# The interpretation of complementation data 

By OLIVER J. GILLIE*<br>Institute of Animal Genetics, Edinburgh, 9<br>(Received 6 September 1965)

## 1. INTRODUCTION

Previous interpretations of complementation data (Crick \& Orgel, 1964; Kapuler \& Bernstein, 1963) have not considered (a) the relationship between the number of mutants used to define a complementation map and the structure (linear, circular or complex) of the map; (b) the significance of exceptional pairwise tests which are inconsistent with a linear map. In this paper these two points have been investigated by constructing sample maps, using mutants with known complementing properties picked at random from those already described at the leu-2 locus in Neurospora crassa (Gross, 1962), and by comparing this with data from all known complementing loci in micro-organisms.
It will be shown that at the majority of loci investigated many complementing groups remain to be discovered, and that the evidence for most loci having linear complementation maps is quite inconclusive. Indeed, it has not been possible to find more than two loci, out of thirty-five which have been studied, which can be considered with any reasonable likelihood to have a linear complementation map. The results obtained will act as a useful guide to the number of complementing mutants which must be tested at a 'typical' locus before a circular or complex map is likely to be established.

Some of the more detailed non-linear complementation maps are compared and an interpretation of them is made in terms of the theory of interallelic complementation involving axes of symmetry, as proposed by Crick \& Orgel (1964). No attempt has been made to explain complementation data by any model other than that of protein-protein interaction (Brenner, 1959; Fincham, 1960) in view of the mounting evidence in favour of this model (Schlesinger, Torriani \& Levinthal, 1963; Coddington \& Fincham, 1965).

## 2. RULES AND DEFINITIONS USED IN CONNEXION WITH COMPLEMENTATION MAPS

As the rules and definitions used in map construction differ in the literature, those adhered to in this paper are given in detail below:

## Rules:

1. Non-complementing mutants overlap, and complementing mutants do not overlap.

[^0]2. A mutant or group of mutants is represented by a line, or a circle, or any continuous combination of lines and circles, such that the number of complementation units covered by a mutant, and the number of branching points and number of units in the map, is the minimum consistent with Rule 1.

## Definitions:

1. A complementation unit is the distance used to separate two adjacent points on the map, defined by the ends of lines representing complementing groups.
2. A complementation group consists of a number of mutants which all have the same qualitative complementing properties.
It must be pointed out that the complementation unit and the complementation group are always defined operationally with reference to a particular set of mutants, and that an increase in the number of mutants in the sample will usually make it necessary to redraw the map and at least some of the units and groups will be redefined. This will become abundantly clear when the results, given below, obtained from sample sets of mutants are seen.

Rule 2 has not always been recognized, although it has been used by Catcheside (1960), Leupold \& Gutz (1963), and used in part by Dorfman (1964) and Costello \& Bevan (1964). The advantage of using Rule 2 is that no information is lost in the translation of the complementation matrix into the map and the assessment of the data may then be unbiased. Furthermore, Rule 2 leads to a minimization of the complexity of the map and a reduction in the number of possible ways in which a given matrix may be represented. Although Rule 2 minimizes three different features of the map, no conflict has been found in attempting to do this. Whether or not Rule 2 can be satisfied for all possible matrices or only for a certain type of matrix is a mathematical question not without interest; it need not, however, concern us here.

Although the use of Rule 2 leads to the representation of all the data in the form of a map, it is still convenient to consider the small proportion of tests which lead to a considerable complication of the map as exceptional tests. In order to facilitate consideration of maps in this way the terms 'proposed plot', 'positive exception' and 'negative exception' are introduced and definitions, not unduly rigorous, are given below.

## Supplementary definitions:

1. 'Proposed plot'-a preliminary, comparatively simple, complementation map which is consistent with more than $99 \%$ of pairwise tests.
2. 'Positive exception'-a comparatively rare positive (i.e. complementing) pairwise test which is inconsistent with the proposed plot. In consideration of the leu-2 data this term has been extended to cover rare pairwise tests inconsistent with a particular hypothesis rather than a particular plot.
3. 'Negative exception'-a comparatively rare negative (i.e. non-complementing) pairwise test which is inconsistent with the proposed plot. Comple-
mentation units defined by such negative tests never contain groups located exclusively in those units.
Further consideration of positive and negative exceptions and the detailed implications of Rule 2 in map-making are given below.

Figure 1 shows an artificial complementation map constructed to show the types of features which may be encountered in maps constructed using Rule 2. All these features have been encountered in constructing the maps illustrated in shorthand form in Fig. 5, but the size and complexity of those maps which show all these features make them unsuitable to use as examples here. The map consists of four major portions, A, B, C and D. Portions B and C are defined by 'positive' and 'negative exceptions' respectively, and if these exceptions were ignored the map would consist of a circle with a tail, i.e. portions A and D.
'Positive exceptions.' The complementation of group 7 with group 8, but not with any of the other groups crossing units $f$ and $g$, leads to the new units, $i$ and $j$, being defined. Units $i$ and $j$ are equivalent to units $g$ and $f$ except for the presence of group 7 in the former two units and


Fig. I. A hypothetical complementation map to illustrate some of the special features of complex maps. Complementation units are labelled with lower case letters and complementation groups with numbers. The map comprises 20 complementation groups, of which 8 are numbered. See text for explanation and discussion.
group 8 in the latter two units. Groups common to all these units are joined at the ends by special joining lines. Positive exceptions lead to portions of the map being drawn as concentric lines in this way. These concentric lines tend to have the same number of units when the number of positive exceptions is small, but when the number of positive exceptions is larger, as in the ad $5 / 7$ (Dorfman) map (see Fig. 5), then there are a number of units unique to each of the concentric lines of the map; these additional units are defined by virtue of the positive exceptional groups placed uniquely in one or other of the lines.


#### Abstract

'Negative exceptions' and the minimization of branching points. The negative tests $3 \times 5$, $4 \times 5,4 \times 6$ cannot be represented in the map in the normal way because of the more frequent occurrence of positive complementation tests with other mutants separating them on the map; for this reason they must be specially represented by portion C of the map. It is important to note that portion C of the map may be attached to the main A portion of the map in several different ways. This is frequently the case with negative exceptions. Furthermore, it should be noted that portion $C$ could have been represented by two separate loops joining $A$ in four different places, but only one loop has been used in accordance with the requirements of Rule 2, to minimize the number of branching points in the map.

