A MAP OF DISTANCES ALONG THE DNA MOLECULE OF PHAGE T4

GISELA MOSIG

Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37203

Received August 29, 1967

THE comparison of genetic map distances with distances along the DNA molecule requires measurements which do not depend on frequencies of recombination between genetic markers. Several investigators have used various methods to estimate the amount of DNA in certain segments of bacteriophage genomes (MESELSON and WEIGLE 1961; KAISER 1962; STAHL *et al.* 1964; JORDAN and MESELSON 1965; HOGNESS *et al.* 1966; GOLDBERG 1966; MOSIG 1966). This paper describes non-recombination distance measurements which depend on the probability of cutting DNA fragments between marker pairs of bacteriophage T4 during maturation. These measurements have led to the construction of a map of physical distances between genetic markers of this phage which is a complete circle.

Populations of bacteriophage T4 include small fractions of non-viable particles of lower than average buoyant density ("light" particles), each containing a single DNA fragment of fixed length which can be measured. Although these light particles are not viable in single infection, in mixed infection with normal particles they transmit their genetic information to viable progeny. Since the fragments represent continuous random segments of the circular genome, the "light" particles can be used to measure distances between genetic markers (MOSIG 1966). Previous experiments using this approach had provided preliminary measurements of distances between an *rII* reference marker and 22 amber (*am*) mutations in other genes which are located within one third of the map on either side of the *rII* reference marker (MOSIG 1966). However, some inconsistencies among these data were not yet understood at that time (MOSIG 1966, Table 2). The light particles used in some of the previous experiments had not been examined for the length of their DNA molecules. Apparently this led to the inconsistencies. More thorough investigations revealed two additional classes of nonviable light T4 particles, in addition to those described by MOSIG (1963). These particles contain DNA fragments which are 0.77 or 0.9 of the normal molecular length.

This paper reports an extension of the previous distance measurements, using only light particles whose DNA molecule had been measured, and also using more *am* mutations and *rI* and *rIII* mutations as additional reference markers. The results of these experiments can be summarized in three fractional maps, each of which represents the distances of certain *am* mutations from one of the reference markers. When the three fractional maps are combined, they yield a self-consistent circular T4 map.

MATERIALS AND METHODS

Escherichia coli strains B, CR63, K(λh) no. 3, and S/6 were obtained from A. H. DOERMANN; strain Bsu+ (011′) was isolated by S. BRENNER and obtained from F. W. STAHL. Most *am* mutants of bacteriophage T4D were isolated by R. H. EPSTEIN or R. S. EDGAR (EPSTEIN *et al.* 1963) and obtained from R. S. EDGAR. The *am*+ osmotic shock resistant mutants, the *rII* mutant r_{73}, and some of the *am* mutants were obtained from A. H. DOERMANN. The *rI* mutant r_{48} and the *rIII* mutant r_{67} were derived from phage stocks obtained from F. C. WOMACK. In addition to the *rI* or to the *rIII* mutation, these phage had contained several *am* mutations. They were crossed to wild-type phage, and *r am*+ phage were selected from the progeny of these crosses.

General procedures and media for the growth of phage stocks were described by CHASE and DOERMANN (1958). The *am* stocks were grown in *E. coli* CR63 or 011′. Most of the techniques utilized in the present work and the general procedure for the isolation of the light particles were described previously (MOSIG 1966). In some of the CsCl density gradient runs, the previous procedure was modified and eight-ml-samples of purified lysate were centrifuged for 40 hrs at 17,000 rpm at 23°C in a Spinco type SW25 rotor, and several different fractions containing light particles were recentrifuged under the same conditions in a Spinco type SW39 rotor. From aliquots of different CsCl gradient fractions, DNA was extracted with phenol according to the method of FRANKEL (1963). The relative lengths of the extracted DNA molecules were determined from their relative sedimentation through sucrose gradients, according to the equation

$$\frac{D^1}{D^2} = \left(\frac{M^1}{M^2}\right)^{0.35} \tag{1}$$

(BURGI and HERSHEY 1963), using DNA molecules from viable particles as reference.

In a typical experiment for measuring distances, *E. coli* B were infected simultaneously with an average of one complete *r*+ *am* particle and <0.03 light particles (genotype *r am*+) per bacterium. Neither the *am* nor the light particle alone can grow in this strain. However, mixedly infected bacteria which receive the respective *am*+ allele from the light particle yield viable phage progeny. Nine minutes after infection, the infected bacteria were plated on *E. coli* S/6 which permits all *am*+ and only *am*+ phage to grow. The resulting plaques were classified as pure *r*+ or mixed *r* and *r*+ (mottled) by inspection. The pure *r*+ plaques arise from mixed infection with light particles, isolated from the *r am*+ stock, which contribute to the progeny the *am*+ allele but not the *r* marker. They are taken to represent the fraction of DNA fragments in light particles which contain the *am*+ allele but have lost the *r* marker.

Theory of distance measurements: If DNA fragments of uniform length cover any segment of the circular genome at random, the fraction of those which contain the *am*+ allele but have lost the *r* marker (Y) depends on the fractional length of the DNA fragments (L), on the fractional distance between the two markers (D), and on the length of the gene containing the *am* mutation (X). When $D \leqslant (1-L)$ and $L \geqslant (D+X)$, Y is proportional to D and can be used to measure D. When $D \geqslant (1-L)$, Y becomes constant and can be used to measure L.

