MARKER-DEPENDENT RECOMBINATION IN T4 BACTERIOPHAGE. II. THE EVALUATION OF MISMATCH REPAIRABILITIES IN CROSSES WITHIN INDICATOR DISTANCES

V. P. SHCHERBAKOV, L. A. PLUGINA, E. A. KUDRYASHOVA, O. I. EFREMOVA AND S. T. SIZOVA

Institute of Chemical Physics, Academy of Sciences of the U.S.S.R., Chernogolovka 142432, Moscow Region, U.S.S.R.

OLEG G. TOOMPUU

Institute of Experimental Biology, Academy of Sciences of the Estonian S.S.R., Harku 203051, Estonia, U.S.S.R.

Manuscript received August 20, 1981
Revised copy accepted June 2, 1982

ABSTRACT

The contribution of mismatch repair to genetic recombination in T4 phage has been evaluated by three independent approaches: (1) testing for non-additivity of recombinant frequencies; (2) measurements of double exchange frequencies in three-factor crosses: (3) comparisons of recombination abilities of mutations occupying the same site. Quantitative agreement among the results of these approaches suggests that within distances much less than the mean length of hybrid regions, mismatch repair accounts perfectly for high negative interference as measured in three-factor crosses and as manifested by non-additivity in two-factor crosses. The mismatch repair mechanism readily recognizes only particular mismatches, the repair frequency being dependent on the base sequence in both strands of the mismatched region.

IN the companion paper (SHCHERBAKOV *et al.* 1982) a chromosome segment much less than the mean length of the hybrid region was shown to be involved in a single repair event in T4 phage. Here use is made of this observation for the development of methods aimed at evaluating the contribution of mismatch repair to genetic recombination.

RESULTS

Mismatch repair in two-factor crosses: Basic for our methodology is the concept of indicator distances. We define an indicator distance as a genetic interval of small length compared to the mean length of hybrid regions, but exceeding the length of the DNA segment involved in a single repair event. The mismatches separated by an indicator distance are spaced closely enough to fall preferentially into the same hybrid region, but are sufficiently far apart to rule out their joint repair.

Within indicator distances, the mismatch repair contribution to recombina-

tion reaches its maximum: when the distance between two markers is shorter than the repair region, their joint repair (not resulting in recombinant formation) becomes possible; when the distance increases to a length comparable to that of a hybrid region, the mismatch repair contribution also diminishes because of the reduced probability that both markers fall simultaneously into the same hybrid region.

Consider the marker sequence *i-a-j* at indicator distances. For the cross $iA \times Ia$, the recombinants *IA* are expected to arise *via* both nonrepair and repair pathways. We denote by ρ_{IA} the partial frequency of *IA* recombinants arising from the nonrepair pathway. To designate the partial frequencies of the same recombinants arising instead from mismatch repair ($a \rightarrow A$ or $i \rightarrow I$), we adopt the symbols $\kappa_{a\rightarrow A}$ or $\kappa_{i\rightarrow I}$, respectively.

It can be inferred from the basic theory (TOOMPUU and SHCHERBAKOV 1980) that, regardless of the coupling of two recombination pathways, their contributions within the limits of indicator distances approximately obey the additivity relation

$$R_{IA} = \rho_{IA} + \kappa_{i\rightarrow I} + \kappa_{a\rightarrow A},$$

where R_{IA} signifies the total frequency of *IA* recombinants.

For the crosses $aJ \times Aj$ and $iJ \times Ij$, similar approximations are legitimate:

$$R_{AJ} = \rho_{AJ} + \kappa_{a\rightarrow A} + \kappa_{j\rightarrow J},$$

$$R_{IJ} = \rho_{IJ} + \kappa_{i\rightarrow I} + \kappa_{j\rightarrow J}.$$

Since at short distances

$$\rho_{IJ} = \rho_{IA} + \rho_{AJ},$$

a simple formula

$$\kappa_{a\rightarrow A} = \tfrac{1}{2}(R_{IA} + R_{AJ} - R_{IJ}), \tag{1}$$

which allows us to calculate $\kappa_{a\rightarrow A}$, follows.