Application of Rule 2 to the minimization of the number of units. Group 1 has been drawn overlapping group 2 only in unit $a$, whereas it could also have been made to overlap group 2 in another unit drawn in between $c$ and $d$; in other words it is arbitrary in this case whether unit $a$ as defined by the pairwise interaction of group 1 and 2 is drawn at $a$ (i.e. between $h$ and $b$ ) or between $c$ and $d$, but in order to comply with rule 2 and minimize the number of units it is only drawn in the position shown because of the existing complementation pattern of group 5 and 2.


## 3. CONSTRUCTION OF MAPS FROM MUTANTS RANDOMLY SAMPLED AT THE LEU- 2 LOCUS

Data of Gross (1962) obtained from mutants at the leu-2 locus in Neurospora was chosen for detailed examination by construction of sample maps, as his data was obtained by making all possible pairwise tests between a large number of complementing mutants. Mutants (complementing and non-complementing) were chosen, using random numbers, 16 at a time, and the complementation map plotted for each increment of 16 mutants. Figure 2 shows how the number of complementation groups increases linearly with the number of mutants in the sample. The slope of the graph shows that for every 10 new complementing mutants mapped at the leu-2 locus approximately 5 new groups are defined. If the number of complementation groups (in Fig. 2) were plotted against the total number of mutants in the sample (complementing and non-complementing), rather than only against the complementing mutants in the sample, the relationship only changes in so far as more scatter is introduced.

There is no sign of the number of complementation groups being defined in diminishing numbers as the number of complementing mutants increases up to 78 , and so it seems certain that many more than 42 complementation groups exist at the leu-2 locus. The reason for this is that if all of the possible complementation groups were found (assuming that there are a finite number of complementation groups), more complementary mutants will still be found, but these on testing will fall into one of the existing groups. As the number of recognized groups increases towards this limit, then new mutants are more likely to fall into one of the already recognised groups, and hence the number of complementing mutants already found will be relatively greater in proportion to the number of new groups defined till the limit is reached.

In two trials the map had to be plotted circularly when the number of complementing mutants in the sample was increased from 21 to 31 in one case and from 17 to 25 in the other case; the map had to be plotted in a complex manner when the number of complementing mutants in the sample were increased from 37 to 45 in one case and from 25 to 33 in the other case. In other words, the maps change from


Fig. 2. The number of complementation groups is plotted against the number of complementing mutants in maps made from random samples of mutants at the leu- 2 locus. Mutants (complementing and non-complementing) were picked 16 at a time in two series. Series 1 is plotted as circles, Series 2 as squares. Open circles or squares indicate a linear map, solid circles or squares a circular map, and crosses with circles or squares indicate a complex map. The line is drawn from the origin to the point defined by the leu-2 map from which the mutants were sampled. The number of complementing mutants is defined operationally in terms of the sample of mutants obtained, and does not represent mutants in the sample which are found to complement in the final map.
being linear to circular, or from being circular to complex, when a mean sample size of 24 complementing mutants (giving 276 pairwise tests and about 14 groups) or 35 complementing mutants (giving 595 pair-wise tests and about 19 groups) are used respectively. Clearly loci which in fact have circular or complex complementation maps would not be recognized as such unless a sufficient number of mutants were tested. This problem will be discussed in relation to the known data in the following section.

Figure $3 a$ shows the number of units plotted against the number of groups found in sample maps at the leu-2 locus. It can be seen that for every 10 groups defined about 6 units are defined, and that the relationship is linear. Figure $3 b$ shows the number of units plotted against the number of groups for data from all known loci (see Table 1). The relationship is much the same as in Fig. $3 a$, up to about 45 groups, and then the slope increases when the data from the complex maps ( $a d 5 / 7$ and $\mathrm{his}-\mathrm{B}$ ) is plotted. The reason for this is that when the data is plotted to accommodate all 'exceptions' many groups as plotted on the map have more than two ends, and


Fig. 3a. The number of complementation groups is plotted against the number of complementation units in maps constructed from randomly selected mutants at the leu-2 locus, as explained in the legend to Fig. 1. Series 1 is plotted as triangles, Series 2 as squares. The line is drawn from the minimum number of units and groups (two) to the point defined by the number of units and groups in the complete leu-2 map.
Fig. 3b. The number of complementation groups is plotted against the number of complementation units for all loci from Table 1. Line A represents the maximum possible number of groups for a given number of units in a linear map, and line B represents the maximum possible number of groups for a given number of units in a circular map (Carlson, 1961).
hence they define a larger number of units. If, for example, the $a d 5 / 7$ data is plotted, ignoring the exceptions, then the proportion of units (33) to groups (67) is quite normal. The ratio of units to groups will be affected by a number of variables, including the frequency of groups of length $1,2, \ldots n$ units, the complexity and the symmetry of the 'ultimate' map from which any map is a sample. In view of this it is perhaps surprising that the ratio of units to groups for most loci is fairly similar to that found at the leu-2 locus (cf. Figs. $3 a$ and $3 b$ ). Except for maps made from
very small samples of mutants, the maximum number of groups (Carlson, 1961) possible with a given number of units is never found (whatever the assumption about the actual form of the map-linear, circular or more complex). This suggests that all maps are very asymmetrical, very incomplete, or both.

## 4. ANALYSIS OF DATA FROM COMPLEMENTING LOCI

The available data from all known complementing loci in micro-organisms is summarized in Table 1. The data has been obtained from 35 loci in 6 different organisms. It must be pointed out that a variety of mutagens and selective methods


Fig. 4. The number of complementation groups is plotted against the number of mutants for all loci from Table 1. Open circles represent linear maps, solid circles represent circular or complex maps.
have been used to obtain the mutants, and it cannot always be assumed that the mutants mapped are a random sample of those obtained with a given mutagen. A number of workers have chosen mutants in special ways (e.g. temperature sensitive, leakiness) to select in favour of complementation, and moreover, in several cases all possible pairwise tests have not been performed, and hence the complementation map has had to be constructed from incomplete data. Wherever possible these special conditions have been recorded in the notes appended to Table 1.

