distance from fd	Selection coefficient (reference phage: fd)
i	$d_{H\,fd\,i}$	s_i
fdU1		0.035
δA		0.025 ± 0.01
fd	0	0
fdT	1	− 0.031 ± 0.01
M13	193	− 0.05 ± 0.01
f1	180	− 0.11 ± 0.01
fd107–111		− 0.13
M13mp11		− 0.36 ± 0.01
fd107		− 0.38

Environmental conditions: 37°C, Davis minimal medium. Host: *E. coli*S26, $h_0 = 2 \times 10^8$ CFU/ml, $D_c = 3.3h^{-1}$, with vigorous aeration.

in molecular evolution is such a distribution of mutants(a quasi-species), not any single molecular species(*1, 2*). The important role of slightly deleterious mutants in molecular evolution has been discussed (*71, 72*). Such a distribution of mutants is convenient to survey the base sequence space for the optimization of molecular function, or Darwinian evolution.

In contrast to a population of small effective size N_e, in a cellstat where N_e is very large(10^9–10^{12}), a neutral mutant can hardly exist conceptually, because the neutrality criterion of selection coefficient s is $s < s_{th} = 1/N_e = 10^{-9}$–$10^{-12}$. On the other hand, s is $k^*/k - 1$ where k^* and k are kinetic constants of the reproduction reaction of mutant phage and wild type phage, respectively.

Now consider single point mutation. If the major step of reproduction reaction consists of an elongation in replication, $1/k$ is nearly equal to $1/k_1 + 1/k_2 + \ldots + 1/k_\nu$ and all k_i are nearly the same, where k_i is the replication rate of a single nucleotide. Then, $k^*/k - 1$ is nearly equal to $(\Delta G^{*\neq} - \Delta G^{\neq})/\nu k_B T$, where $\Delta G^{*\neq}$, $\Delta G^{\neq}$, k_B and T are the activation free energy of replication for a single nucleotide at the mutation sites of mutant and wild type, Boltzmann constant, and absolute temperature, respectively. If there are L equal weight rate-limiting steps in reproduction reaction and mutant DNA takes advantage or disadvantage at each step, $s = s_m = (\Delta G^{*\neq} - \Delta G^{\neq})/\sqrt{L}\nu k_B T$. Almost all molecular interactions with different bases may generate a difference of at least $0.01 k_B T$ at room temperature in $\Delta G^{\neq}$. For Ff phage $\nu = 6{,}408$. When $L = 100$, $s_m > 10^{-7}$. Almost all base substitutions have a difference in kinetic constant of reproduction reaction. If by chance there exists a neutral mutant, a long generation period(10^{9-12}) is required before it is fixed through random drift. There are, however, many neutral mutants in an operational sense. Duration of the cellstat culture limits the resolution of measured s. Neutrality criterion in an operational sense is $s < s_c = 1/D_c t$, where D_c is the dilution rate and t is cultivation time. When $s < s_c$, we cannot judge which phage is advantageous in a competition experiment. In a normal experiment s_c is about 0.003, which is much larger than the above mentioned s_m.

Therefore evolution in a cellstat may not simulate evolution in nature on earth, in which neutral mutants and their fixation through random drift might have played an important role(*72*). It provides, however, some basic data for a theory of the actual evolution process. With large N_e, we can measure small s deterministically, which cannot be done in the field because of fluctuations of small N_e. For in-

stance, a model in which frequency of selection coefficients is distributed as exponential function or as gamma function(*73*) may be made more realistic by laboratory evolution experiments, although our data are still insufficient for this.

The landscape apparently become deformed when environmental conditions change. The Dykhuizen-Hartl effect, so-called by M. Kimura (*72*), suggests that some selectively neutral or nearly neutral allozymes have a latent potential for selection, which may be realized under certain environmental conditions (*74*). One of many unfinished study plans using a cellstat is the temperature dependence of the landscape.

Even in a cellstat, stochastic phenomena play a role. Emergence and washout of mutants with large d_H are not deterministic, but stochastic.