Being readily accessible to experimental determination, the partial frequency $\kappa_{a\rightarrow A}$ seems to be of value in mismatch repair investigation. Although remaining within the framework of the basic theory except for the adoption of marker-specific mismatch repair, the quantity mentioned can be shown to be characteristic of the $a \rightarrow A$ conversion and independent of the *i* and *j* markers used, provided the latter are poorly repairable. However, $\kappa_{a\rightarrow A}$ should not be taken for an elementary characteristic, since it actually refers to two reciprocal mismatches having the *a* allele in the sense and in the antisense strand, respectively.

Hereafter, $\kappa_{a\rightarrow A}$ will be referred to as the repairability of a/A to A/A or as the repairability of allele *a* when mismatched against allele *A*. Essentially, as a measure of repairability we introduce a partial quantity contributed by mismatch repair to the yield of pertinent recombinants within an indicator distance. By definition, this is the highest contribution possible.

In Figure 1, abstracted from the companion paper, the arrows show the approximate chromosome segments involved in repair of the alleles *FC1*, *op360*,

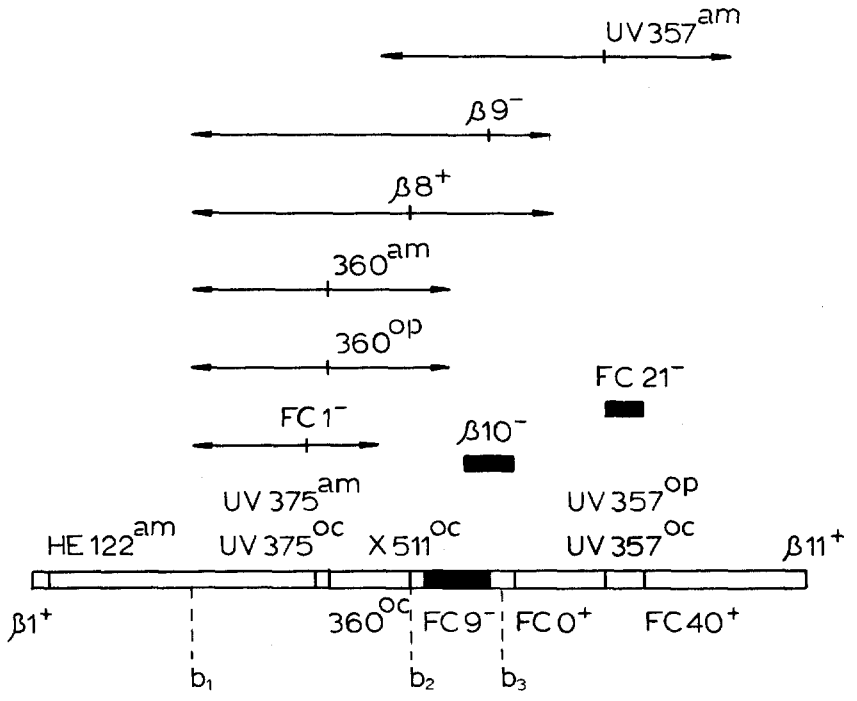

FIGURE 1.—Map of *rIIB* mutations of T4 bacteriophage. The distances were determined on the ρ scale: $\rho_{IJ} = R_{IJ} - \kappa_{i\to I} - \kappa_{j\to J}$. The ρ values were calculated for the reference mutations *HE122*, *UV375*, *β10*, *opUV357* and *β11*, which are sufficiently widely apart to be repaired independently at low κ values (for *HE122* and *β11*, κ values were assumed to be zero). The remaining mutations were mapped according to recombination frequencies in crosses with the nearest reference mutations and with each other. Phase-shift mutations are marked with the superscripts + or − to show the direction of the phase shift with reference to FC0 as + (BARNETT *et al.* 1967); amber, ocher, and opal mutations are labeled *am*, *oc* and *op*, respectively. The barriers b_1, b_2 and b_3 correspond to the terminating codons UAA, UGA, and UAA engendered by the plus-sign phase shift (BARNETT *et al.* 1967; KATZ and BRENNER 1975). Arrows show the approximate sizes of the chromosome segments involved in a single mismatch repair event.

