INTERPRETATIONS OF SOME LARGE NON-LINEAR COMPLEMENTATION MAPS

OLIVER J. GILLIE

National Institute for Medical Research, The Ridgeway, Mill Hill, London, N.W. 7, England

Received July 31, 1967

A general theory of interallelic complementation has been proposed by CRICK and ORGEL (1964), and this has been used by ISHIKAWA (1965) to interpret the relationship between the complementation map and the recombination map of the *ad-8* locus in *Neurospora crassa* and has also been used to interpret a more complex complementation map at the *leu-2* locus in *Neurospora crassa* (GROSS 1962; GILLIE 1966).

It seems that most complementation maps appear linear because they are constructed using rather small numbers of complementing mutants, and that random reduction of numbers of mutants in a large non-linear complementation map may result in the map becoming linear (GILLIE 1966). Since most loci may have complex complementation maps if sufficient data were available, it appears important that any theory of complementation should be able to account for such maps. Full confirmation of the CRICK-ORGEL theory may only be possible when chemical, physical and genetical methods are used on one protein. However, the general usefulness of the theory may be tested by seeing whether it is applicable to the more complicated maps in a consistent way. This paper shows how the theory may be used to interpret four more complementation maps.

A previous analysis of the *leu-2* map showed that complementation groups (or complementing mutants) could be considered to belong to one of two clusters, located at one or other pole of the circular complementation map (GILLIE 1966). Mutants located in groups at any one pole of the map showed no complementation amongst themselves but complemented with mutants in certain groups at the other pole of the map (Figure 1). Using these observations it was found to be possible to interpret the *leu-2* map as a topologically circular interface between two protein monomers in a dimer with a single axis of symmetry. Gross (1962) has interesting evidence to suggest that the *leu-2* and *leu-3* gene products interact to form a multimer, possibly a tetramer, of the type $\alpha\alpha\beta\beta$ where α is the gene product of *leu-2* and β the gene product of *leu-3*. This view does not conflict with the interpretation of the *leu-2* map given here; the *leu-2* gene product is referred to here as a dimer, as this is the simplest situation pertinent to our discussion. The poles of the *leu-2* map at which complementation groups were found to cluster were identified with the point at which the polypeptide chain intersected the axis of symmetry of the dimer; because, according to the CRICK-ORGEL theory, mutants involving defects at a particular intersection of the polypeptide chain

FIGURE 1.—Complementation map of the *leu-2* locus in *Neurospora crassa* redrawn from data of GROSS, 1962. One cluster of mutually non-complementing complementation groups is represented by plain lines, and the other cluster by dotted lines. A few groups (negative exceptions) which cannot be represented by a single line have been represented by branched lines which meet in the middle of the circle. In a previous, less detailed representation of this map (GILLIE 1966) these negative exceptions were represented by additional lines outside the circle; both representations are formally equivalent.

with the axis of symmetry should not show complementation amongst themselves, but may complement with mutants causing defects in other parts of the polypeptide chain, depending on the extent to which the various defects spread along the polypeptide chain. However, mutations which cause defects in the polypeptide chain away from the interface between the polypeptide monomers would be non-complementing, unless the defect spreads along the polypeptide chain to the interface, which would be expected to happen in certain cases. The defect might then be corrected, or partially corrected, if the part of the defect at the interface lies next to a normal portion of polypeptide chain in a hybrid mole-

cule. This interpretation of complementation data (CRICK and ORGEL 1964) considers the length of a complementation group on the map to be a function of the extent to which the polypeptide chain is defective. If, as seems likely, amino acid substitutions are the commonest type of mutational change which will complement, then the length of a complementation group will most often be a measure of the degree of uncoiling and associated steric effects which spread along the polypeptide chain (and across disulphide bonds) at the interface between the component monomers. A complementation group may, of course, appear as a branched line on the map in some cases, if the defect spreads *via* a disulphide bond.

On this interpretation the existence and relative position of axes of symmetry would be detected in sufficiently large complementation maps by the recognition of clusters of mutually non-complementing complementation groups. The difference between clusters of this kind, which are hierarchies of non-complementing complementation groups, and complementation groups themselves, which consist of one or more mutants with the same complementation response to all other mutants, must be appreciated. The differences between the complementation groups within one cluster are defined by virtue of their diverse interactions with the complementation groups within other clusters.