If the number of complementation groups is plotted against the number of complementing mutants using the data from all known loci in Table 1, mostof the loci (see
Table 1. Summary of published data on complementation in micro-organisms

 Ishikawa, 1962
Kapuler, 1963 Kapuler, 1963
Fincham, 1963

 $=$
菏 8.8
$0_{2}$
0.
0.0
0
0
0
0
0
 $=$ Gross, 1962
Gross, 1963
Muray, 1960
Case, 1960
Giles, 1963
Woodward, 1962
Ahmad, 1960
Suyama, 1962
Ahmad, 1964
Catcheside, 1964
 de Serres, 1963
de Serres, 1965
Woodward, 1958 Woodward, 1958
Ishikawa, 1962 Linear
Circular Linear
Linear
Linear
Linear
Complex
Linear
Linear
Linear
Linear
Linear
Linear





 0
 $\infty \infty$ Percentage of mutants comple-
menting
 $\infty$
 $\quad$ No. of
mutants
comple-
menting

 .

 $\begin{array}{lr}\quad \text { Locus } & \text { mutants } \\ \text { Neurospora crassa } \\ \text { ad-3b } & 151 \\ \text { ad-3b } & \sim 235 \\ \text { ad-4 } & 124 \\ \text { ad-8 } & 385\end{array}$
 ت


Schizosaccharomyces pombe
공 家 家

Linear
Linear
Complex
Complex
Linear
Linear
Complex
Complex
$L=$ leaky mutants. $\quad \dagger$ First author's name only is quoted.
For notes to table see following page.

| 4 | UV, DS, SP |
| :---: | :---: |
| 9 | UV |
| 14 | UV |
| 21 | UV, X, NA |
| 5 |  |
| 4 | UV |
| 66 | SP, EMS |
| 40 | UV, EMS |
| $\left.\begin{array}{l} 22 \\ 17 \end{array}\right\} 59$ | UV, SP, X |
| 3 |  |
| 4 | SP, X, UV, A |
| 30 | Various |
| 2 | SP, X, UV, A |
|  | Various |
| 4 | SP, X, UV, A |
| 5 | Various |
| 6 |  |
| 7 (7) | AP, NA, BUD |
| 7 (5) | AP, NA, BUD |
| 7 (7) | AP, NA, BUD |
| 10 (10) | AP, NA, BUD |







## Notes on the compilation of the data in Table 1

The data represents all the examples in the literature of complementation in micro-organisms known to the author up to about May, 1965. Where weak or doubtful complementation responses have been noted in the literature these have always been scored as positive although in fact these cases are sufficiently uncommon that the figures are not substantially affected. Maps are classified as linear, circular or complex; the latter category consists of all non-linear or non-circular maps. All non-linear maps are shown in 'shorthand' form in Fig. 5. Original maps have been replotted as necessary to conform to the rules stated at the beginning of the paper. Individual notes explaining unusual features of the data at each locus are given below.

Abbreviations used for mutagens are as follows:
UV-ultraviolet, X-X-rays, $\gamma-\gamma$-radiation, SP-spontaneous.
Chemicals: AP-2-aminopurine, BUDR-5-bromodeoxyuridine, DS-diethyl sulfate, ETTM-ethyleneimine, EMS-ethyl methane sulfonate, FUDR-5-fluorodeoxyuridine, $\mathrm{HP}+\mathrm{F}$-hydrogen peroxide with formaldehyde, L-lysozyme, NA-nitrous acid, $\beta$ PL-$\beta$-propriolactone, TBHP-tert-butylhydroperoxide.
$a d-3 b$-Figures for the number of groups and units defined by the leaky and non-leaky mutants are given as defined by the mutants within these groups and not as defined in the combined map.
ad-8-308 primary mutants and 77 secondary mutants were investigated. Secondary mutants were obtained by treating several different homocaryotic revertants obtained from primary mutants with UV. All possible pairwise tests were not made.
ad-4-44 out of 51 complementing mutants were secondary mutants induced in three different revertants.
arg-10-Rice (1963) found only three complementation groups when the same alleles were tested by complementation in the St Lawrence wild-type background as opposed to the Emerson wild-type background used by Catcheside (1965b).
his-1-A more recent investigation (Jessop, 1965) has shown that the map is probably not complex, but was thought to be due to a high frequency of reversion to prototrophy of one of the mutants.
$i v-2$-Only 832 out of 4371 possible pairwise tests were made.
$i v$-3-Only 10,500 out of the 93,096 possible pairwise tests were made.
leu-2-Mutants used in construction of the map were selected for complementing properties by trial tests with 13 known complementers.
ad-6-Only 771 of the 946 possible pairwise tests were made for the pooled data of Leupold and Gutz.
ad $5 / 7$ (Costello)-Of the 77 ad- 7 mutants isolated only 59 were tested by complementation.
ad-2-Of 94 complementing mutants isolated 48 non-leaky mutants and 35 leaky mutants were used to construct separate maps for the non-leaky and leaky mutants. 78 mutants were used to construct the final map shown in Fig. 5,5 leaky mutants with anomalous complementation responses being omitted. Figures for the number of groups and units defined by the leaky and non-leaky mutants are given as defined by the mutants within these groups and not as defined in the combined map.
$r$ IIA-The mutants used were selected for leakiness.
37, 10, 12, 34-were with a few exceptions obtained as temperature-sensitive mutants. Two mutants at locus 12 and one at locus 34 were isolated as amber mutants. The figures are given for a burst size $0.75 \%$ or greater of the wild-type burst size being scored as positive complementation. In brackets figures are given for a burst size $2 \cdot 5 \%$ of wild-type or greater.

Fig. 4) fall roughly on the line expected from the analysis of the relationship between the number of groups and number of mutants found in the sample maps at the leu-2 locus (cf. Fig. 2). The most marked exceptions to this are $a d-8$ and $i v-3$ (not shown in Fig. 4), which may be discounted, as only a portion of the possible pairwise tests were made in mapping these mutants, and many groups no doubt have not been identified.

It is also clear from Fig. 4 that maps with the most number of mutants or groups tend to be circular or complex, and those with the smallest number of mutants or groups tend to be linear. This encourages one to believe that many linear maps would become circular or complex if more mutants were mapped at the loci in question.