am360, *β8*, *β9* and *amUV357* to the wild-type ones. Evidently, when the repairabilities of the respective markers are tested, the indicator distances should exceed the intervals covered by the arrows whereas the appropriate upper limits can be dispensed with, because the total genetic length of the *rIIB* segment involved is as small as 20% of a hybrid region.

In keeping with formula (1), we combined a large number of triads *i-a-j* within the limits of indicator distances. To avoid marker interference (SHCHERBAKOV *et al.* 1982), we used only mutations of low recombination ability as the test-markers *i* and *j*. The characterized mutations *a* were either of high recombination ability (HR-mutations) with known repair regions (Figure 1) or of low recombination ability (LR-mutations). We assumed the repair regions of the LR-mutations not to exceed those for their HR-type homoalleles.

The $\kappa_{a\to A}$ values were calculated for all $iA \times Ia$, $aJ \times Aj$, and $iJ \times Ij$ crosses described in the companion paper (SHCHERBAKOV *et al.* 1982) and averaged over

the triads involving the same mutation *a*. The results are summarized in Table 1 and in Figure 2, diagram I, presenting the repairability values on a linear scale.

Mismatch repair in three-factor crosses: It can be inferred from the basic theory that, within indicator distances, the frequency of *IAJ* recombinants from the cross *iAj* × *IaJ* adheres to the following approximation:

$$R_{IAJ} = \kappa_{a \to A}. \tag{2}$$

The formula suggests that in the vast majority of cases the *IAJ* recombinants originate from the repair conversion $a \to A$, the chance that the $i \to I$ and $j \to J$ conversions coincide being negligible.

In Table 2 the *IAJ* frequency values are given within indicator distances determined as described in the companion paper (SHCHERBAKOV *et al.* 1982). Ten different mutants *a* were crossed against the same *HE122β11* double mutant ($i = HE122$, $j = \beta 11$). In compliance with formula (2), the resultant wild-type recombinant frequencies should be considered as the repairabilities of the respective mutations to the wild-type alleles. As shown in Figure 2, this proves to be the case: the repairability diagram II as defined by formula (2) shows a good fit to the independent diagram I constructed with the aid of formula (1).

Comparative repairabilities: We have already concerned ourselves with comparing the recombination abilities of the mutations *a'* and *a"* occupying a common site *a* when crossed against a set of test-markers *i* (SHCHERBAKOV *et al.* 1982). With increase of *a-i* distance, the absolute value of the difference $R_{A'I} - R_{A''I}$ as related to $a'I \times A'i$ and $a''I \times A''i$ crosses rose sharply at first and reached a final value within indicator distances.

Since for any *i* the interval *a'-a"* is negligible as compared to the interval *a-i*, the partial frequencies of the wild-type recombinants arising from the nonrepair pathway of both crosses are approximately equal to one another:

$$\rho_{A'I} = \rho_{A''I}.$$

Hence, making use of the additivity relations

$$R_{A'I} = \rho_{A'I} + \kappa_{a' \to A'} + \kappa_{i \to I}$$

and

$$R_{A''I} = \rho_{A''I} + \kappa_{a'' \to A''} + \kappa_{i \to I},$$

we can derive the following relation:

$$\kappa_{a' \to A'} - \kappa_{a'' \to A''} = R_{A'I} - R_{A''I}. \tag{3}$$

Although the latter formula does not allow us to find the absolute value of repairability, it can be used to evaluate the repairability difference for mutations occupying the same site. In practice, the formula is of value when the difference $\kappa_{a' \to A'} - \kappa_{a'' \to A''}$ sought is an average over the results of several crosses involving various *i* markers. To rule out marker interference, all the *i* markers used should be of the LR-type.