THE IDENTIFICATION OF CLUSTERS

In order to clarify what is meant by a cluster, definitions of group and cluster are given below, and methods used to identify clusters are examined.

A complementation *group* is a number of mutants which all have the same qualitative complementing properties. It should be noted that the number and constitution of such groups is therefore dependent on the percentage of wild type growth above and below which we arbitrarily choose to define positive and negative complementation responses.

A *cluster* of complementation groups (or of mutants) has the following properties: a) that no member of the cluster complements with any other member of the cluster; b) the cluster is formed from those groups which are most similar to each other with respect to the other groups with which they will complement; c) the cluster must consist of those groups with these properties which will make the cluster contain the maximum number of groups. A complementation group *may* belong to more than one cluster but a cluster is characterized by those groups which belong uniquely to it.

Thus, we may say that complementation matrices consisting of a large number, m , (at least 50 complementing mutants) of independent mutants at one locus may be shown to consist of:

- a) a number of groups, g , defined as above, such that $g \leq m$;
- b) a unique set of clusters, c , with the properties defined above, such that $c \ll g$ and as small as possible.

If the groups were arranged randomly on the map (or the matrix filled in randomly), it would be improbable that a unique set of clusters could be identified

in this way. *Each cluster then consists of similar, mutually non-complementing mutants and the clusters are chosen such that their total number is the least possible with any given set of mutants.*

This type of operational definition makes rather heavy weather of the identification of clustering in complementation maps. There are two simple ways of identifying clusters in practice apart from direct inspection.

The first method has been explained before (GILLIE 1966, p. 23 in the legend to Figure 6) and consists of plotting the central point of the line representing a complementation group on the map and ignoring the line itself. This has been done for the *ad-5/7* map in Figure 2b, and reference to this figure shows that five fairly well defined groups can be seen as a result. It is found that the groups identified by this method (which depends on position on the map only) do in practice fall within the clusters defined as above and identified independently by the second method.

For the second method of identifying clusters, proceed as follows using coloured pencils: locate a complementation group on the map which covers a small number of units, colour it in and colour with the same colour all those groups which overlap it and also overlap each other (i.e., mutually non-complementing groups). When there are no more groups left which overlap the original "small" group and each other, then start again with another separate "small" group, not overlapping the first, and colour that in another colour, and so on. It is essential to start each time with a "small" group, as small groups must be located near the centre of clusters and will lead to the choice of non-complementing mutants which are most similar and which will form a cluster consisting of the largest number of mutually non-complementing groups. When the map is entirely coloured in, it will be seen to consist of a number of clusters of groups each with its own colour. These clusters will have the properties given in the operational definition above. There will be some groups which are coloured with two or more colours because they overlap two or more clusters, and there may be occasional groups which seem to belong to no cluster. If groups were arranged on the map at random, however, most groups would become multi-coloured and clusters of uniquely coloured groups would not be found.

Some exceptional groups may be found which show a small number of positive complementation reactions with a few members of one cluster but clearly belong to no other cluster, and other single groups may be found which complement well with the majority of mutants in all other clusters but, since they are alone, can hardly be considered to be a cluster by themselves. The existence of exceptions of these types is anticipated by the CRICK-ORGEL model, and their significance is discussed below.

RESULTS AND DISCUSSION

A search for clusters of mutually non-complementing mutants in large complementation maps was successful (Table 1). These clusters were not quite so obvious as in the case of the *leu-2* map because in these maps 4–7 clusters per map were found. As there are from 10–20 times as many complementation groups as

TABLE 1

Analysis of clustering of complementation groups in some large, non-linear complementation maps