## 5. ATTEMPT TO ESTIMATE THE TOTAL NUMBER OF COMPLEMENTATION GROUPS (KNOWN AND UNKNOWN) AT A LOCUS

If the total number of groups at a locus could be estimated from the frequency of mutants per group already known, then some idea of the completeness of the map and the amount of work necessary to complete it could be obtained. This would be particularly helpful in deciding whether maps were genuinely linear or not.
General statistical methods (e.g. Good, 1953) exist which can be used to calculate the number of species in a population from the number of species represented by one individual. Unfortunately this method does not seem to be of use in estimating the total number of complementation groups at a locus, as the groups are not defined independently of each other and the discovery of new groups is often accompanied by the splitting up of old groups. At best such methods may be used to estimate the minimum number of unknown groups at a locus. However, the existence of complementation groups represented by only one mutant at a locus must be taken as an indication of the existence of further, as yet undefined, complementation groups at that locus. The number of groups at each locus which are represented by only one mutant is given in Table 1, column 6.

All the loci examined in T4 phage show a particularly high proportion of groups with only one mutant. This may no doubt be attributed to the special means of selection used, which seems to favour complementing mutants.

## 6. CRITERIA FOR LINEARITY OF A COMPLEMENTATION MAP

It seems likely for the reasons already discussed that many linear maps will, as a result of further tests on larger samples of mutants, become complex. There may nevertheless exist maps which are genuinely linear after a large number of tests have been performed. It was determined at the leu-2 locus that a mean figure of 24 complementing mutants, or 13 complementation groups, had to be exceeded before the map became non-linear; these figures can be taken as fairly representative of other loci since leu-2 has about $3.0 \%$ of pairwise tests exceptional to linearity, this figure varying between 1 and $8 \%$ for various other loci (see Table 2). The only linear maps to have more than 13 complementing groups are $p y r-3$, pan-2, $a d-3 b$, $a d-4$ and $h i s-3$ (Ahmed et al., 1964). Of these only his-3, with 25 groups, and ad-3b, with 35 groups, are sufficiently large at present for it to be likely that they may be genuinely linear, or at least have a lower frequency of pairwise tests exceptional to linearity.

The specificity of various mutagens with respect to the type of complementary mutants they induce (de Serres, 1964; Costello \& Bevan, 1964; Gutz, 1963b), the
likelihood of various selection procedures of producing leaky and non-leaky mutants (de Serres, 1964) and the techniques and scoring criteria used in different laboratories (cf. Catcheside's and Ahmed's data for the number of complementation groups at the his-3 locus in Table 1) may all affect the type of complementation groups identified and hence the form of the map. For this reason it will always be more difficult to establish that a map is linear than that it is non-linear.

## 7. NON-LINEAR COMPLEMENTATION MAPS

In view of the preceding discussion it seems likely that with the possible exception of $a d-3 b$ and $h i s-3$, many linear maps would be non-linear if the number of mutants comprising the map were increased or the sensitivity of the tests improved. This is a most important conclusion as it may account for much of the difficulty which has been found in correlating genetic maps and functional data with complementation maps. Although precise correlations of these different types of data are not expected on a priori grounds, interpretations will often be misleading so long as the complementation map is of undetermined form.

In the course of the collection of all the data presented in Table 1, several maps which had been plotted linearly or circularly were replotted according to the rules given above. These and most other non-linear maps are shown in Fig. 5, and the frequency of exceptions necessitating the non-linear plots are given in Table 2. Leu-2, his-B, ad-2, ad-6 and both ad 5/7, maps are of special interest as they have exceptions not only to linearity but also to circularity in the case of leu-2 and his-B, and exceptions to the general form of the map defined by most complementation groups in the case of the other loci. These exceptions to circularity or to the general form of the map have been described as exceptions to the 'proposed plot', as they are much less frequent than exceptions to linearity such that if about $0.5 \%$ of selected pairwise tests were ignored, then the matrix could be completely represented by the 'proposed plot', i.e. by ignoring comparatively rare exceptions a 'proposed plot' may be made which is comparatively simple (see above for the definition of 'proposed plot' and more detailed discussion of the term). The 'proposed plot' is in fact the simplified version of the maps which in most cases have been already published (leu-2, ad-6 (Leupold), ad 5/7 (Dorfman), ad 5/7 (Costello), his-B.) In replotting the data of Leupold \& Gutz (1964) to accommodate the additional data of Gutz (1963a) the 'proposed plot' was found to be of the same general form as the plot used in the 1964 paper.

Exceptions to the 'proposed plot' are of two kinds: those which are positive and hence show complementation instead of non-complementation, and those which are negative and hence show non-complementation instead of complementation. When data is replotted according to the rules given above, this distinction between positive and negative exceptions may still generally be made, but some of the negative exceptions may become accommodated with replotting of positive exceptions, and a few exceptional interactions may be regarded as positive or negative depending on which individual groups are regarded as exceptional and what is considered to be


Fig. 5. The complementation maps of circular and complex loci are drawn in a shorthand form in this figure. The number associated with each complementation unit represents the number of groups which are inactive in this unit, i.e. the number of lines which cross this unit in a complementation map. Dotted lines are features of the map which are defined by 'negative exceptions' (see text). The map for ad-6 has been drawn twice, once using only the data of Leupold (1964) and a second time using the data of both Leupold (1964) and Gutz (1963); the former map is labelled $a d-6(\mathrm{~L})$ and the latter map $a d-6(\mathrm{~L} \& \mathrm{G})$.
the 'proposed plot'. There is sometimes ambiguity also about the positioning of negative exceptions on the map when the exception involves two large groups on the map, as the interaction may sometimes be drawn between either of the two sets

Table 2. Loci requiring non-linear plots

| Locus | No. of pairwise tests made | No. of pairwise tests exceptional to linearity | Percentage exceptional to linearity | No. of pairwise tests exceptional to 'proposed plot' |  | Percentage exceptional to 'proposed plot' |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | -ve | +ve | -ve | + ve |
| leu-2 | 3003 | 97 | $3 \cdot 2$ | 6 | 0 | 0.2 | 0 |
| leu-4 | 1176 | 19 | $1 \cdot 6$ |  |  |  |  |
| $a d-6$ (Leupold) | 378 | 12 | $3 \cdot 2$ |  |  |  |  |
| ad. 6 (Leupold | 771 | 24 | $3 \cdot 1$ | 2 | 0 | 0.3 | 0 |
| \& Gutz) |  |  |  |  |  |  |  |
| ad 5/7 | 5460 | 347 | 6.4 | 17 | 19 | 0.3 | $0 \cdot 3$ |
| (Dorfman) |  |  |  |  |  |  |  |
| ad 5/7 | 1711 | 140 | 8.2 | 7 | 2 | 0.4 | $0 \cdot 1$ |
| (Costello) |  |  |  |  |  |  |  |
| his B | 2485 | 53 | $2 \cdot 1$ | 6 | 11 | 0.2 | 0.4 |
| $r$ IIA | 55 | 1 | 1.8 |  |  |  |  |