Several other methods are also available to map a local landscape: 1. Fingerprint analysis of molecular species distribution of nucleic acid as performed for a $Q\beta$ natural population (*20*). 2. Competition experiment among a whole spectrum of mutants having $d_H < 4$ as described in the next section, and 3. A comparative study of natural DNA as follows:

We studied (*75*) the base substitution frequency along the DNA chain, comparing known sequences of fd, M13, and f1(39–41). There are 193 base substitutions between fd and M13 of which only 12 have different amino acids. Most substitutions are therefore phenotypically silent on the protein level, so we can ignore their protein differences in following discussions on the statistics of base substitutions. In this case, random base substitutions along the DNA strand are expected, because an evolutional feedback through protein functions does not work. A running averaged (window = 180 bases) distribution map of substituted bases is, however, far from being random as shown in Fig. 13. The base sequence of gene 4 and the latter half of the intergenic region(IG) is most variable, that of gene 1 and 2 is fairly variable, but that of genes for coat proteins and functional regions of DNA such as the origin of replication or boundaries of genes are conservative. The thermal stability map of fd replicative form DNA has been determined and there is negative correlation (correlation coefficient = -0.47) between the base substitution map and thermal stability map(*75*). The thermostable regions of DNA are also evolutionary stable. Based on this phenomenon we assume that the structure of nucleic acids itself can be a phenotype of genetic information on the molecule and that the selective value depends on the local DNA or

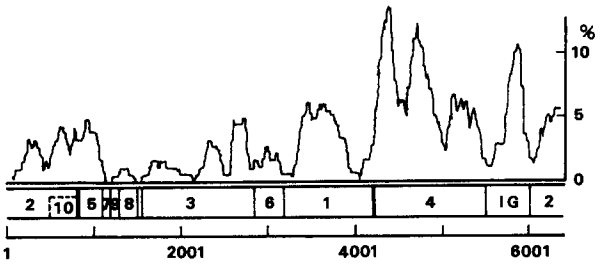

Fig. 13. Base substitution map of Ff phage (*75*). Abscissa: base-site number from the unique *Hind*II site. Top: base substitution map along DNA strand of fd, f1, and M13 phages. A site in which one of the three phages has a different base is scored as 1 point, and a site in which each of three phages has a different base as 2 points. Running averaged (window = 180 bases) 100 × (the sum of scores in each region of 180-base length on these three DNAs)/180 is plotted against the location of the region. Bottom: genomic organization of bacteriophage fd. Digits denote the gene number. IG denotes an intergenic region.

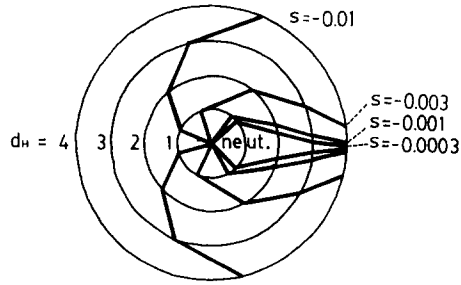

Fig. 14. Selection coefficient distribution around fd. d_H denotes Hamming distance from fd sequence. These are concentric pie charts showing the fraction of base sequences having s-value between s_i and s_{i+1}.

mRNA structure, although other explanations are also possible. We assume further that each segment of DNA contributes to the selection coefficient independently and homogenously. Then, using the s-value of fd, M13, f1, and the base substitution map, we can draw a selection coefficient contour around fd as shown in Fig. 14. This contour map is very tentative and quantitatively meaningless, but it does confirm that slightly deleterious mutants are important in molecular evolution.

3. A Model of Landscape and Cooperative Selection

The landscape obtained as above is, so to speak, a landscape for mu-

tation rate $\mu = 0$. In this case, a selective value is assigned to each individual sequence and a $W(O)$-landscape is drawn, where $W(O)$ denotes the intrinsic selective value at $\mu = 0$. For high mutation rate, however, this is not the case, because flows from surrounding sequences with $d_H = 1$ through the mutation cannot be ignored. μ and neighbor populations which reflect local landscape influence the apparent selective value of the sequence, which is determined phenomenologically by competition experiments. These flows generate a cooperativity in the selection process. Cooperative selection may not be of primary importance in present-day DNA evolution, but it may be of value in RNA evolution and in the origins of life. It may also become important in evolutionary molecular engineering using the cellstat as an evolution reactor, in which μ is very large and s is small.

Here, we demonstrate a model of cooperative selection based on Eigen's quasi-species theory (*1–6*). The subject of selection is not an individual sequence, but a normal combination of individual sequences, or quasi-species determined by the eigen vector of selective-value/mutation-rate matrix(W_{ij}) defined below (*76*). Thus, no apparent selective value can be assigned to an individual sequence; rather, it is assigned to a quasi-species and determined by the eigen value of the matrix. In a case where the mutation rate is high and the selective value landscape in the sequence space has certain characters as delineated below, the major component in a population is not necessarily the sequence having the highest intrinsic selective value.