The homoallelic *a'/a"* marker pairs and pertinent *i* markers are listed in Table 3. Owing to extremely dense mutation clustering within particular gene

TABLE 1

Mismatch repairabilities as determined by formula (1)

Characterized mutation *a*	Test-markers *i*	Test-markers *j*	No. of *i-a-j* triads studied	Repairability, $\kappa_{a\to A} \times 10^4$ (mean over the triads)
FC1	*β1*	*X511, FC9, β10, opUV357*	8	15.5
	HE122	*X511, FC9, β10, opUV357*		
UV375	*HE122*	*β10, opUV357, β11*	3	1.4
amUV375	*β1*	*β10, opUV357, FC21*	8	1.7
	HE122	*β10, FC0, opUV357, FC21, FC40*		
360	*β1*	*β10, opUV357, β11*	7	1.4
	HE122	*β10, UV357, opUV357, β11*		
op360	*β1*	*β10, opUV357, FC21, β11*	11	8.0
	HE122	*β10, FC0, UV357, opUV357, FC21, FC40, β11*		
am360	*β1*	*β10, opUV357, FC21, β11*	11	4.8
	HE122	*β10, FC0, UV357, opUV357, FC21, FC40, β11*		
am360[a]	*HE122*	*opUV357*	1	5.2
X511	*β1*	*opUV357, β11*	4	0.9
	HE122	*opUV357, β11*		
β8	*β1*	*opUV357, β11*	5	21.2
	HE122	*opUV357, FC40, β11*		
FC9	*β1*	*opUV357*	2	−1.1
	HE122	*opUV357*		
β9	*β1*	*opUV357, FC21*	5	15.5
	HE122	*UV357, opUV357, FC21*		
β10	*β1*	*opUV357, FC21*	5	−0.1
	HE122	*UV357, opUV357, FC21*		
FC0	*HE122*	*opUV357, β11*	2	1.9
UV357	*HE122*	*β11*	2	−0.2
	360	*β11*		
opUV357	*β1*	*β11*	5	1.8
	HE122	*β11*		
	UV375	*β11*		
	UV375[b]	*β11*		
	360	*β11*		
amUV357	*β1*	*β11*	3	28.8
	HE122	*β11*		
	360	*β11*		
FC40	*HE122*	*β11*	1	−1.0

The alleles *I*, *A* and *J* are wild type in all $iA \times Ia$, $aJ \times Aj$, and $iJ \times Ij$ crosses except the cases denoted by [a] or [b].

[a] *A* is *op360*.

[b] *I* is *amUV375*.