Organism	Locus	Number of complementing mutants in map	Number of complementation groups	Number of pairwise tests exceptional to postulated clusters	Number of pairwise tests exceptional to map only ¹	Total number of pairwise tests made	Percentage of tests exceptional to postulated clusters	Number of complementation groups in each cluster ²	Number of mutants in each cluster ²	Reference
<i>Neurospora crassa</i>	<i>leu-2</i>	78	42	2	1	0	3003	0.03	21,21	GROSS 1962, GILLIE 1966
<i>Saccharomyces cerevisiae</i>	<i>ad-5/7</i>	105	67	5	21	32	5460	0.38	a b c d e 3,8,7,21,12	DORFMAN 1964
	<i>ad-5/7</i>	59	49	3 (or 5)	4	2	1711	0.93 (or 0.23)	a b c 6,22,18 ⁴	COSTELLO and BEVAN 1964
	<i>ad-2</i>	78	61	7	3	11	2964	0.10	4,6,10, 7,6,7,2	WOODS and BEVAN 1966
<i>Salmonella typhimurium</i>	<i>his-B</i> ³	71	47	4	7	9	2485	0.28	12,9,5,4 16,18,12,9	LOPER <i>et al.</i> 1964

¹ Certain exceptional positive pairwise tests occur between mutants located in different clusters rather than between mutants located in the same cluster. This type of positive exception is considered exceptional to the map rather than to the cluster—see text for further discussion.

² Some complementing mutants or groups are not assigned to a particular cluster as they overlap more than one cluster. Such mutants are not exceptions, as to be an exception a mutant must occur between two clusters, overlapping neither of them. This, however, results in a discrepancy between the number of groups per cluster expressed this way and those found by plotting the mutants in the special way shown in Figure 2b which is unselective and includes groups which overlap more than one cluster. The data for the *ad-5/7* maps are labelled to correspond with the labelling in Table 2 and Figure 2.

³ The four clusters found in this map are roughly equivalent to the "four basic complementation classes" mentioned by LOPER *et al.* (1964).

⁴ Figures are given here only for three clusters since if five clusters are postulated the majority of groups then overlap two clusters—see text for full discussion.

clusters in these maps (Table 1), a considerable simplification is imposed by analysing maps into clusters. In Table 1 the number and frequency of pairwise tests exceptional to the postulated number of clusters is given, and in all cases is less than 1.0%. These exceptions may be accounted for in two ways: a) local compensation of particular amino-acid substitutions in the polypeptide chain may occur (CRICK and ORGEL 1964). Compensation of this kind may be thought of, with respect to two opposed lengths of polypeptide chain, as "bad plus bad makes good" rather than as "good plus bad makes good", as is understood to occur in normal complementation; b) in some cases the polypeptide chain may pass near to or across the axis of symmetry, but only a very small portion of it be actually at the interface between two monomers. Under these circumstances the chances of picking up a complementing mutation at that point may be rather small compared with the chances of picking up mutations in places where a large portion of the polypeptide chain exists at the interface on either side of the axis of symmetry. When such rare complementing mutants are picked up they might be recorded as exceptions rather than as new clusters. Exceptions of type a) would be expected to occur within clusters involving a few positive exceptional tests, whereas exceptions of type b) would be expected to involve a group which showed good complementation with most members of other clusters.

The number of positive exceptional pairwise tests (positive exceptions are pairwise tests which show complementation when non-complementation would be expected on the basis of all the other tests on the mutants concerned) is of the same order within and between clusters (Table 1), although the *leu-2* map is exceptional in having no positive exceptions between clusters. If, then, positive exceptions are interpreted as due to *compensation* of particular amino acid substitutions in the polypeptide chain, this compensation may be able to act over appreciable distances in the molecule or possibly involve the secondary steric consequences of the defects rather than just the defects themselves.

The number of complementation groups and mutants in each cluster is recorded in Table 1. It may be seen that a large proportion of all the complementation groups in each case occur in one or other cluster; however, some complementation groups straddle two or more clusters and then often cannot be assigned to one cluster rather than another. Such complementation groups are not exceptional to the clusters. To be exceptional a group would have to overlap no clusters rather than overlapping more than one, as these do. Groups of this kind, which straddle two or more clusters, may be understood to involve defects which have spread along the polypeptide chain such that the defect crosses the axis of symmetry in more than one place.