See text for explanation of the term 'proposed plot'. The proposed plot of leu-2 and his B is circular and that of both ad 5/7 maps is the circle with tail form. The proposed plot of ad-6 is the original plot of Leupold (Fig. 4). It must be noted that there is no unique way of resolving a complex map into a linear one and so the number of pairwise tests exceptional to linearity have been enumerated by 'cutting' the map in the place where least mutants are involved and counting the number of negative tests which are eliminated by the 'cut'. ad-2 has not been included in this table as there is no obvious way of linearizing the plot without cutting the map in three places and representing it as two quite separate portions.
of ends (see discussion of Fig. 1 above). When the map is arrived at by the application of the rules discussed at the beginning of the paper it should be possible to re-derive the original matrix from the map without any loss of information. Nevertheless small variations, discussed above, exist in the way in which a map may be drawn; these variations exist because alternative ways of drawing the map may result in the same number of mutants being exceptional to the proposed plot, or what amounts to the same thing, the number of branching points and number of units being the same. These variations are, however, rare or trivial, and have no serious effect on the form of the map. Since there is at present no mathematical method of effecting the transformation between matrix and map, maps are constructed by a process of inspection, and so there is no rigorous way of showing that a map has only a limited number of variations or that the particular map chosen makes the smallest number of groups exceptional.

Figure 5 shows the complementation maps drawn in a shorthand form. Each complementation unit has a number attached to it which represents the number of groups which cross that unit. Using Fig, 1 as an example, unit $a$ would have the number 5 associated with it and unit $d$ the number 7 . This number may be thought of as some function of the frequency with which complementing mutants are likely to occur in that part of the map. Negative exceptions to the 'proposed plot' are
shown in these maps by dotted lines, positive exceptions to the 'proposed plot' lead to portions of the map being drawn as concentric lines, i.e. positive and negative exceptions to the 'proposed plot' are included in the maps (Fig. 5) by application of the rules already discussed.

## 8. THE INTERPRETATION OF COMPLEX COMPLEMENTATION MAPS

In order to attempt to account for complementation data which must be plotted as a complex map in terms of any current hypothesis, the following features of the map must be explained: the general form (i.e. circle, double circle, circle with tail,


Fig. 6. A plot to show the polar distribution of leu-2 mutant groups. The central point (or mean position) of each mutant is represented by a square. The central point of a mutant group is defined as that point on the line representing the mutant on the map, which divides the line into two halves with equal proportions of complementation units. As groups may have odd or even numbers of units, the number of points at which groups may be placed in this way is twice the number of complementation units in the map. The diagram shows that the mutants may be divided into two types, A and B. These types have the property that (with one exception) mutant groups within the types A or B do not complement one another.
etc.) of the map, the distribution of mutants on the map, and the occurrence of positive and negative exceptions to the general form ('proposed plot') of the map. An attempt has been made to do this for the data of Gross (1962) at the leu-2 locus of Neurospora in terms of the theory of Crick \& Orgel (1964).

According to the concept of Crick and Orgel, each complementing mutant must
have a derangement of structure (probably of folding structure) near the axis of rotation of a multimeric protein. If the derangement was away from the axis of symmetry, then the fault in each monomer would be expected to be correctable by another monomer of the same kind, in which case there would not be a mutant phenotype in the first place. In the subsequent discussion of the leu-2 data it will be assumed that the multimer in question is in fact a dimer.

It seems reasonable to suggest that the two 'poles' of the Neurospora leu-2 complementation map (see Figs. 5 and 6) represent two regions of the protein monomer which come close to the axis of symmetry of the dimer. On this interpretation one would expect no complementation among the mutants at any one pole of the map, since all must have defects in their respective protein monomers overlapping the same region near the dyad axis. It is indeed the case that mutants at pole A (Fig. 6) show no complementation amongst themselves but only with mutants at pole B, while mutants at B (with only one exception involving mutants D6 and D18) also show no complementation amongst themselves but only with A mutants.

What has to be explained is that some A-B pairs are non-complementary and that there seem to be two ways of getting such non-complementation, represented by the two 'sides' of the circular map. It seems possible to offer an explanation in terms of a model such as that illustrated in Fig. 7, with the interface between the


Fig. 7. Diagrammatic representation of protein molecule to illustrate axes of symmetry. Part $a$ shows a dimer consisting of two identical monomers joined at the interface A, B, C, D, E, F. Part $b$ shows the two monomers separated to show the interface at which they are normally joined together. See text for discussion.
two protein monomers in the form of a closed loop. Owing to the reversed orientation of one monomer relative to the other (to give the postulated symmetry), homologous portions of the monomers are in contact only at $A$ and $B$, which are
taken as corresponding to the 'poles' of the complementation map. Away from the dyad axis non-homologous portions are in contact, C with D and D with C . It is supposed that a derangement of structure of one monomer can be corrected provided it is 'covered' by a normal portion of the other monomer, whether that normal portion is homologous with the damaged portion or not. As pointed out above, the damaged region must include a portion near the dyad axis, otherwise the mutant monomer would be correctable by another monomer of the same mutant kind.

Given this model there are two possible ways in which A and B mutants could be non-complementary. A derangement including A could spread into C and overlap with a derangement in $B$ spreading into $D$. Or the derangement in $A$ could spread into D and overlap with a derangement in B spreading into C . Some mutants, for instance centred on A , could have effects spreading both into C and into D , in which case they might overlap some $B$ mutants on the one side and other $B$ mutants on the other. The complementation map affords a number of examples of mutants which would have to be explained in this way. As the protein is represented in Fig. 7 the spreading effect from B into C would have to be transmitted through some interpolypeptide bond, such as a disulphide bridge.