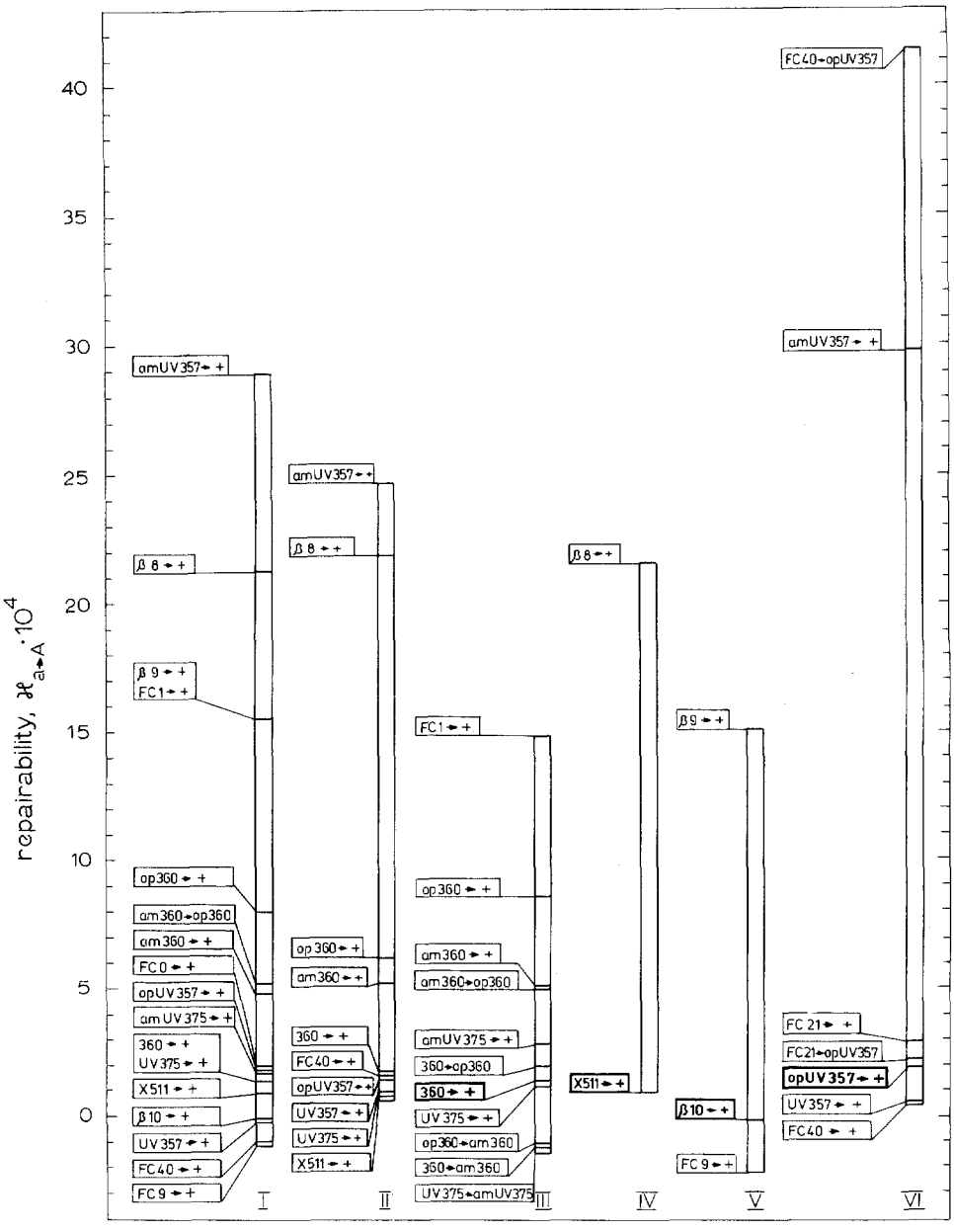

FIGURE 2.—Repairability diagrams for *rIIB* mutations of T4 phage. The following symbols are used: *amUV357* → + signifies the repair conversion of *amUV357* to the wild-type allele; *360* → *op360* denotes the repair conversion of *360* to *op360*, etc. Diagrams I and II are constructed by formulas (1) and (2), respectively; diagrams III, IV, V and VI are constructed according to formula (3) and connected with the absolute scale through the repair conversions *360* → +, *X511* → +, *β10* → +, and *opUV357* → + (bold lettering).

TABLE 2

Frequencies of IAJ *recombinants from* iAj × IaJ *crosses*

Allele *a*	$R_{IAJ} \times 10^{4a}$
UV375	0.937 ± 0.034
360	1.48 ± 0.09
op360	6.12 ± 0.19
am360	5.23 ± 0.54
X511	0.846 ± 0.088
β8	21.8 ± 3.2
UV357	0.993 ± 0.044
opUV357	1.31 ± 0.03
amUV357	24.7 ± 0.4
FC40	1.36 ± 0.11

In all crosses the alleles *I*, *A* and *J* are wild type; the alleles *i* and *j* are *HE122* and *β11*, respectively.
[a] The mean with respect to three determinations and the standard deviation of the mean are given.

segments, repairability comparisons in several pairwise combinations are legitimate for most of the mutations. The comparison of conversions of the same mutation to different alleles is possible as well. For example, the repair conversion *360* → + can be compared with the homoallelic conversions *UV375* → +, *UV375* → *amUV375*, *amUV375* → +, *op360* → +, *op360* → *am360*, *am360* → +, and *am360* → *op360*, as well as with the conversions of the same mutation *360* → *op360* and *360* → *am360*. Via common comparison partners, all the repair conversions studied can be integrated into four comparison groups.