When $D \leqslant (1-L)$, a fragment containing the *am*+ gene can include the *r* marker only by spanning the shorter of the two distances on the circular map. However, two considerations need to be taken into account: (1) The normal T4 DNA molecule is redundant for a small fraction (R) of the T4 genome (STREISINGER, EDGAR and DENHARDT 1964; MCHATTIE *et al.* 1967). Thus, when the length L_{DNA} is measured as fraction of the normal DNA molecule, the length L expressed as fraction of the genome is

$$L = L_{\mathrm{DNA}}\,(1+R) \tag{2}$$

(2) Supposedly, an *am*+ allele on a fragment can function only when the fragment contains the complete *am*+ gene or an even larger functional unit (unless it recombines with the complete genome before expressing its function). If so, all *am* mutations in the same gene should give the same value of Y, independently of their position within the functional unit. Y is measured experimentally as the fraction of pure *r*+ plaques. L_{DNA} is measured as described above. When the complete genome is taken as unity, the average size of the functional unit

as X and the distance between the r-reference marker and the proximal end of the functional unit as D, for $D \leqslant (1-L)$,

$$Y = \frac{D}{L-X} \tag{3}$$

or

$$D = LY-XY \tag{4}$$

The size of X is different for each *am* mutation and cannot be measured accurately. Since the term XY is negligible as compared to LY, the approximation of equation (4):

$$D' = LY \tag{5}$$

was used to calculate a given distance. Thus D' is actually longer by the value of XY than the distance between the reference marker and the proximal end of the functional unit under study; i.e., D' depends on the distance between the r marker and the proximal end *as well as* on the size of the functional unit(s) to be complemented, but it is not the distance between the reference marker and the distal end of the gene, as erroneously stated in a previous paper (Mosig 1966). The resulting error is, however, negligible.

When $D \geqslant (1-L)$ and $L \geqslant 0.5$, a fragment containing the am^+ gene can include the r marker by spanning either distance on the circular map. Then the fraction of DNA fragments which contain the am^+ allele and have lost the r marker depends only on L and X, and becomes constant except for small variations due to different sizes of X.

$$Y = \frac{1-L}{L-X} . \tag{6}$$

This provides a measure for L, which is independent of the physical measurements of lengths of DNA fragments:

$$L = \frac{1}{Y+1} + \frac{XY}{Y+1} . \tag{7}$$

RESULTS

In different fractions of lighter than average density from CsCl gradients, three classes of light, nonviable particles of different density can be distinguished. Sucrose gradient analysis of DNA from CsCl gradient fractions lighter than the peak of viable particles showed, in addition to DNA molecules of normal length, discrete peaks corresponding to fragments 0.9, 0.77, or 0.67 of the normal length (Mosig to be published). Due to some spreading of particles of different classes in the CsCl gradient, some fractions of the density gradient contained several classes of DNA fragments in different relative amounts. However, many CsCl gradient fractions yielded only one class of DNA fragments in detectable amounts, and these fractions were selected for the distance measurements. Light particles containing the 0.67 length DNA fragments were used in the distance measurements from the *rII* and *rIII* reference markers, because they could be purified to a high degree and because the genome segment over which distances can be measured is larger, the smaller the length of the DNA fragment, when $L > 0.5$. The DNA fragments of this class of particles appear uniform in length. Figure 1 shows a typical sedimentation pattern of this class of DNA fragments.

If there were length variations among these fragments, one would expect particles containing the longer fragments to accumulate preferentially in the denser fractions of a CsCl gradient. Contrary to this prediction, DNA extracted

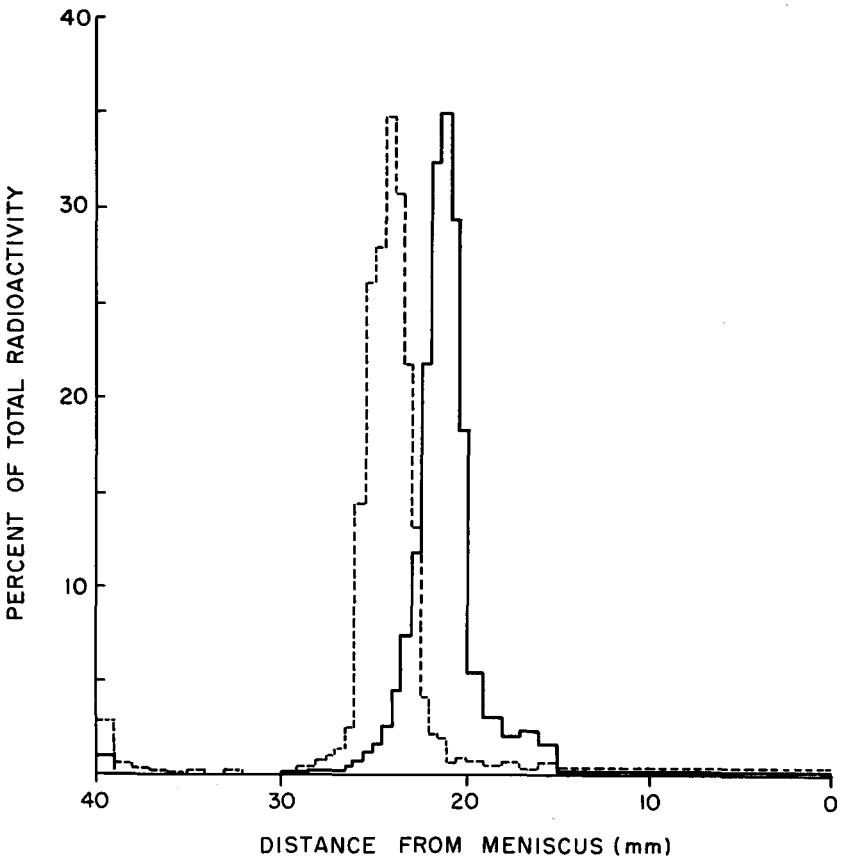

FIGURE 1.—Zone sedimentation of P^{32} labeled DNA from light particles used for the distance measurements from the *rII* marker (solid line) as compared with H^3 labeled reference DNA (broken lines).—The percent of radioactivity is normalized for fractions representing 1 mm of the total liquid column, though some fractions represent smaller or larger distances.

from "lighter" or "denser" CsCl gradient fractions (containing this class of particles) showed peaks of identical sedimentation rates.