A special study has been made of the *ad-5/7* complementation maps in *Saccharomyces cerevisiae*, since two versions of this map have been made independently (DORFMAN 1964; COSTELLO and BEVAN 1964). The *ad-5/7* map of DORFMAN clearly shows five clusters, whether the clusters are identified by inspection (Figure 2a) or by plotting the mean position of groups on the map (Figure 2b). The *ad-5/7* map of COSTELLO and BEVAN may be considered as consisting of three or five clusters. Table 2 shows that the number of exceptional

TABLE 2

*Comparison of two interpretations of the two *ad-5/7* complementation maps constructed by DORFMAN (1964) and COSTELLO and BEVAN (1964)*

Hypothesis	Map	Number of positive exceptional pairwise tests					% of all pairwise tests	
		a	b	Clusters*				Total
				c	d	e		
Three clusters	DORFMAN (1964)	1	73	122	..	..	196	3.59
	COSTELLO and BEVAN (1964)	0	4	12	..	..	16	0.93
Five clusters	DORFMAN (1964)	2	0	2	10	7	21	0.38
	COSTELLO and BEVAN (1964)	0	0	0	0	2	4	0.23

* The letters assigned to the clusters may be used to identify the corresponding clusters in Figure 2 and Table 1. It must be noted that five of the ten exceptions noted under d and five of the seven exceptions noted under e of the DORFMAN data do not appear on the map but can only be seen with reference to DORFMAN's original material. This is entirely due to the difficulty in representing matrices by maps fully discussed in GILLIE (1966).

pairwise tests (considered cluster by cluster) is least for both maps when there are five clusters. However, the maps differ in that the *ad-5/7* (COSTELLO and BEVAN), when considered as five clusters, has many groups which straddle two clusters and rather few groups only occurring in any one of the clusters. As DORFMAN studied spontaneous and ethyl methane sulfonate (EMS) induced mutations, and COSTELLO and BEVAN ultraviolet and EMS induced mutations, and since differences were found in the complementation reactions of UV- and EMS-induced mutations by COSTELLO and BEVAN in their map, it is not surprising that the two maps are not precisely comparable.

We may consider how the five clusters found in the *ad-5/7* map may be arranged to intersect axes of symmetry such as to satisfy the CRICK-ORGEL hypothesis (CRICK and ORGEL 1964), as has been done previously for the *leu-2* data (GILLIE 1966). The arrangement of these clusters on the map is shown in a formal way in Figure 3A; in Figure 3B the map (which we equate with a topological representation of a polypeptide chain plus disulphide bridges) is deformed, as is permitted for topological figures, such that all the clusters now intersect a postulated axis of symmetry. Since all the clusters may be made to intersect a single axis of symmetry, the map may be considered to represent a dimer, although other, more complicated multimers could also of course be constructed to satisfy the facts. An example showing how the *ad-5/7* map may be arranged to form a multimer with two twofold and one fourfold axis of symmetry and sixteen subunits is given in Figure 4. This type of molecule might associate into larger formations due to the unpaired bonds (at c and d in Figure 4) at the edges, leading to the formation of a sheet. Other molecules with a basic repeating unit consisting of four subunits as shown in Figure 4B are possible. The *his-B* map of Salmonella (Table 1), which consists of four clusters on a circular map, may be accounted for in a similar way.

The arrangement of clusters postulated in the *ad-2* map (WOODS and BEVAN) is illustrated in a formal manner in Figure 3C. In Figure 3D, the map is deformed to bring all the clusters to intersect an imaginary axis of symmetry generated by

FIGURE 2.—a. Complementation map of the *ad-5/7* locus in *Saccharomyces cerevisiae* redrawn from data of DORFMAN, 1964. Five distinct clusters are represented here; two clusters opposite each other in the circular portion of the map, by dashed lines; two other clusters, also opposite each other in the circular portion of the map, by dotted lines; and one cluster in the straight portion of the map, by dot-dashed lines. Groups which overlap more than one cluster, and so cannot be assigned to a particular cluster, are represented by plain lines. The clusters are labelled

FIGURE 3.—A. Simple formal representation of the *ad-5/7* complementation map. The thin line represents the general form of the map identified with the polypeptide chain and the thickened portion of the line indicates the position of non-complementing clusters of complementation groups.