Recent developments in the quasi-species theory showed that the mode of evolution of the Eigen model is a phase transition between the localized state and delocalized state in base sequence space (*4*). Here, we show a simple model of landscape which simultaneously provides both the above features, that is, the non-strongest champion and the phase transition-like evolution (*70, 77*).

We consider (R, Y) or (0, 1) binary sequences of length 4. The sequence space is a 4-dimensional hypercube (Fig. 15). We assume that the mutation rate μ and the decay reaction rate constant are independent of sequences and that the replication reaction rate constants A_k are functions of the Hamming distance d_{Hok} from the site 0 (binary sequence 0000). In the present model we assume the intrinsic selective value landscape has two peaks: one is sharp and high, and the other peak is not as high and broad. The population dynamics of j-th sequence under a constraint of constant total population is described by(*1*):

$$\dot{x}_j = \sum_{k=0}^{15} W_{jk} x_k - x_j \sum_{k=0}^{15} A_k x_k, \tag{23}$$

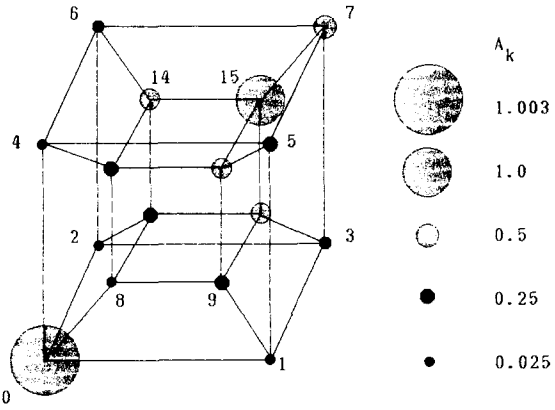

Fig. 15. A selective value landscape in 4-dimensional sequence space. (*70, 77*). A_k is the intrinsic selective value of each site at zero-mutation rate. There are two peaks: one is high and sharp, and the other not as high and broad. Number at each point is a decimal translation of a 4-digits binary sequence.

where x_j is the mole fraction of j-th sequence and W_{jk} is the intrinsic selective value (for $j = k$) and mutational transition rate (for $j \neq k$) given by

$$W_{jk} = A_k \mu^{dH_{jk}} (1-\mu)^{4-dH_{jk}}. \tag{24}$$

The result of computer simulation is shown in Fig. 16 where the steady state mole fractions are plotted against the mutation rate. We can see a sharp localization/localization transition at μ_{c2} and a diffuse localization/delocalization transition at μ_{c1}. The latter will become sharp as the length ν of the sequence increases (*3*). The μ_{c1} is calculated by a formula containing ν, μ and intrinsic selective value landscape (*4*):

$$\mu_{c1} = \frac{\ln[1 + \exp < \ln \dfrac{W_{00} - W}{W} >]}{\nu}, \tag{25}$$

where $< . . . >$ indicate an average.

When the mutation rate is low enough, that is, below the critical point μ_{c2}, the major component in the population is located at the sharp maximum peak on the landscape. When for the mutation rate μ inequalities $\mu_{c2} < \mu < \mu_{c1}$ are satisfied, the major component in the population is located at the center of the broad peak on the landscape. Its strength is due to its cooperation with surrounding sequences. The parameter which measures the strength in competition experiments

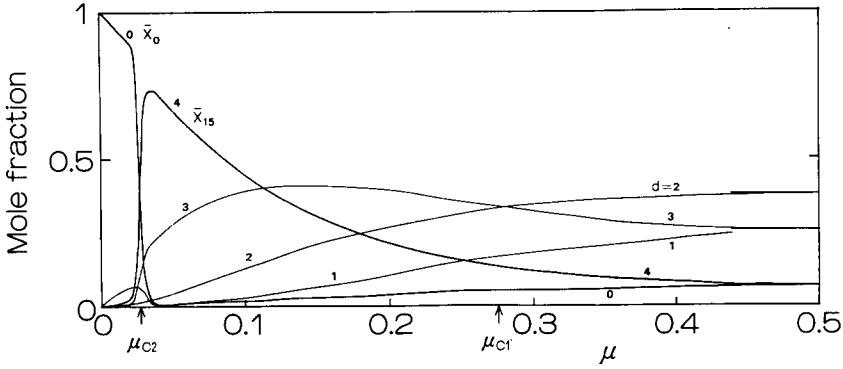

Fig. 16. Mutant-localization transitions in a sequence space (*70, 77*). The selective value landscape used is shown in Fig. 15. Abscissa: mutation rate. Ordinate: mole-fraction of each sequence (for Hamming distance 0 and 4) or each class of sequences (for Hamming distance 1, 2, and 3) at selection equilibrium. μ_{c1} indicates localization/delocalization transition. μ_{c2} indicates the transition between localization at intrinsic selective value maximum and localization at the broad peak.

is not the intrinsic selective value, but the function of local selective value landscape and mutation rate, the eigen value of (W_{ij}). The situation is analogous to the free energy concept in statistical thermodynamics. Free energy is a function of local energy landscape and temperature. Both the mutation rate and the temperature correspond to the fluctuation power. Note that in the competition experiments reported in Section VI, the mutation rate is very small and the apparent selective value is nearly equal to the intrinsic selective value.