Thus, on the interpretation offered here, the circular complementation map is taken as indicating a 'circular' interface between the monomers in the dimer. A superficially similar argument from the shape of the complementation map to the shape of the protein monomer has been made by Kapuler \& Bernstein (1963) in the case of the Neurospora ad- 8 locus. However, the model proposed by these authors is quite different from that proposed here. They envisaged a stack of more or less planar monomers, like a pile of pennies all 'heads'-side up. In such an aggregate homologous parts of monomers are superimposed over the whole interface, and each portion of the complementation map can be interpreted as corresponding to a particular part of the monomer. There is no axis of symmetry in an aggregate of this sort. In the model presented here the map does not represent the different parts of a single monomer but rather the interface between two monomers, i.e. the relative positions of the groups at pole $A$ of the lex-2 map are defined by the other mutants at pole $B$ of the map and vice versa, where any mutant from a group at pole $A$ is always in one monomer and any mutant from a group at pole $B$ in the other monomer of the dimer. At a given pole A or B of the map the groups are regarded as being superimposed upon one another so as to show the relative extent to which they cause a derangement of the folding of the monomer. The weight which one gives to this interpretation as against the apparently more straightforward model of Kapuler and Bernstein must depend on the strength of one's conviction that dimeric (or, more generally, multimeric) proteins should have axes of symmetry, and with particular reference to the leu-2 data, on the ability of the two models to account for the complexity of the map and the asymmetry in the distribution of the complementation groups.
Using the model presented here, it is possible to account for the negative exceptions which form the 'ears' of the leu-2 map in Fig. 5. Mutants giving negatively exceptional tests may be assigned to portions E and F of the monomer in

Fig. 7, on the grounds that a mutant at E or F will usually interact with other mutants via the axes of rotation (possibly via disulphide bonds linking E with A and F with B ) at A or B , and hence be placed with these mutants. A mutant at E may, however, interact directly with a mutant at $F$ to prevent complementation, and such mutants would consequently be classed as 'negative exceptions'. A detailed explanation of this kind would have to account for the existence of two types of negative exceptions represented by each of the two 'ears'. Such a model is quite possible, but it does not seem worth while to give a detailed description in view of the comparatively small number of pairwise tests which are involved. This interpretation of the leu-2 data then accounts for all the pairwise tests except one ( $\mathrm{D} 6 \times \mathrm{D} 18$ ), which has been attributed to a special interaction of the type where 'two blacks make a white'. The frequency of tests of this kind at the leu-2 locus is then $0.03 \%$, which supports the conclusion of Crick \& Orgel (1964) that this type of interaction was unlikely to be the usual explanation of complementation.

It is important to note that this interpretation has assumed that in general misfolding spreads along the polypeptide chain and via disulphide bonds, and it is on these grounds that a dimer with a circular interface has been proposed to explain the data. However, it must be remembered that the complementation map is a topological projection, and that the circular portion of the leu-2 map representing the protein interface may have innumerable kinks and bends, provided that these kinks and bends are made by deforming the structure without cutting it.

The interpretation of the leu-2 data might be queried on the grounds that the leu-2 mutants used to construct the map were pre-selected for complementation with 13 tester strains, and that this procedure might introduce a bias which would account for the symmetry of the map. The structure of the leu-2 map may be simply tested experimentally by looking for any mutants which will complement with both of two testers, one representing pole $A$ and the other pole $B$; on the basis of the present model such mutants would be expected to be very rare. Another example of a complex complementation map which consists basically of two sub-groups within which little or no complementation occurs is the his-D locus in Salmonella (Loper et al., 1964).

Clustering of complementing mutants on recombination maps might be expected in certain cases, as a result of complementing mutants being located at axes of symmetry. For example, in the case of the leu-2 locus, if the polypeptide chain were coiled as is shown in Fig. 7, then two major clusters of complementing mutants would be expected on the recombination map. The complementing mutants involved in the 'negative exceptions' to the complementation map would be expected to be located somewhere outside the two major clusters of mutants on the recombination map. Unfortunately the situation need not be as simple as this, and the circular amino-acid chain forming the interface of the dimer need not, of course, be formed from an unfolded length of polypeptide chain. No recombination map has yet been made for the leu-2 locus, and so the picture must remain speculative. However, there are other possible hypotheses to account for the clustering of sites on the recombination map (Holliday, 1964), and if clustering of sites is caused by
several factors, as seems likely, then any simple correlation of the type mentioned might be precluded.

It is not possible to account for all the other complex complementation maps in such a detailed way. The number of mutants needed to resolve and define axes of symmetry in the same way as has been done in the leu-2 case, will increase with the number of axes of symmetry; and furthermore, in certain polymeric arrangements different axes of symmetry may overlap to confuse the interpretation of the complementation map obtained. A situation of this kind is suggested by the complex general form of certain maps ( $a d-6, a d 5 / 7, a d-2, l e u-4$ ).

The general form of the ad-6 map in Schizosaccharomyces pombe has been found to be the same when 16 additional complementing mutants tested by Gutz (1963) were included (both maps are illustrated in Fig. 3). In view of the general correlation between the complementation map and the recombination map at this locus (Gutz, 1963; Leupold \& Gutz, 1964) it seems particularly reasonable to assume that the complementation map represents a portion of the polypeptide chain which rests at a protein-protein interface. A map of this kind could be generated by a dimer, but equally well by a coiled protein molecule equivalent to a two-turn spiral. Such a model might appear to be similar to the kind of structure which has been proposed by Kapuler \& Bernstein (1963) to explain the results of Ishikawa (1962) for the $a d-8$ locus in Neurospora. However, to explain the ad- 6 data, using the axes of symmetry model of Crick and Orgel, the spiral monomer would have two different interfaces $A$ and $B$, corresponding to the two circles in the map, such that face $A$ would combine with another A face and similarly $B$ with $B$; this arrangement would result in two uncombined faces at the ends of the stack, unless there were an even number of monomers arranged in a ring such that every face combined with another like face to form a closed polymer. The closed polymer type of structure as opposed to a stack of monomers has been considered more likely on general grounds (Crick \& Orgel, 1964; Monod, Wyman \& Changeux, 1965).