Referring to formula (3), we calculated the repairability differences for all legitimate pairs of mutations and averaged the results over *i* markers used (Table 3, column 5). The average differences were integrated into the comparison groups mentioned above and displayed as linear diagrams III, IV, V and VI in Figure 2. Being tied in with the absolute scale via the conversions *360* → +, *X511* → +, *β10* → +, and *opUV357* → +, respectively, the diagrams give an attractive fit to the preceding diagrams I and II based on independent considerations.

Mismatch repair probabilities: An approximate linear equation

$$\kappa_{a\to A} = \tfrac{1}{2} R(\xi)\mu_{a\to A} \tag{4}$$

offers a way to transform $\kappa_{a\to A}$ values to their respective $\mu_{a\to A}$ values. We define the latter parameter as the conditional probability for the repair conversion $a \to A$ to occur once the a/A mismatch has been formed. With reference to the previous parameter μ (TOOMPUU and SHCHERBAKOV 1980), it can be represented as follows:

$$\mu = \mu_{a\to A} + \mu_{A\to a}.$$

Formula (4) implicates the coefficient $R(\xi)$, which can be regarded as the probability that, in the line of descent of a progeny phage particle picked out at random, the a/A mismatch has occurred. If we take this probability as 0.030

TABLE 3

Mismatch repairability differences as determined by formula (3)

Compared homoalleles		Test-markers	No. of *i* markers used	Repairability difference, $(\kappa_{a'\to A'} - \kappa_{a''\to A''}) \times 10^4$ (mean ± S.D.)
a'	*a''*	*i*		
FC1	*UV375*	*HE122, X511, FC9, β10, opUV357*	5	13.8 ± 1.6
UV375	*360*	*HE122, β10, opUV357, β11*	4	−0.3 ± 0.3
amUV375	*360*	*β1, HE122, β10, opUV357*	4	1.4 ± 0.8
UV375[a]	*360*	*β10, opUV357, β11*	3	−2.7 ± 0.6
op360	*360*	*β1, HE122, β10, UV357, opUV357, β11*	6	7.1 ± 0.4
am360	*360*	*β1, HE122, β10, UV357, opUV357, β11*	6	3.8 ± 0.4
360[b]	*360*	*HE122*	1	0.5
360[c]	*360*	*opUV357*	1	−2.5
op360[c]	*360*	*opUV357*	1	−2.4
am360[b]	*360*	*HE122, opUV357*	2	3.7
β8	*X511*	*β1, HE122, opUV357, β11*	4	20.4 ± 1.1
FC9	*β10*	*β1, HE122, opUV357*	3	−2.2 ± 2.9
β9	*β10*	*β1, HE122, UV357, opUV357, FC21*	5	15.2 ± 0.8
UV357	*opUV357*	*HE122, 360, β11*	3	−1.1 ± 0.4
amUV357	*opUV357*	*β1, HE122, 360, β11*	4	27.5 ± 1.1
FC21	*opUV357*	*β1, HE122, amUV375*	3	1.2 ± 2.2
FC21[d]	*opUV357*	*HE122, amUV375*	2	0.4
FC40	FC21	*HE122, amUV375*	2	−2.4
FC40[d]	*FC21*	*HE122, amUV375*	2	38.3

The alleles *A'*, *A''*, and *I* are wild type in all $a'I \times A'i$ and $a''I \times A''i$ crosses except the cases denoted by [a], [b], [c] or [d].