Replicate sucrose gradients of DNA from the same CsCl gradient fraction showed as much variation as did sucrose gradients of DNA from different CsCl gradient fractions. Part, if not all, of the variability can be attributed to statistical errors in the radioactivity counts; the variability is larger in DNA samples of low radioactivity. In one set of experiments, the average value of D^1/D^2 (cf. equation 1) from 14 sucrose gradients with DNA extracted from "lighter" fractions was 0.869 ± 0.015, the average value in 7 sucrose gradients with DNA extracted from "dense" fractions was 0.870 ± 0.009.

The fractions of pure r^+ plaques resulting from *E. coli* B bacteria which were mixedly infected with light *rII* am^+ particles and various normal *am* mutants are summarized in Table 1. The results of the corresponding experiments with light *rIII* am^+ particles are summarized in Table 2. The following facts in Tables 1 and 2 should be noted:

TABLE 1

Distances between the rII *marker and different genes*

Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')	Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')
52	H 17	3162	3.2	2.2	60	HL 626	3794	4.3	2.9
38	B 262	5322	6.7	4.5	39	N 116	2917	7.7	5.2
38/37	B 262/52	868	7.2	4.8	56	E 51	5869	16.7	11.3
37	52	1279	8.0	5.4	41	81	2943	21.3	14.4
35	B 252	2352	9.6		43	B 22	1155	23.2	15.7
	E 2	2873	9.5		44	82	3344	26.6	18.0
		5225	9.6	6.5	46	130	2609	30.5	20.6
					47	A 456	4129	34.3	23.2
34	A 455	1594	12.7						
	58	620	13.2						
	B 25	1344	12.4		e	H 26	721	49.2	>32.3
		3558	12.8	8.7	1	B 24	1866	48.5	>32.3
33	N 134	3453	16.7	11.3	6	251	1226	48.5	>32.3
32	A 453	1450	16.8	11.4	10	B 255	1458	49.3	>32.3
63	E 1072	3049	26.2	17.7	14	B 20	1750	49.5	>32.3
31	N 54	5788	35.5	24.0	17	N 56	1781	49.9	>32.3
30	H 39	4726	40.6	27.5	22	B 270	1408	48.9	>32.3
54	H 21	1292	44.7	30.3	23	17	1595	49.9	>32.3
29	6	1567	45.8	31.0	24	N 65	1706	48.2	>32.3
26	N 131	3850	46.4	31.4			13511	49.3	

The numbers of plaques shown have been corrected for contamination of defective particles with viable particles and for am^+ revertants. They are the sum of plaques counted in different experiments, since the experimental error depends only on the number of plaques counted (cf. Mosig 1966). Distances were calculated using equation (5) and assuming L to be 0.677 of the complete genome. This table includes some but not all of the data which had been reported in the previous paper (Mosig 1966). Some of those data had been obtained by using light particles which had been characterized by their P^{32} content and by their bacteria-killing ability. However, the length of their DNA fragments had not been measured. Presumably some particles had contained 0.77 or 0.9 length DNA fragments rather than the assumed 0.67 length, and this led to the discrepancies mentioned in the introduction. Consequently, data from experiments in which light particles of undetermined DNA content had been used are not shown in Table 1. Some of the omitted data would yield distances similar to the ones reported here, if L were assumed to be 0.778. For example, previous experiments with *am* H21 in gene 54 gave 36.8% pure r^+ plaques. D' would be 28.6% as compared to 30.3% calculated from the data presented here.

TABLE 2

Distances between the rIII *marker and different genes*

Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')	Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')
31	N 54	5245	1.2	0.8	63	E 1071	4695	6.2	4.2
30	H 39	2896	5.3	3.6	32	A 453	1877	17.4	11.8
54	H 21	4577	11.3	7.7	33	N 134	2982	17.7	12.0
29	S- 71	4428	12.5		34	B 25	4973	21.4	
	6	3423	13.3			58	6425	22.1	
		7851	12.8	8.7		A 455	2677	22.4	
					35	B 252	4520	21.1	
26	N 131	6875	15.0	10.2		E 2	3585	20.4	
24	N 65	6507	22.9	15.5			22180	21.6	14.6
23	17	3889	23.6	16.0	37	52	6991	25.8	17.5
20	N 50	1854	25.8	17.5	38	B 262	6051	26.7	18.1
16	N 66	3052	29.4	19.9	52	H 17	4181	30.5	20.6
14	B 20	7796	30.2	20.4	60	HL 626	3543	37.0	25.1
10	B 255	3878	37.6	25.4	39	N 116	8723	40.9	27.7
8	N 132	3757	41.0	27.8					
					56	E 51	1393	47.3	>32.3
5	135	1024	42.1		41	81	1840	48.8	>32.3
	256	1121	44.1		44	82	672	48.6	>32.3
		2145	43.3	29.3	46	130	3482	50.6	>32.3
					e	H 26	2910	48.4	>32.3
6	251	3565	45.3	30.7	1	B 25	2136	48.5	>32.3
7	B 16	1561	46.2	31.3	64	E 1102	1068	50.3	>32.3
					4	N 112	1499	49.3	>32.3
							15000	49.2	

The numbers of plaques shown have been corrected for contamination of defective particles with viable particles and for am^+ revertants. They are the sum of plaques counted in different experiments, since the experimental error depends only on the numbers of plaques counted (cf. MOSIG 1966). Distances were calculated using equation (5) and assuming L to be 0.677 of the complete genome. Genes 34 and 35 show similar distances from the *rIII* marker. The reasons are discussed in the text.