The map of the $a d$-2 locus (Fig. 5) in Saccharomyces cerevisiae also shows a general symmetry similar to that shown by the ad-6 map and may be accounted for in the same general way. There are, however, three major portions of the ad-2 map which might be taken as implying three major areas of the ad-2 monomer concerned with the binding of homologous polypeptides in the multimer. As it does not seem to be possible to construct a closed polymer using more than two homologous binding sets (Monod et al., 1965), it seems likely that either all the portions of the ad-2 map form one binding set in a dimer, or two portions of the map form one binding set and the third portion the second binding set in a larger closed multimer with an even number of sub-units, or alternatively three binding sets are involved in forming a large unclosed polymer. At least two binding sets (implying that the molecule is at least a tetramer) seem likely, as it is not easy to see how one axis of symmetry could generate such a complex map.

The his-B locus in Salmonella typhimurium shows some symmetry in the position of the positive and negative exceptions, but both ad 5/7 maps are markedly asymmetrical in this respect. A larger sample of mutants was used to construct the ad $5 / \gamma$
(Dorfman) map, and so a correspondingly large number of positive exceptions were found. This accounts for the major difference between the two ad $5 / 7$ maps. Positive and negative exceptions are sufficiently rare ( $<1 \%$ of all pairwise interactions) for it to be expected that their distribution might be markedly affected by sampling error in cases where these exceptions may be scattered over the map. Negative exceptions have been accounted for in the case of leu-2, and there seems to be no reason why this type of explanation should not account for negative exceptions in other cases. Positive exceptions represent the type of interaction referred to by Crick \& Orgel (1964) as 'two blacks make a white', and seem most likely to be accounted for in terms of particular amino-acid substitutions which favour compensation. It is important to note that detailed consideration of the leu-2 data revealed a 'positive exception' which was defined with reference to a particular interpretation of the map, rather than with reference to the map itself.

Recently, negative complementation (Bernstein et al., 1965; R. A. Woods, 1963) has been discovered, in which it has been found that two leaky mutants grow more vigorously apart than they do together. A study of the type of amino-acid substitutions which give rise to negatively complementing pairs of mutants (two greys make a black) and to pairs of mutants positively exceptional to the 'proposed plot' (two blacks make a white or grey) should give valuable information about the nature of the bonding interactions between protein monomers.

Although most complementation maps may be non-linear or complex when large samples of mutants are analysed, the occurrence of linear maps containing three or more complementation groups seems to reflect an underlying tendency towards linearity in the maps, since circular maps may be constructed artificially with as few as three groups. This situation is most easily understood if we consider circles, double circles and so on as topological structures which may be reduced to a linear form by cutting them in one or more places. With small samples of mutants, cuts or gaps in circular complementation maps are very likely to occur, and so the structure will inevitably be reduced to a simpler form. This tendency will be increased if the distribution of mutants is asymmetrical. In so far as complex maps reflect the organization of the polypeptide chain with its secondary and tertiary structure, linear maps constructed from small samples of mutants may be said to reflect the underlying linearity of the complex maps (as discussed above) and hence the linearity of the protein primary structure.

## SUMMARY

The evidence for complementation maps being linear is examined by analysis of all known complementation maps in micro-organisms, and by constructing maps from mutants randomly sampled from amongst those at the leu-2 locus in Neurospora with known complementing properties. Eleven loci out of thirty-five examined in six micro-organisms have non-linear complementation maps. Two linear maps, $h i s-3$ and $a d-3 b$ (having 25 and 35 complementation groups respectively) have a sufficiently large number of groups for it to be likely that if they do not remain
linear on testing further mutants, they will at least have a lower frequency of mutants exceptional to linearity than known non-linear loci. On the basis of maps made from mutants sampled from the leu-2 data, it seemed unlikely that nonlinearity would be observed with less than 24 complementing mutants or 13 complementing groups in the sample, and therefore many loci with linear maps are likely to be found to have non-linear maps when larger samples of mutants are tested. This conclusion is important in attempting to correlate the structure of complementation maps with recombination maps and with functional data concerning enzyme activities.
The relationship between the number of complementing mutants, number of groups and number of units at the leu-2 locus is described and a statistical method of determining the total number of groups at a locus is discussed.

Known complex complementation maps have been replotted according to consistent rules, and are illustrated in a shorthand form. The form of the complex maps is discussed in relation to current hypotheses concerning the interpretation of complementation maps. In particular an interpretation of the 'circular' leu-2 map is given in terms of the theory of complementation proposed by Crick \& Orgel (1964).

I should like to thank Dr Alan Robertson for drawing to my attention the work of I. J. Good; and also many other people who have helped me to trace the literature, answered specific questions about their data and sent me unpublished material, and in particular Dr J. R. S. Fincham for reading the manuscript and making many helpful comments and constructive criticisms.

## REFERENCES

Ahmad, M. \& Catcheside, D. G. (1960). Physiological diversity amongst tryptophan mutants of Neurospora crassa. Heredity, Lond. 15, 55-64.
Ahmad, M., Khamil, M. D., Khan, N. A. \& Mozmadar, A. (1964). Structural and functional complexity at the tryptophan-1 in Neurospora crassa. Genetics, 49, 925-933.
Ahmed, A., Case, M. E. \& Glies, N. H. (1964). The nature of complementation among mutants in the histidine-3 region of Neurospora crassa. Brookhaven Symp. Biol. 17, 53-65.
Bernstein, H. \& Muler, A. (1961). Complementation studies with isoleucine valine mutants of Neurospora crassa. Genetics, 46, 1039-1052.
Bernstein, H., Edgar, R. S. \& Denhardt, G. H. (1965). Intragenic complementation among temperature sensitive mutants of bacteriophage T4D. Genetics, 51, 987-1002.
Brenner, S. (1959). In Symposium on Biochemistry of Human Genetics (G. E. W. Wolstenholme \& C. M. O'Connor, eds.), p. 304. London: Churchill.
Carlson, E. A. (1961). Limitations of geometrical models for complementation mapping of allelic series. Nature, Lond., 191, 788-790.
Case, M. E. \& Giles, N. H. (1960). Comparative complementation and genetic maps at the pan-2 locus in Neurospora crassa. Proc. natn. Acad.Sci. U.S.A. 46, 659-676.
Catcheside, D. G. (1960). Complementation among histidine mutants of Neurospora crassa. Proc. R. Soc. B, 153, 179-194.
Catcheside, D. G. (1964). Interallelic complementation. Brookhaven Symp. Biol. 17, 1-14.
Catcheside, D. G. (1965). Multiple enzyme functions of a gene in Neurospora crassa. Biochem. biophys. Res. Commun. 18, 648-651.
Champe, S. P. (Personal communication.)
Coddington, A. \& Fincham, J. R. S. (1965). Proof of hybrid enzyme formation in a case of inter-allelic complementation in Neurospora crassa. J. molec. Biol. 12, 152-161.
Costello, W. P. \& Betan, E. A. (1964). Complementation between ad $5 / 7$ alleles in yeast. Genetics, 50, 1219-1230.