B. The *ad-5/7* map has been deformed to align all clusters such that they intersect a postulated axis of symmetry (shown by arrows).

two protein monomers. In order to do this we have to imagine, quite reasonably, that parts of the polypeptide chain (dotted lines) go behind the interface between the two monomers. It is disappointing that so complex a map as *ad-2* seems to impose little restriction on the type of structural interpretation which can be made.

Using the complementation maps at present available, it is possible to predict the number (or, more cautiously, the minimum number) of times the polypeptide chain intersects the axis of symmetry of the molecule. This information, together with the general form of the complementation map, does not seem to be

a, b, c, d, e, to correspond to the clusters in Figure 2*b* and in Tables 1 and 2.

b. Special plot to show the clustered distribution of *ad-5/7* complementation groups. The central point (or mean position) of each group is represented on the map by a square. The central point of a group is defined as that point on the line representing the group 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 plot shows that the mutants may be divided into five clusters (labelled a, b, c, d, e) on this basis. The labelling corresponds with that of Figure 2*a* and Tables 1 and 2.

C. Simple formal representation of the *ad-2* complementation map with conventions as for the *ad-5-7* drawings. The *ad-2* map has a more complicated general form and seven clusters.

D. The *ad-2* map has been deformed to align all clusters such that they intersect a postulated axis of symmetry. The dotted line indicates parts of the polypeptide chain which go behind the interface between the two monomers.

Corresponding clusters in the "undeformed" and "deformed" maps of both the *ad-5/7* and *ad-2* maps are labelled with lower case letters.

sufficient to give us more than a very general two-dimensional picture of subunit interaction. However, studies of the relationship between complementation and recombination maps should yield further information about the secondary structure of the protein involved. Although co-linearity of the complementation map and recombination map could not be expected on the CRICK-ORGEL model, it would be expected that mutants occurring in clusters on the complementation map would often occur in clusters on the recombination map, since complementing mutants might be expected to be found in clusters at points on the recombination map, equivalent to parts of the polypeptide chain which cross the axis of symmetry. There is a shortage of data in which both complementation and recombination maps have been made using large numbers of mutants at one locus. The best of such data is for the *ad-8* locus in *N. crassa* (ISHIKAWA 1965) which shows that the mutants found in the same position on the complementation map

FIGURE 4.—The *ad-5/7* map is arranged to show how it might be generated by a large multimer with three axis of symmetry. The multimer consists of sixteen subunits, drawn to resemble the *ad-5/7* map, as a circle with a tail at right angles to it. In A the multimer is seen in plan showing eight subunits each of which is matched by an identical subunit below. In B, one edge of the molecule is seen from the side, looking from point B (marked by arrow in 4A), showing the circular portion of the map as a straight line with the tail at right angles to it. In C the molecule is seen again from the side, looking from point C (marked by arrow in 4A), showing once more the circular portion of the map as a straight line with the tail at right angles to it.

The points of the molecule which would generate clusters of complementation groups are labelled with lower case letters to correspond with clusters in Figure 3A. Lower case letters *in brackets* indicate portions of the molecule which are hidden from a particular aspect by an overlying heterologous portion of the amino-acid chain. Portions a, b, e of the molecule always lie adjacent to homologous amino acid chains, whereas heterologous portions d and c lie adjacent to each other, but also pair homologously in some cases with other d, c regions such that two d and two c portions lie together.

tend to be found in the same position on the recombination map, although the maps are definitely not co-linear. Unfortunately, although ISHIKAWA has used a large number of mutants located at many different positions on the recombination map, the resolution obtained in the complementation map is, for unknown reasons, low compared with that obtained for the loci discussed earlier in this

paper. Nevertheless there are three "presumptive clusters" in the *ad-8* complementation map which, as ISHIKAWA has carefully shown, are related to the genetic map in a way which is entirely consistent with the CRICK-ORGEL model. Correlated clustering of mutants on complementation and recombination maps is also found at the *ad-6* and *ad-7* loci of *Schizosaccharomyces pombe* (LEUPOLD and GUTZ 1964; RAMIREZ, FRIS and LEUPOLD 1963), at the *pan-2* locus of *Neurospora crassa* (CASE and GILES 1960) and at genes 10 and 37 of bacteriophage T4D (BERNSTEIN, EDGAR and DENHARDT 1965), although strict co-linearity between the complementation and recombination maps is found at none of these loci.