(1) The fractions of pure r^+ plaques range from 1.2% to 50.6%.

(2) Different mutations within the same gene give similar results (cf. A455, 58, and B25 in gene 34, 6, and *s*-71 in gene 29).

(3) Different mutations in the adjacent genes 34 and 35 give similar results, when measured against the *rIII* marker, but different results when measured against the *rII* marker.

From the data in Tables 1 and 2, distances between the reference markers and the various genes were calculated, using equation (5). Since the length of the DNA fragments in the light particles used was 0.67 of the normal molecular length and since normal T4 DNA molecules have been estimated to be redundant for about 1% of the genome (STREISINGER *et al.* 1964; MACHATTIE *et al.* 1967), L (the length of the fragments as fraction of the genome) was assumed to be 0.677. Using equation (7), L can be estimated independently from the average fraction of pure r^+ plaques obtained with *am* mutants at $D > (1-L)$, where Y becomes constant. The average for Y in the experiments with the *rII* and the *rIII* reference markers (cf. Tables 1 and 2) is 49%. (The weighted averages are 49.3% and 49.2%, respectively.) Therefore:

$$L = 0.67 + \frac{XY}{Y+1} .$$

Thus, the two independent estimates of L are equal within the errors of the measurements involved. (It may be noted that the two estimates of L are identical if the average size of a functional unit X is assumed to be 0.02 of the total genome.)

The experiments described here measure the transmission of markers on the DNA fragments to viable progeny. This transmission is very efficient (MOSIG 1963 and 1966), but one might consider the possibility that markers located near the ends of the DNA fragments do not appear as efficiently in progeny particles. If the probability that two markers are transmitted to progeny were lower than the probability that they are contained in one fragment, the "effective" length of the fragment, which is calculated using equation (7), would be smaller than the measured length. Since both estimates of L are equal within the errors of measurement, it appears that all, or nearly all, markers, when present on a fragment do appear in viable progeny.

Lysates of the *rI* am^+ mutant contain only a very small fraction of nonviable light particles. No particles with DNA fragments 0.67 of the normal length could be isolated or detected by the physical measurements. Therefore, particles containing DNA fragments which measure 0.77 of the normal length were used. The results of these experiments are summarized in Table 3.

The light *rI* am^+ particles could not be purified to the same degree as the light particles previously used, and the DNA fragments from the CsCl fraction were slightly less uniform in length than the fragments used in the previous experiments. Figure 2 shows the sedimentation pattern of these DNA fragments. Variation in lengths should result in variations in Y for genes located in a segment between $D = (1-L_{max})$ and $D = (1-L_{min})$ from the reference marker, when L_{max} and L_{min} are the lengths of the largest and of the smallest fragments respectively. L_{max} and L_{min} could not be measured. However, the larger variability of Y at distances near $(1-L)$ from the reference marker in Table 3 correlates with the length variations of the DNA fragments observed in the sucrose gradient pattern.

TABLE 3

Distances between the rI *marker and different genes*

Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')	Gene	Mutant	Number of plaques	Percent r^+ plaques	Distance (D')
47	A 456	1958	12.1	9.4	e	H 26	3731	7.1	5.6
46	130	3074	12.9	10.3	1	B 24	2578	12.2	9.5
44	82	2152	15.6	12.1	64	E 1102	3662	15.7	12.2
43	B 22	1873	18.4	14.3	4	N 112	2710	16.0	12.4
41	81	2162	21.1	16.4	7	B 16	3669	17.0	13.2
56	E 51	7043	23.1	18.0	6	251	1171	19.6	15.2
					5	256	3159	21.0	16.4
39	N 116	891	29.6	>22.8	8	N 132	1087	23.4	18.2
38	B 262	1298	32.4	>22.8	10	B 255	1013	26.3	20.5
37	52	925	33.7	>22.8					
34	58	1520	31.5	>22.8					
34	B 25	1436	33.5	>22.8					
33	N 134	1233	30.7	>22.8					
54	H 21	682	32.7	>22.8					
26	S- 105	1786	28.4	>22.8					
24	N 65	1603	28.2	>22.8					
		11374	31.0	...					

The numbers of plaques shown have been corrected for contamination of defective particles with viable particles and for am^+ revertants. They are the sum of plaques counted in different experiments, since the experimental error depends only on the numbers of plaques counted (cf. Mosig 1966). Distances were calculated using equation (5) and assuming L to be 0.778 of the complete genome.