In the simulation, we saw a critical slowing down near the transition point. The time (generation) at which the mole fraction of the defeated sequence decreases to $1/e$ of the initial value($= 0.5$) is plotted against the mutation rate; it has a peak at μ_{c2}. Monte Carlo simulation near the transition point showed very large and slow fluctuations of mole fractions. This seems to correspond to the fixation process of a neutral mutant by random drift. In this case, however, the neutrality depends on the mutation rate.

In any event, the present case is only the tip of the iceberg revealed by gained from computer simulation. The general theoretical treatment proved that the Eigen model is mathematically equivalent to the two-dimensional Ising model (*5*); the cooperativity or phase transition-like evolution can thus be understood in this general context. The question remains of how the physicochemical

properties of proteins and nucleic acids affect an actual landscape. A preliminary answer is shown in Table I and Fig. 14. Another preliminary answer was given by computer simulation using the RNA folding model(*78*).

IX. APPLICATION TO EVOLUTIONARY MOLECULAR ENGINEERING: THE NEXT STEP

The cellstat may be applied to various fields besides the original area of phage evolution. As suggested in Section VII, it can be used to measure the instability of recombinant DNAs. Instead of phage and bacteria, we will be able to culture viruses and cultivated eukaryotic cells, although the flow rate is very slow. It will be possible to study the evolutionary variability of a virus which is responsible for some infectious diseases on a laboratory time scale. It may also simulate a persistent infection state(*80*). Besides these future applications, here we propose the application of cellstat to evolutionary molecular engineering.

Present-day protein engineering may be characterized as a structural approach, because the conformation of a biopolymer is studied as an entity correlating both its function and its genetic information (sequence). Thinking based on the three dimensional structure of biopolymers is the tradition of molecular biology and its effectiveness has been proved repeatedly. What is finally needed in engineering, however, is not the structure, but the function. Moreover, the data input to a genetic engineering system is not conformational data, but linear sequence data. Therefore, it is worth asking whether the intermediate term, conformation, is really effective to connect the function to the sequence.

Biological evolution has also proceeded based on the relationship between sequence and function. If we can make the rate of molecular evolution fast enough in a laboratory, we can create a new protein having a new function without computing its structure.

Eigen (*6*, *81*) proposed a concept of evolutionary molecular engineering based on his theory on molecular evolution, the quasispecies theory. He asked how to optimize the function of a biopolymer. He also proposed an evolution machine or an evolution reactor, with which the scientific feasibility of optimization of the function of a biopolymer could be estimated. He and his colleagues are constructing two kinds of such machines, a serial transfer machine (*82*) and a parallel cloning/screening machine for replicating RNA molecules

(*83*). The former is an automated and rationalized varsion of Spiegelman's *in vitro* RNA evolution experiments. It may be classified as a natural selection type machine, in which the value criteria for the working replicon coincides with the value criteria for the experimenter. Then a simple cultivation will produce the most optimized replicon. The latter may be classified as an artificial selection type machine, in which neither set of value criteria coincide and the experimenter must select a clone himself.

Although the *in vitro* machine is a chemically clear system and will offer suggestions on the origins of life, or the origins of ribozymes, its role in molecular evolution will be limited, because it will be quite difficult to introduce a cell-free translational system.

The cellstat can be regarded as an *in vivo* evolution machine which uses a bacteriophage as a working replicon; it is a natural selection type machine. A chemostat for plasmid cultivation may also be regarded as an *in vivo* natural selection type machine. Both have an endogenous translational system. A bacteriophage is better as a working replicon than a plasmid because of its selfish behavior besides reasons described in p. 3. Phage utilizes a part of the host cell machinery without consideration of the consistency of the whole system. The cellstat selects the most rapidly growing phage, that is, the optimized protein without considering its part in the system.

We are also planning to develop an *in vivo* artificial selection type machine, a parallel cloning/screening machine for a Ff phage.