[a] *A'* is *amUV375*.

[b] *A'* is *op360*.

[c] *A'* is *am360*.

[d] *A'* is *opUV357*.

(TOOMPUU and SHCHERBAKOV 1980), $\mu_{a\to A}$ values ranging from less than 0.02 to 0.25 can be calculated. According to these calculations, in standard crosses involving the mutation *amUV357*, as many as 20% of mismatched *amUV357* alleles are repaired to the wild type; still higher (about 25%) is the repaired fraction in the case of *FC40* → *opUV357*. The mismatched alleles *β8*, *FC1*, *β9*, *op360* and *am360* are repaired to the wild type in 15, 10, 10, 5 and 3 cases out of 100, respectively. In the case of LR-mutations, mismatch repair to the wild type is below 2%.

The latter finding is of particular importance as far as the legitimacy of our test-marker selection is concerned. A reflection too cumbersome to be presented here shows that, in general, the complex variable $\mu_{a\to A}(1 - \mu_{I\to i})$ rather than the simple one $\mu_{a\to A}$ should be used as the argument of function (4). The former variable reduces to the latter when $\mu_{I\to i}$ is sufficiently small compared to unity, and this should be expected in the case of *i* markers of the LR-type.

This consideration, on merely statistical grounds, can also be shown to account partly for marker interference, the phenomenon which first compelled our attention to the unfitness of HR-mutations as test markers *i* (SHCHERBAKOV *et al.* 1982).

DISCUSSION

In this paper the involvement of mismatch repair was postulated in three recombination phenomena in T4 phage: positive values of $R_{IA} + R_{AJ} - R_{IJ}$ were used to calculate allele repairabilities by formula (1); in three-factor crosses, high negative interference resulting in experimental R_{IAJ} values of the order of 10^{-4} to 10^{-3} (rather than 10^{-6} to 10^{-5} as expected from coincidences of independent single exchanges) allowed us to estimate repairability according to formula (2); and finally, by formula (3), repairability values were inferred from the observation that, in two-factor crosses on the same gene fraction, recombinant frequency depends on the markers involved.

The results as juxtaposed in Figure 2 are in fair agreement, suggesting that for distances much less than the mean length of hybrid regions, mismatch repair perfectly accounts for non-additivity of recombination frequencies, high negative interference in three-factor crosses, and marker-specificity of recombination. Quantitatively, there seems to be no need for additional speculations on the origins of these phenomena. This conclusion is strictly valid in the framework of a model with a constant length of hybrid region. We cannot exclude, however, some marker-independent background contribution from hybrid regions much shorter than the mean.

It is interesting to compare these repairability data with earlier results on the recovery of *rIIB* mutants from heterozygous T4 particles (DRAKE 1966). The mutations *FC1*, *UV375*, *X511*, *FC9*, *FC0* and *FC40* have been characterized both by DRAKE and by us. Quantitative comparison of the data is not feasible, since we measured repairabilities to the wild-type alleles only, whereas in the experiments of DRAKE, the marker occurrence in heterozygotes depended inversely on mismatch repair to both wild-type and mutant alleles. Nevertheless, a qualitative fit is apparent: *UV375*, *X511*, *FC9*, *FC0* and *FC40*, occupying lower

TABLE 4

The nucleotide structure of mismatches formed by rIIB *mutations of* T4 *phage as inferred from the DNA base sequence determined by* PRIBNOW et al. *(1981).*

Alleles *a/A*	Reciprocal mismatches	Rate of *a*→*A* conversion
HE122/+	G/G; C/C	Low
UV375/+	T/C; A/G	Low
UV375/amUV375	A/C; T/G	Low
amUV375/+	TAG/CTT; ATC/GAA	Low
360/+	T/G; A/C	Low
360/op360	A/C; T/G	Low
360/am360	A/C; T/G	Low
op360/+	TG/GT; AC/CA	High
op360/am360	GA/TC; CT/AG	Low
am360/+	TAG/GTT; TTC/CAA	High
am360/op360	TC/GA; AG/CT	High
X511/+	T/C; A/G	Low
UV357/+	A/A; T/T	Low
opUV357/+	GA/TA; CT/AT	Low
amUV357/+	G/A; C/T	Very high

positions in our repairability diagrams, were shown by DRAKE to exhibit high recovery frequency from heterozygotes, whereas the readily repairable *FC1* was recovered less frequently.