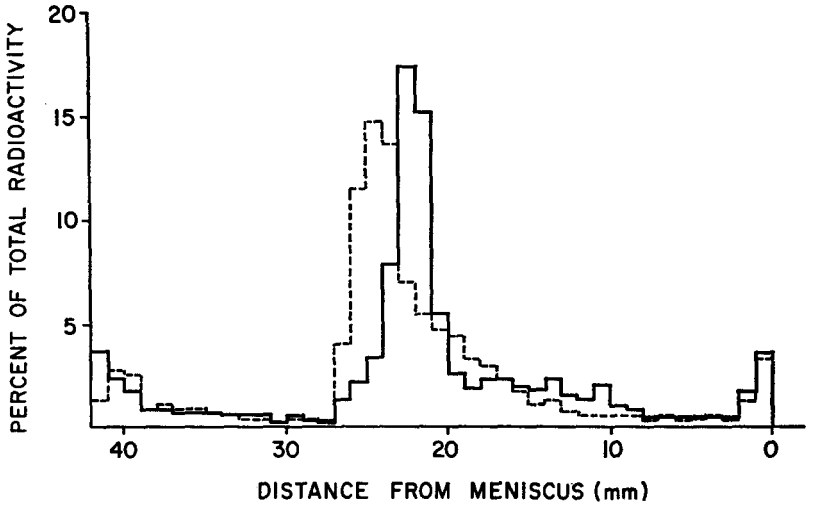

FIGURE 2.—Zone sedimentation of P^{32} labeled DNA from light particles used for the distance measurements from the *rI* marker (solid line) as compared with H^3 labeled reference DNA (broken line).—The percent of radioactivity is normalized for fractions representing 1 mm of the total liquid column. The H^3-labeled reference DNA used here was partially broken due to radiation damage. By comparison in sucrose gradients not shown here, with P^{32}-labeled DNA extracted from viable particles, it was ascertained that the peak of this reference DNA corresponds to the position of normal T4 DNA molecules. The more homogeneous a DNA preparation, the larger is the relative amount of radioactivity in the peak fraction. In contrast to the DNA preparation shown in Figure 1, only 18% of the P^{32} of the DNA preparation shown here is found in the peak fraction.

As expected, when light particles with DNA fragments of 0.77 length are used, the fractions of pure r^+ plaques in Table 3 range only from 7% to 33%. The average fraction of pure r^+ plaques obtained with *am* mutants at $D > (1-L)$ is 0.31 from which L can be calculated, using equation (7) as

$$L = 0.765 + \frac{XY}{Y+1} .$$

This value is in good agreement with the estimate of L based on sedimentation rates.

From the distances given in Tables 1–3, three fractional maps were constructed. The fractional maps containing the *rII* or the *rIII* reference markers were aligned at the position of gene 31 and the maps containing the *rIII* or the *rI* reference markers at the position of gene 10. This alignment gives the combined circular map shown in Figure 3, in which the *rI* and the *rII* maps are also aligned so that the region between genes 56 and 47 overlaps correctly. Thus, distances from any two reference markers are congruent within the errors of the measurements. The marker order on this map is the same as on the recombination map, with the exception of genes 4–8. But distances between the markers are only approximately similar on the two types of maps.

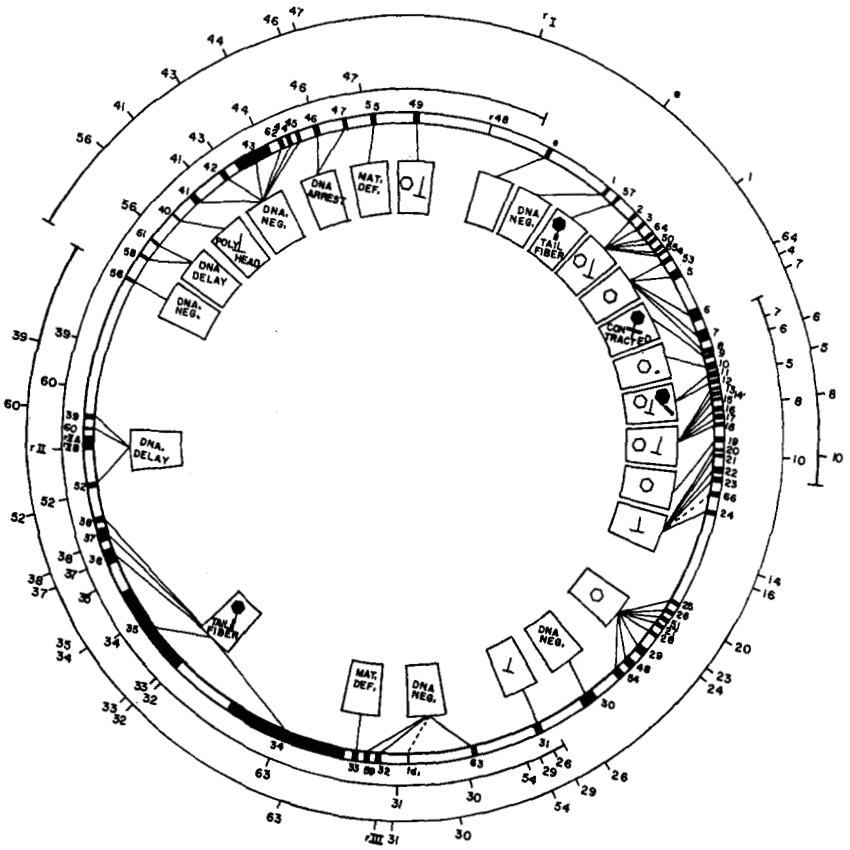

FIGURE 3.—Maps of distances between genetic markers in the T4 chromosome.

The inner circle shows functions of genes and recombination distances between them (EDGAR and WOOD 1966). The three outer arcs show distances obtained by the method described in this paper.