The following findings support the possibility of fast molecular evolution in the laboratory: firstly, Okada *et al.* (*84*) isolated several *Flavobacterium* strains, which grow on nylon oligomers as a sole source of carbon and nitrogen, in the waste water from a nylon factory. One strain contains a plasmid called pOAD2 which is responsible for degrading the nylon. Analysis of the DNA of this plasmid revealed that substitution of a few amino acids on one of the duplicated genes produced a new nylon hydrolyzing enzyme (*85*). Secondly, Chakrabarty *et al.* (*86, 87*) cultured soil bacteria carrying a plasmid which contains an enzyme that degrades an organochlorine compound. By increasing the concentration of the herbicide, 2,4,5-T, in the medium, they succeeded in 10 months to breed a plasmid which can degrade 2, 4, 5-T. Thirdly, a second gene for β-galactosidase in *E. coli* has been evolved in a serial transfer on culture plates (*88*). The active site of the *ebg* gene product thus obtained is homologous to the *lacZ* product (*89*). Fourthly, thermostable proteins have been produced using gene engineering and chemostat cultures of thermophilic bacteria

(*90*). Finally, the above mentioned evolution experiments on RNA performed by Spiegelman's group and Eigen's group realized very fast molecular evolution.

Factors determining the rate of molecular evolution are 1. length of the DNA(RNA), ν, 2. mutation rate, μ, 3. replication rate, k_g, 4. selective value landscape on the base sequence space, 5. population size, N, and 6. characteristic time of fluctuation or drift of μ and k_g.

According to Eigen's theory of quasi-species, μ can be set large when ν is short, as discussed in Section VIII. The screening efficiency of mutants is one of the most important factors in speeding up molecular evolution. Therefore a replicon of short ν has the potential for rapid evolution. We have therefore been using a bacteriophage as an experimental organism. In order to optimize the function of a biopolymer as quickly as possible, we consider it wise to treat a single gene instead of a whole genome. Mutation rate μ can be increased more than 10^4 times by using a single gene of 300 bp instead of the *E. coli* whole genome (4.7×10^6 bp). Even a bacteriophage genome is too large. Insertion of the gene into an Ff phage genome like a cassette is a realistic technique for this approach. The gene cassette is mutagenized *in vitro* randomly but at a controlled strength, and is ligated with the mutation-free Ff phage genome, and the chimera phage is cultivated in a cellstat. The whole spectrum of 3-point mutants of the gene can be obtained and tested with this process, because the population size N in a cellstat is about 10^{12}, and the population of i-th mutant of Hamming distance d_H is given by

$$n_{d_{H^i}} = N\mu^{d_H}(1-\mu)^{\nu-d_H}.\ 3^{-d_H}. \tag{26}$$

If we assume that $n_{d_{H^i}}$ corresponds to d_H amino acid substitutions and that the complete degeneracy of the third letter of codons, the number of neutral mutant of $n_{d_{H^i}}$ of this sense is given by

$$n_{>d_{H^i}} = n_{d_{H^i}}\{1 + \sum_{k=1}^{\nu/3} \binom{\nu/3}{k} \mu^k(1-\mu)^{-k}\}. \tag{27}$$

For $\nu = 219$, $N = 10^{12}$, $d_H = 3$, and $\mu = 0.015$, we get $n_{3i} = 5{,}000$ and $n_{>3i} = 15{,}000$. These values give us hope because, as mentioned earlier, a few amino acid substitutions can cause new protein functions in some cases.

The cellstat can select a little more optimized mutant gene, whereas

no batch culture method has such a resolution. Extending the model in Section VI to a mixed population of vast kinds of mutants having various selection coefficients, we get an upgrowth curve of an advantageous mutant of which the population was initially very small. For example, a mutant having $s = +0.03$ grows from a population of 10^5 to 10^{12} within 50 h. By repeating such a mutation(*in vitro*)-selection (in a cellstat) cycle, an optimized gene will be obtained. The speed of optimization depends on the selective value landscape between the initial sequence and the optimized sequence.

It is obvious that the general characteristics of the landscape reflect the general physicochemical properties of proteins or nucleic acids. The structural approach and the evolutionary approach to protein engineering are thus complementary to each other. And the results in evolutionary protein engineering will reveal deeper insight into the molecular evolution theory. We remember that Darwin himself took the suggestion offered by the breeding technology of cultivated plants and animals when he proposed the theory of natural selection (*91*).

In conclusion, mapping local landscapes around various concrete genes or genomes is of fundamental importance both in evolutionary molecular engineering and in the molecular evolution theory.