What determines the repairability differences of T4 markers? As exemplified by the occurrence of amber → + conversions at low (*amUV375*), moderate (*am360*), and high (*amUV357*) rates, the base sequence of the allele repaired seems not to be critical in itself. This observation suggests the importance of the sequence to which an allele is opposed in hybrid DNA. The more than fortyfold preponderance of *FC40* → *opUV357* over *FC40* → + (Figure 2, diagram VI) is striking evidence in support of the latter suggestion. However, the finding that homoallelic mutations are of different repairabilities when opposed to the same (e.g., wild-type) requires that we consider the allele repaired as well.

In studies with artificial DNA heteroduplexes, the *in vitro* susceptibility of base substitution mismatches to single-strand specific endonucleases was found to increase progressively with the increase in the number of mismatched bases from one to six (DODGSON and WELLS 1977). No similar conclusion could be inferred from our repairability diagrams when the data on base sequence in the *rIIB* segment (PRIBNOW *et al.* 1981) were taken into consideration (Table 4). Although among single-base mismatches most of the transition-like (*UV375/amUV375*, *360/op360*, *360/am360*, *360/+*) and transversion-like (*HE122/+*, *UV375/+*, *X511/+*, *UV357/+*) structures are poorly repaired, the transversion-like mismatch *amUV357/+* converts to +/+ at a very high rate. Among two-base mismatches, *amUV375/+* and *opUV375/+* are poorly repaired whereas *am360/+* and *op360/+* convert to +/+ at a high frequency. Equally, consideration of the base sequences around the mismatched regions gives no clue to the specificity of the marker-discriminating mechanism. We will not be surprised if further investigations prove mismatch repair to be conformation-specific rather than length- or sequence-specific.

LITERATURE CITED

BARNETT, L., S. BRENNER, F. H. C. CRICK, R. G. SHULMAN and R. J. WATTS-TOBIN, 1967 Phase-shift and other mutants in the first part of *rIIB* cistron of bacteriophage T4. Philos. Trans. R. Soc. Lond. (Biol.) **252:** 487–560.

DRAKE, J. W., 1966 Heteroduplex heterozygotes in bacteriophage T4 involving mutations of various dimensions. Proc. Natl. Acad. Sci. USA **55:** 506–512.

KATZ, E. R. and S. BRENNER, 1975 Genetic map of the beginning of the *rIIB* cistron of bacteriophage T4. Mol. Gen. Genetics **143:** 101–104.

PRIBNOW, D., D. C. SIGURDSON, L. GOLD, B. S. SINGER, C. NAPOLI, J. BROSIUS, T. J. DULL and H. F. NOLLER, 1981 *rII* cistrons of bacteriophage T4. DNA sequence around the intercistronic divide and positions of genetic landmarks. J. Mol. Biol. **149:** 337–376.

SHCHERBAKOV, V. P., L. A. PLUGINA, E. A. KUDRYASHOVA, O. I. EFREMOVA, S. T. SIZOVA and O. G. TOOMPUU, 1982 Marker-dependent recombination in T4 bacteriophage. I. Outline of the phenomenon and evidence suggesting mismatch repair mechanism. Genetics **102:** 615–625.

TOOMPUU, O. G. and V. P. SHCHERBAKOV, 1980 Genetic recombination: formal implications of a crossed strand-exchange between two homologous DNA molecules. J. Theor. Biol. **82:** 497–520.

Corresponding editor: J. W. DRAKE