DISCUSSION

Vegetative T4 DNA can be larger than DNA in mature phage particles (FRANKEL 1966, 1968; ALTMAN 1967) and is thought to be cut during maturation to measure discrete lengths which constitute the amount of DNA that fits into a phage head (STAHL 1965; STREISINGER, EMRICH and STAHL 1967). The cutting device must be responsible for the different molecular lengths which are obtained, such as the unit-length of viable particles ("whole molecules") and fragments of 0.67, 0.77, or 0.9 of that length and for the fact that the gene sequences and the base sequences of the whole molecules or of the fragments are circularly permuted over the T4 genome (STREISINGER *et al.* 1964; MOSIG 1966; MACHATTIE *et al.* 1967). KELLENBERGER (1966) has described a variant of the T4 head which is smaller than the head of a viable T4 particle. The particles which contain DNA fragments 0.67 of the normal length appear in the electron microscope like the small-head-particles described by KELLENBERGER. Prelimi-

nary measurements indicate that the heads of the other two classes of light particles are intermediate in size (MOSIG, CARNIGHAN, and BAXANDALL to be published). Length variations of DNA molecules and variations in head sizes have also been found in phage P1 (IKEDA and TOMIZAWA 1965; WALKER 1966).

The data presented above show that one can construct a complete circular map of the T4 genome in which distances are measured by the frequencies with which the cutting events separate two markers of a pair. They provide distance measurements only between, not within genes, since a fragment presumably has to carry the complete corresponding wild type gene or functional unit in order to compensate for an *am* mutation. This is assumed, since two mutations within the same "late" cistron cannot complement each other when in *trans*-position (STREISINGER and FRANKLIN 1956; EPSTEIN *et al.* 1963). Support for this assumption comes from the fact that several mutations within the same gene give similar distances (cf. genes 34, 35, 29, and 5) and, in particular, from the distance measurements from genes 34 and 35 in opposite directions. They are discussed below in connection with Figure 4. The statistical error for measurements of small intervals is large. Distances within a gene or within any short interval can be measured much more accurately by the method devised by GOLDBERG (1966). The marker order on the map is uniquely defined when the distances can be checked against different reference markers. For most of the markers used, this order is colinear with the marker order on the recombination map of T4 (STAHL, EDGAR and STEINBERG 1964; EDGAR and WOOD 1966). However, the order of genes 5–7 appears inverted, when compared to their order on the recombination map of EDGAR and WOOD. This might be due to insufficient accuracy in either of the measurements involved. On the other hand, the *rI* and *rIII* phage strains from which the light particles were isolated, may indeed contain an inversion. As described above, these strains were derived from strains containing many *am* mutations. One or several of these mutations might have facilitated pairing of DNA molecules which resulted in an inversion (THOMAS 1966). This possibility is being investigated further.

The use of equation (5) to calculate distances implies that the ends of DNA fragments are cut at random positions of the genome. The following arguments lead to the conclusion that this assumption is correct and that these distances reflect true physical distances:

(1) The distances are additive in two respects: (a) The sum of all fractional distances yields a complete circle. (b) Distances measured from two reference points are congruent; for example, the difference in distance from genes 52 or 38 to the *rII* reference marker is the same as the difference in distance from these genes to the *rIII* reference marker. This is observed for almost all distances which can be measured from reference markers in opposite directions. The small deviations can be ascribed to the statistical error of the measurements and to the fact that in either direction D is the distance between the reference marker and a position closer to the proximal end than to the distal end of the corresponding gene (cf. equation 5).

(2) Some regions are represented with smaller or larger distances than the

corresponding distances on the recombination map. If this were due to a particularly low or high probability of cutting molecular ends in these regions, one should observe similar discrepancies in regions at distance L clockwise or counter-clockwise from the primary region because the DNA fragments are uniform in length and therefore lower than average cutting probabilities for left ends should be correlated with lower than average cutting probabilities for right ends at distance L and vice versa. The most pronounced discrepancy occurs in the region including genes 34 and 35, where distances are smaller than distances on the recombination map. At distance L counter-clockwise from that region, in the interval 41–47–(*rI*), both distance measurements agree remarkably well. At distance L clockwise from the first region, in the interval 7–26, recombination distances appear even smaller than physical distances. This indicates that the discrepancies are not due to differential cutting probabilities.

In most but not all sections on the map, the recombination and physical distances agree remarkably well. This argues for the overall validity of the mapping functions which convert recombination frequencies to map distances. However, recombination distances do not necessarily reflect true physical distances, and mapping functions do not correct for true noncorrespondence between them. As pointed out above, some distances disagree significantly on the two maps. In particular, the region including genes 34 and 35 shows larger recombination distances, the region between the *rII* marker and gene 56, and between the *rI* marker and gene 20 shows smaller recombination distances than physical distances. Thus it appears that in several large areas of the T4 genome local factors other than distance influence genetic recombination frequency. TESSMAN (1965) has made similar observations for very small intervals in the *rII* region. Two other findings might be the consequence of the unusually high recombination frequencies in the region including genes 34 and 35: WOMACK (1965) found that wild type alleles of *am* mutations in these genes are rescued from irradiated phage with higher than average efficiency; BERGER (1965) and ROTTLÄNDER, HERMANN and HERTEL (1967) observed exceptionally high frequencies of partial heterozygotes for markers in gene 34.

The genome of bacteriophage λ appears to contain a region of considerable physical length within which genetic recombination is rare (JORDAN and MESELSON 1965).

The nature of the factors which influence recombination frequencies is at present unknown. But one might consider that certain base sequences affect genetic recombination; and unequal nucleotide content, unequal glucosylation of hydroxymethylcytosine, or unequal methylation of certain bases, possibly combined with a differential response to recombination enzymes might be the reason for exceptionally high or low recombination frequencies in certain regions.