SUMMARY

Objectives of this work were as follows: 1. to establish a laboratory experimental system utilizable in a biophysical approach to molecular evolution; and 2. to provide real world parameters to theories of molecular evolution, especially to Eigen's theory of quasi-species.

Secretion type bacteriophage fd of *E. coli*, closely related phages and artificial chimera phages of fd, and a virulent phage Qβ of *E. coli* were cultured continuously in a specially designed fermenter called a "cellstat". A phage is cultured in a flow of host bacterial cells. Due to its high dilution rate, the mutant cell could not be selected in the cellstat. It was therefore recognized that the cellstat is suitable for study of the selection and evolution process of a bacteriophage under well-defined environmental conditions without interference from host cell mutations.

Population dynamics of bacteriophages of various types in the cellstat were studied theoretically by computer simulation and experimentally. A genetically invariable pure population of phage behaves

like an open non-linear chemical reaction system. An invariable mixed population shows a selection process, while a variable population generates an evolution process.

Kinetic constants describing the dynamics were determined by curve fitting between the theoretical and the experimental curve obtained from competition experiments and from biological relaxation experiments. One of the most important kinetic parameters thus obtained was the selection coefficient, and its dependence on the base sequence of phage DNA. We drew a local landscape of the selection coefficient near the fd sequence on the base sequence space. From this landscape we were able to confirm the importance of slightly deleterious mutants in molecular evolution. We also confirmed the possibility of developing an evolutionary molecular engineering using a cellstat as an evolution reactor and fd phage as a working replicon.

Novelties of this work were as follows: 1. the first stable continuous culture of a bacteriophage was achieved with a cellstat; 2. a local landscape of selection coefficient near the fd sequence on the sequence space was the first experimental drawing of such a map; 3. a biological relaxation method was realized to measure kinetic constants of a biological kinetic process, or molecular evolution; and 4. a practical engineering process of evolutionary molecular engineering was proposed.

Acknowledgments

I thank my colleagues, K. Nishigaki, Y. Kinoshita, H. Kihara, and many students for their continued cooperation. I thank Professors A. Wada and T. Tanaka for encouragement. I also appreciate the interest and discussions of Professors M. Eigen and W. Gardiner and their colleagues. I am grateful to Y. Saito for making a database of the literature. This work was supported by grants from the Ministry of Education, Science and Culture, Japan.

REFERENCES

1 M. Eigen, *Naturwissenschaften*, **58**, 465(1971).
2 M. Eigen and P. Schuster, *Naturwissenschaften*, **64**, 541(1977).
3 J. Swetina and P. Schuster, *Biophys. Chem.*, **16**, 329(1982).
4 J.S. McCaskill, *J. Chem. Phys.*, **80**, 5194(1984).
5 I. Leuthäusser, *J. Chem. Phys.*, **84**, 1884(1986).
6 M. Eigen, *Chem. Scripta*, **26B**, 13(1986).
7 D.R. Mills, R.L. Peterson, and S. Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.*, **58**, 217 (1967).
8 S. Spiegelman, *Q. Rev. Biophys.*, **4**, 213(1971).