As made clear by CRICK and ORGEL, any detailed verification of these types of predictions must await X-ray analysis of suitable proteins. However, the investigation of complementation reactions themselves may be taken much further by using larger samples of mutants to test the consistency of various types of exceptions, and by using a variety of mutagens in order to study both the types of map obtained and the specificity and relative "damage" attributable to various mutagens. Future investigations may also be able to make use of the quantitative data available about the intensity of complementation interactions which have not been considered in this or in the earlier publication (GILLIE 1966). Finally, statistical studies of the evolution of proteins (FITCH 1966) should reveal whether mutations in certain parts of the molecule (for example at the subunit interfaces at or away from the axes of symmetry) may have occurred more frequently in the course of evolution of the molecule. FINCHAM (1966) has considered the general evolutionary significance of complementation; if, however, the CRICK-ORGEL theory of complementation proves to be correct, a more detailed consideration of the evolutionary implications of this theory will be necessary.

I should like to thank Dr. R. HOLLIDAY and Dr. H. J. ROGERS for reading the manuscript and for general encouragement, and Mr. R. PETO for valuable discussion.

SUMMARY

Five large and complex complementation maps are shown to consist of a number of clusters of non-complementing complementation groups. This clustering is interpreted and discussed in terms of the CRICK-ORGEL theory of complementation.

LITERATURE CITED

- BERNSTEIN, H., R. S. EDGAR, and G. H. DENHARDT, 1965 Intragenic complementation among temperature sensitive mutants of bacteriophage T4D. *Genetics* **51**: 987-1002.
- CASE, M. E., and N. H. GILES, 1960 Comparative complementation and genetic maps of the *pan-2* locus in *Neurospora crassa*. *Proc. Natl. Acad. Sci. U.S.A.* **46**: 659-676.
- COSTELLO, W. P., and E. A. BEVAN, 1964 Complementation between *ad-5/7* alleles in yeast. *Genetics* **50**: 1219-1230.
- CRICK, F. H. C., and L. E. ORGEL, 1964 The theory of interallelic complementation. *J. Mol. Biol.* **8**: 161-165.
- DOEFMAN, B., 1964 Allelic complementation at the *ad-5/7* locus in yeast. *Genetics* **50**: 1231-1243.

- FINCHAM, J. R. S., 1966 *Genetic Complementation*. W. A. Benjamin Inc., New York and Amsterdam.
- FITCH, W. M., 1966 An improved method of testing for evolutionary homology. *J. Mol. Biol.* **16**: 9-16.
- GILLIE, O. J., 1966 The interpretation of complementation data. *Genet. Res.* **8**: 9-31.
- GROSS, S. R., 1962 On the mechanism of complementation at the *leu-2* locus of *Neurospora*. *Proc. Natl. Acad. Sci. U.S.* **48**: 922-930.
- ISHIKAWA, T., 1965 A molecular model for an enzyme based on the genetic and complementation analyses at the *ad-8* locus in *Neurospora*. *J. Mol. Biol.* **13**: 586-591.
- LEUPOLD, U., and H. GUTZ, 1964 Genetic fine structure in *Schizosaccharomyces*. *Proc. 11th Intern. Congr. Genet.* **2**: 31-35.
- LOPER, J. C., M. GRABNAR, R. C. STAHL, Z. HARTMAN, and P. E. HARTMAN, 1964 Genes and proteins involved in histidine synthesis in *Salmonella*. *Brookhaven Symp. Biol.* **17**: 15-52.
- RAMIREZ, C., J. FRIS, and U. LEUPOLD, 1963 Allelic recombination and complementation in adenine-requiring mutants of *Schizosaccharomyces pombe*. *Proc. 11th Intern. Congr. Genet.* **1**: 7.
- WOODS, R. A., and E. A. BEVAN, 1966 Interallelic complementation at the *ad-2* locus of *Saccharomyces cerevisiae*. *Heredity* **21**: 121-130.