Distances measured to the adjacent genes 34 and 35 from different directions are not additive. It is suggested that this result is obtained because genes 34 and 35 belong to one unit of transcription (STAHL *et al.* 1966). STAHL *et al.* observed that under restrictive conditions complete phage genomes which have an *am* mutation in gene 35 give a normal burst size after mixed infection with gene 34

ts mutants, whereas complete genomes which have an *am* mutation in gene 34 give a depressed burst size after mixed infection with gene 35 *ts* mutants. They concluded that genes 34 and 35 are read together in clockwise direction.

The data presented here show that the fraction of fragments which have lost the *rII* marker is larger among those which have to compensate for a gene 34 mutation than among those which have to compensate for a gene 35 mutation (cf. Table 1). But the fraction of fragments which have lost the *rIII* marker is the same among those which have to compensate for a gene 34 or gene 35 mutation (cf. Table 2). Apparently, a fragment which has to compensate for a gene 34 *am* mutation needs to span both genes 35 and 34 and cannot function at all when it is located on fragments which have their right end within gene 34 (cf. Figure 4). (If, after mixed infection with gene 35 *am* mutants, fragments terminating in gene 34 could direct synthesis of some gene 35 product, at least one viable progeny particle should have been released.) It is concluded that gene 35 can function only on fragments which also contain at least the complete gene 34, but gene 34 can function regardless of the presence of gene 35. The simplest interpretation of this situation would be that an intact starting region for transcription and/or translation (promoter region) is necessary for functional expression of phage genes as well as of bacterial genes (JACOB, ULLMAN and MONOD 1964; BECKWITH 1967). Apparently, genetic recombination does not efficiently connect an inactive gene 35 to the necessary starting region at the beginning of gene 34, before its action is required for phage growth. This agrees with the observation by STREISINGER and FRANKLIN (1956) who found that host range (*h*) mutations in *trans*configuration are not efficient for the production of the *h* phenotype.

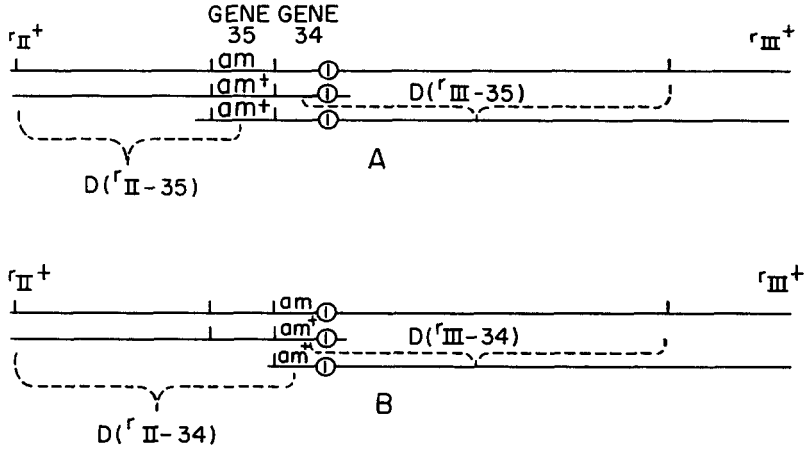

FIGURE 4.—The arrangement of genes 35 (A) or 34 (B) on fragments, which appears necessary for these genes to express their function.—① indicates the position of the necessary starting region. The two lower solid lines indicate the right hand and left hand extensions of fragments which are necessary for expression of genes 35 or 34. The broken lines reflect the apparent distances which result from these restrictions.

Thanks are due to DR. A. D. HERSHEY for his encouraging interest, DRS. R. S. EDGAR, A. H. DOERMANN, and F. C. WOMACK for mutant phage strains and information about them, DRS. P. MARGOLIN and D. PARMA for helpful discussions, DRS. T. BIBRING, A. D. HERSHEY, R. M. LITMAN, F. W. STAHL, and J. WEIL for their criticisms of the manuscript, and MRS. MARIANNE HENDRICKSON for excellent technical assistance. This work was supported by a grant from the National Institutes of Health (GM-13 221).

SUMMARY

Gene sequences of DNA fragments in light nonviable particles of bacteriophage T4 are circularly permuted over the T4 genome. The frequency with which the DNA fragments terminate between marker pairs was used to measure distances along the DNA molecules. These measurements do not depend on recombination frequencies. Distances so determined are additive and yield a circular map of the T4 genome. A comparison of this map with the recombination map of T4 shows that the gene order on both maps is the same with the possible exception of genes 5 through 7. For most intervals map distances based on recombination frequencies are similar but not identical to physical distances. However, the map interval including genes 34 and 35 shows an abnormally high recombination frequency per physical length of DNA.—The results also suggest that genes on fragments can express their function only when they are in *cis*-position to a necessary starting region for transcription and/or translation. When two genes belong to one unit of transcription or translation from which the distance to markers in opposite directions can be measured, the direction of reading can be inferred.

LITERATURE CITED

ALTMAN, S., 1967 Bacteriophage T4 DNA replication in the absence and presence of 9-aminoacridine. Ph.D. Thesis, University of Colorado.

BECKWITH, J. R., 1967 Regulation of the *lac* operon. Science **156**: 597–604.

BERGER, H., 1965 Genetic analysis of T4D phage heterozygotes produced in the presence of 5-fluorodeoxyuridine. Genetics **52**: 729–746.