9 F.R. Kramer, D.R. Mills, P.E. Cale, T. Nishihara, and S. Spiegelman, *J. Mol. Biol.*, **89**, 719(1974).
10 M. Sumper and R. Luce, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 162(1975).
11 C.K. Biebricher, M. Eigen, and R. Luce, *J. Mol. Biol.*, **148**, 391(1981).
11a C.K. Biebricher, M. Eigen, and R. Luce, *Nature*, **321**, 89(1986).
12 C.K. Biebricher, M. Eigen, and R. Luce, *J. Mol. Biol.*, **148**, 369(1981).
13 C.K. Biebricher, *Evol. Biol.*, **16**, 1(1983).
14 A. Novick and L. Szilard. *Science*, **112**, 715(1950).
15 A. Novick and L. Szilard, *Proc. Natl. Acad. Sci. U.S.A.*, **36**, 708(1950).
16 D.E. Dykhuizen and D.L. Hartl, *Microbiol. Rev.*, **47**, 150(1983).
17 J.Drake, "Evolution in the Microbial World," ed. by M.J. Carlile and J.J. Skehel Cambridge Univ. Press, London, p. 41(1974).
18 D.A. Lauffenburger, *Trends Biotechnol.* **5**, 87(1987).
19 E. Domingo, R.A. Flavell, and C. Weissmann, *Gene*, **1**, 3(1976).
20 E. Domingo, D. Sabo, T. Taniguchi, and C. Weissmann, *Cell*, **13**, 735(1978).
21 E. Domingo, M. Davilla, and J. Ortin, *Gene*, **11**, 333(1980).
22 G. Hobom and G. Braunitzer, *Hoppe-Seyler's Z. Physiol. Chem.*, **348**, 783(1967).
23 M.T. Horne, *Science*, **168**, 992(1970).
24 M. Varon, *Nature*, **277**, 386(1979).
25 Y. Husimi, K. Nishigaki, Y. Kinoshita, and T. Tanaka, *Rev. Sci. Instrum.*, **53**, 517(1982).
26 Y. Husimi and H.-C. Keweloh, *Rev. Sci. Instrum.*, **58**, 1109(1987).
27 V. Bryson, *Science*, **116**, 45(1952).
28 S.J. Pirt, "Principles of Microbe and Cell Cultivation," Blackwell Scientific Pub., Oxford(1975).
29 R.C. Valentine, P.M. Silverman, K.A. Ippen, and H. Mobach, *Adv. Microbiol. Physiol.*, **3**, 1(1969).
30 D.A. Marvin and B. Hohn, *Bacteriol. Rev.*, **33**, 172(1969).
31 D.T. Denhart, D. Dressler, and D.S. Ray(eds.), "The Single-Stranded DNA Phages," Cold Spring Harbor Laboratory, Cold Spring Harbor (1978).
32 I. Rasched and E. Oberer, *Microbiol. Rev.*, **50**, 401(1986).
33 W. Fulford, M. Russel, and P. Model, *Proc. Nucleic Acid Res. Mol. Biol.*, **33**, 141(1986).
33a N.D. Zinder and K. Horiuchi, *Microbiol. Rev.*, **49**, 101(1985).
34 P.H. Hofschneider, *Z. Naturforschng.*, **18B**, 203(1962).
35 H. Hoffmann-Berling, D.A. Marvin, and H. Duerald, *Z. Naturforschng.*, **18B**, 876 (1963).
35a D.A. Marvin and H. Hoffmann-Berling, *Z. Naturforschng.*, **18B**, 884(1963).
35b H. Hoffmann-Berling, H. Duerald, and I. Beulke, *Z. Naturforschng.*, **18B**, 893(1963).
36 T. Loeb, *Science*, **131**, 932(1960).
37 T. Nishihara and I. Watanabe, *Uirus*, **17**, 42(1967).
38 A.A. Schluederberg, B. Marshall, C. Tachibana, and S.B. Levy, *Nature*, **283**, 792(1980).
39 P.M.G.F. van Wezenbeek, T.J.M. Hulesbos, and J.G.G. Schoenmarkes, *Gene*, **11**, 129(1980).
40 E. Beck, R. Sommer, E.A. Auerswald, Ch. Kurz, B. Zink, G. Osterburg, and H. Schaller, *Nucleic Acids Res.*, **5**, 4495(1978).
41 E. Beck and B. Zink, *Gene*, **16**, 35(1981).
42 D.F. Hill and G.B. Petersen, *J. Virol.*, **44**, 32(1982).
43 H. Tzagoloff and D. Pratt, *Virology*, **24**, 372(1964).
44 J.D. Boeke, P. Model, and N. D. Zinder, *Mol. Gen. Genet.*, **186**, 185(1982).
45 C.K. Biebricher and E.-M. Dueker, *J. Gen. Microbiol.*, **130**, 941(1984).
45a C.K. Biebricher and E.-M. Dueker, *J. Gen. Microbiol.*, **130**, 951(1984).
46 I. Ikeda and Y. Husimi, unpublished data.
47 J. Messing, B. Gronenborn, B. Mueller-Hill, and P.H. Hofschneider, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 3642(1977).
48 J. Messing, *Methods Enzymol.*, **101**, 20(1983).