Crick, F. H. C. \& Orgel, L. E. (1964). The theory of inter-allelic complementation. J. molec. Biol. 8, 161-165.
Dorfman, B. (1964). Allelic complementation at the ad $5 / 7$ locus in yeast. Genetics, 50, 1231-1243.
Fincham, J.R.S. (1960). Genetically controlled differences in enzyme activity. Adv. Enzymol. 22, 1-43.
Fincham, J. R. S. \& Coddington, A. (1963). The mechanism of complementation between am mutants of Neurospora crassa. Cold Spring Harb. Symp. quant. Biol. XXVIII, 517-527.
Giles, N. H. (1963). Genetic fine structure in relation to function in Neurospora. Genetics Today. Proc. XI Int. Congr. Genet. 2, 17-30.
Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. Biometrika, 40, 1953.
Gross, S. R. (1962). On the mechanism of complementation at the leu-2 locus of Neurospora. Proc. natn. Acad.Sci. U.S.A. 48, 922-930.
Gross, S. R.\& Webster, R. E. (1963). Some aspects of interallelic complementation involving leucine biosynthetic enzymes of Neurospora. Cold Spring Harb. Symp. quant. Biol. XXVIII, 543-547.
Gutz, H. (1963a). Untersuchungen zur Feinstruktur der Gene ad7 und ad6 von Schizosaccharomyces pombe find. Habilitationschrift. Technischen Universität Berlin.
Gutz, H. (1963b). Abstract. Studies on the genetic fine structure of the $\operatorname{ad}_{7}$ and $\operatorname{ad}_{6}$ loci of Schizosaccharomyces pombe. Proc. XI Int. Congr. Genet. 1, 7.
Hartman, P. E., Hartman, Z. \& Serman D. (1960). Complementation mapping by abortive transduction of histidine requiring Salmonella mutants. J. gen. Microbiol. 22, 354-368.
Holliday, R. (1964). A mechanism for gene conversion in fungi. Genet. Res. 5, 282-304.
Ishikawa, T. (1962). Genetic studies of ad-8 mutants in Neurospora crassa. II. Interallelic complementation of the ad-8 locus. Genetics, 47, 1755-1770.
Jessop, A. P. \& Catcheside, D. G. (1965). Interallelic recombination at the his-1 locus in Neurospora crassa and its genetic control. Heredity, Lond., 20, 237-256.
Kapuler, A. M. \& Bernstein, H. (1963). A molecular model for an enzyme based on a correlation between the genetic and complementation maps of the locus specifying the enzyme. J. molec. Biol. 6, 443-451.

Leupold, U. (1965). (Personal communication.)
Leupold, U. (1961). Intragene rekombination und allele komplementierung. Arch. Julius Klaus-Stift Vererb-Forsch. 36, 89-117.
Leupold, U. \& Gutz, H. (1964). Genetic fine structure in Schizosaccharomyces. Proc. XI Int. Congr. Genet. 2, 31-35.
Loper, J. C., Grabnar, M., Stahl, R. C., Hartman, Z. \& Hartman, P. E. (1964). Genes and proteins involved in histidine biosynthesis in Salmonella. Brookhaven Nat. Lab. Symp. 17, 15-52.
Manney, T. R. (1964). Tryptophan synthetase mutants of yeast: Action of a super-suppressor in relation to allelic mapping and complementation. Thesis, University of California.
Megnet, R. \& Giles, N. H. (1965). Allelic complementation at the adenylosuccinase locus in Schizosaccharomyces pombe. Genetics, 50, 967-971.
Monod, H., Wyman, J. \& Changeux, J. (1965). On the nature of allosteric transitions: A plausible model. J. molec. Biol. 12, 88-118.
Murray, N. E. (1960). Complementation and recombination between me-2 alleles in Neurospora crassa. Heredity, Lond., 15, 207-217.
Rice, J. (1963). The biochemical genetics of Neurospora crassa: Complementation at the arg-10 locus. Ph.D. Thesis, University of Leicester.
Schiesinger, M. J., Torriani, A. \& Levinthat, C. (1963). In vitro formation of enzymatically active hybrid proteins from $E$. coli alkaline phosphatase CRM's. Cold Spring Harb. Symp. quant. Biol. XXVIII. 539-542.
de Serres, F. J. (1963). Studies with purple adenine mutants in Neurospora crassa. V. Evidence for allelic complementation among ad-3b mutants. Genetics 48, 351-360.
de Serres, F. J. (1964). Mutagenesis and chromosome structure. J. cell. comp. Physiol. 64, Sup. 1, 33-42.
de Serres, F. J. (1965). (Personal communication.)

Suyama, Y. \& Bonner, D. M. (1962). Abstract: Complementation, mutation and enzyme alteration in the tryptophan synthetase system of Neurospora. Genetics, 47, 989.
Woodward, V. W. (1962). Complementation and recombination among pyr-3 hetero-alleles of Neurospora crassa. Proc. natn. Acad. Sci. U.S.A. 48, 348-356.
Woodward, D. O. (1958). Complementation at the ad-4 locus in Neurospora crassa. Proc. natn. Acad. Sci. U.S.A. 44, 1237-1244.
Woods, R. A. (1963). Genetic variation in yeast. D.Phil. Thesis, Oxford.
Woods, R. A. \& Bevan, E. A. (1966). Interallelic complementation at the ad-2 locus of Saccharomyces cerevisiae. Heredity, Lond., 21, 121-130.


[^0]:    * Present address: The National Institute for Medical Research, The Ridgeway, Mill Hill, London, N.W. 7.