BURGI, E., and A. D. HERSHEY, 1963 Sedimentation rate as a measure of molecular weight of DNA. Biophys. J. **3**: 309–321.

CHASE, M., and A. H. DOERMANN, 1958 High negative interference over short segments of the genetic structure of bacteriophage T4. Genetics **43**: 332–353.

EDGAR, R. S., and W. B. WOOD, 1966 Morphogenesis of bacteriophage T4 in extracts of mutant-infected cells. Proc. Natl. Acad. Sci. U.S. **47**: 857–868.

EPSTEIN, R. H., A. BOLLE, C. M. STEINBERG, E. KELLENBERGER, E. BOY DE LA TOUR, R. CHEVALLEY, R. S. EDGAR, M. SUSSMAN, G. H. DENHARDT, and A. LIELAUSIS, 1963 Physiological studies of conditional lethal mutants of bacteriophage T4D. Cold Spring Harbor Symp. Quant. Biol. **28**: 375–394.

FRANKEL, F. R., 1963 An unusual DNA extracted from bacteria infected with phage T2. Proc. Natl. Acad. Sci. U.S. **49**: 366–372. —— 1966 Studies on the nature of replicating DNA in T4-infected *Escherichia coli*. J. Mol. Biol. **18**: 127–143. —— 1968 Evidence for long DNA strains in the replicating pool after T4 infection. Proc. Natl. Acad. Sci. U.S. **59**: 131–138.

GOLDBERG, E. B., 1966 The amount of DNA between genetic markers in phage T4. Proc. Natl. Acad. Sci. U.S. **56**: 1457–1463.

HOGNESS, D. S., W. DOERFLER, J. B. EGAR, and L. W. BLACK, 1966 The position and orientation of genes in λ and λdg DNA. Cold Spring Harbor Symp. Quant. Biol. **31**: 129–138.

IKEDA, H., and J. TOMIZAWA, 1965 Transducing fragments in generalized transduction of phage P1. III. Studies with small phage particles. J. Mol. Biol. **14**: 120–129.

JACOB, F., A. ULLMAN, and J. MONOD, 1964 Le promoteur, élément génétique nécessaire à l'expression d'un opéron. Compt. Rend. **258**: 3125–3128.

JORDAN, E., and M. MESELSON, 1965 Discrepancy between genetic and physical lengths on the chromosome of bacteriophage lambda. Genetics **51**: 77–86.

KAISER, A. D., 1962 The production of phage chromosome fragments and their capacity for genetic transfer. J. Mol. Biol. **4**: 275–287.

KELLENBERGER, E., 1966 Control mechanisms in bacteriophage morphopoiesis. pp. 192–228. *Ciba Foundation Symposium on Principles of Biomolecular Organization.* Edited by G. E. W. WOLSTENHOLME and M. O'CONNOR. Little, Brown, Boston, Mass.

MACHATTIE, L. A., D. A. RITCHIE, C. A. THOMAS, JR., and C. C. RICHARDSON, 1967 Terminal repetition in permuted T2 bacteriophage DNA molecules. J. Mol. Biol. **23**: 355–363.

MESELSON, M., and J. WEIGLE, 1961 Chromosome breakage accompanying genetic recombination in bacteriophage. Proc. Natl. Acad. Sci. U.S. **47**: 857–868.

MOSIG, G., 1963 Genetic recombination in bacteriophage T4 during replication of DNA fragments. Cold Spring Harbor Symp. Quant. Biol. **28**: 35–42. —— 1966 Distances separating genetic markers in T4 DNA. Proc. Natl. Acad. Sci. U.S. **56**: 1177–1183.

ROTTLÄNDER, E., K. O. HERMANN, and R. HERTEL, 1967 Increased heterozygote frequency in certain regions of the T4 chromosome. Molec. Gen. Genetics **99**: 34–39.

STAHL, F. W., 1965 Recombination in bacteriophage T4. Heterozygosity and circularity. Symp. Hung. **6**: 131–141.

STAHL, F. W., R. S. EDGAR, and J. STEINBERG, 1964 The linkage map of bacteriophage T4. Genetics **50**: 539–552.

STAHL, F. W., N. E. MURRAY, A. NAKATA, and J. M. CRASEMAN, 1966 Intergenic cis-trans position effects in bacteriophage T4. Genetics **54**: 223–232.

STREISINGER, G., R. S. EDGAR, and G. H. DENHARDT, 1964 Chromosome structure in phage T4. I. Circularity of the linkage map. Proc. Natl. Acad. Sci. U.S. **51**: 775–779.

STREISINGER G., J. EMRICH, and M. M. STAHL, 1967 Chromosome structure in phage T4. III. Terminal redundancy and length determination. Proc. Natl. Acad. Sci. U.S. **57**: 292–295.

STREISINGER, G., and N. C. FRANKLIN, 1956 Mutation and recombination at the host range genetic region of phage T2. Cold Spring Harbor Symp. Quant. Biol. **21**: 103–111.

TESSMAN, I., 1965 Genetic ultrafine structure in the T4 *rII* region. Genetics **51**: 63–75.

THOMAS, C. A., 1966 Recombination of DNA molecules. pp. 315–342. Progress in Nucleic Acid Research. Edited by J. N. DAVIDSON and W. E. COHN. Academic Press, New York.

WALKER, D. H., JR., 1966 Structure and function of bacteriophage P1 and its morphological variants. Ph.D. Thesis, University of Pennsylvania.

WOMACK, F. C., 1965 Cross-reactivation differences in bacteriophage T4D. Virology **26**: 758–760.