49 K.R. Yamamoto and B. M. Alberts, *Virology*, **40**, 734(1970).
50 K. Furuse, A. Hirashima, H. Harigai, A. Ando, K. Watanabe, K. Kurosawa, Y. Inokuchi, and I. Watanabe, *Virology*, **97**, 328(1979).
50a Y. Inokuchi, R. Takahashi, T. Hirose, S. Inayama, A.B. Jacobson, and A. Hirashima, *J. Biochem.*, **99**, 1169(1986).
51 C. Weissmann, *FEBS Lett.*, **40**, S10(1974).
52 R.C. Valentine and M. Strand, *Science*, **148**, 511(1965).
53 Y. Husimi, unpublished.
54 Y. Husimi and H. Kihara, unpublished.
55 W. Gardiner, *Ber. Bunsenges. Phys. Chem.*, **90**, 1024(1986).
55a W. Gardiner, M. Eigen, C.K. Biebricher, Y. Husimi, H.-C. Keweloh, and A. Obst, Proc. 2nd Int. Workshop on Modeling of Chemical Reactions, Heidelberg, p. 17 (1986).
56 C.K. Biebricher, M. Eigen, and W.C. Gardiner, Jr., *Biochemistry*, **22**, 2544(1983).
57 C.K. Biebricher, M. Eigen, and W.C. Gardiner, Jr., *Biochemistry*, **23**, 3186(1984).
58 C.K. Biebricher, M. Eigen, and W.C. Gardiner, Jr., *Biochemistry*, **24**, 6550(1985).
59 M. Tomoeda, M. Inuzuka, and T. Date, *Proc. Biophys. Mol. Biol.*, **30**, 23(1975).
60 Y. Husimi, T. Tanaka, K. Nishigaki, and Y. Kinoshita, *Biophysics*, **20**, 379(1980).
61 C.S. Freitag, J.M. Abraham, J.R. Clements, and B.I. Eisentein, *J. Bacteriol.*, **162**, 668(1985).
62 A. Schwienhorst and C. Biebricher, personal communication.
63 M. Makita and Y. Husimi, unpublished data.
64 M.H. Adams, "Bacteriophages," Interscience Pub., New York(1959).
65 A. Schwienhorst, personal communication.
66 R. Herrmann, K. Neugebauer, H. Zentgraf, and H. Schaller, *Mol. Gen. Genet.*, **159**, 171(1978).
67 J. Messing and J. Vieira, *Gene*, **19**, 269 (1982).
68 S. Wright, "Evolution and the Genetics of Populations," Vol. 3, Univ. of Chicago Press, Chicago, p. 452 (1977).
69 Y. Husimi, H. Kihara, J. Torisu, S. Yoshimura, K. Nishigaki, and Y. Kinoshita, unpublished.
70 Y. Husimi, K. Nishigaki, H. Kihara, and Y. Kinoshita, Proc. 19th Yamada Conference on Ordering and Organization in Ionic Solutions, World Scientific Pub., Singapore, p. 165(1988).
71 T. Ohta, *Nature*, **246**, 96(1973).
72 M. Kimura, "The Neutral Theory of Molecular Evolution," Cambridge Univ. Press, London(1983).
73 M. Kimura and T. Ohta, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 2848(1974).
74 D.L. Hartl and D.E. Dykhuizen, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 6344(1981).
75 Y. Husimi and K. Shibata, *J. Phys. Soc. Jpn.*, **53**, 3712(1984).
76 B.L. Jones, R.H. Enns, and S.S. Rangnekar, *Bull. Math. Biol.*, **38**, 15(1976).
77 Y. Husimi, *Viva Origino*, **16**, 136 (1988).
78 W. Fontana and P. Schuster, *Biophys. Chem.*, **26**, 123(1987).
79 D.A. Buonagurio, S. Nakada, J.D. Parvin, M. Krystal, P. Palese, and W.M. Fitch, *Science*, **232**, 980(1986).
80 J. Holland, K. Spindler, F. Horodyski, E. Gragau, S. Nichol, and S. van de Pol, *Science*, **215**, 1577(1982).
81 M. Eigen and W. Gardiner, *Pure Appl. Chem.*, **563**, 967(1984).
82 G. Bauer, H. Otten, and M. Eigen, personal communication.
83 H. Otten and M. Eigen, personal communication.
84 H. Okada, S. Negoro, H. Kimura, and S. Nakamura, *Nature*, **306**, 203(1983).
85 S. Negoro, H. Kimura, K. Fujiyama, Y.Z. Zhang, H. Kanzaki, and H. Okada, *J. Biol. Chem.*, **259**, 13648(1984).
86 S.T. Kellogg, D.K. Chatterjee, and A.M. Chakrabarty, *Science*, **214**, 1133(1981).

87 D. Ghosal, I.-S. You, D.K. Chatterjee, and A.M. Chakrabarty, *Science*, **228**, 135(1985).
88 J.H. Campbell, J.A. Lengyel, and J. Langridge, *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 1841 (1973).
89 A.V. Fowler and P.J. Smith, *J. Biol. Chem.*, **258**, 10204(1983).
90 H. Liao, T. Mckenzie, and R. Hageman, *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 576(1986).
91 C. Darwin, "On the Origin of Species by Means of Natural Selection," John Murray, London(1859).
92 F. Buchholtz and F.W. Schneider, *Biophys. Chem.*, **26**, 171(1987).

Received for publication April 12, 1988.