PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The Structure of <latex>α</latex>-Chymotrypsin

J. J. Birktoft, D. M. Blow, R. Henderson and T. A. Steitz

Phil. Trans. R. Soc. Lond. B 1970 **257**, 67-76
doi: 10.1098/rstb.1970.0009

Phil. Trans. Roy. Soc. Lond. B **257**, 67–76 (1970)
Printed in Great Britain

I. SERINE PROTEINASES

The structure of α-chymotrypsin

By J. J. Birktoft, D. M. Blow, R. Henderson and T. A. Steitz
Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge

[Plates 51 to 53]

A newly calculated electron density map has allowed a more detailed description of the molecular structure to be given. The structure can be described in detail in relation to the probable existence of hydrogen bonds and the conformations of side chains. The discovery of a new buried acid group which is part of a hydrogen bonding system involving the active serine has indicated how this serine can become a powerful nucleophile, by means of a 'charge relay system'. Diffraction studies of the binding of various substituents, coupled with accurate model building, have defined the catalytic binding site.

Introduction

Bovine α-chymotrypsin is one of the most thoroughly studied of all enzymes. Owing to the simple procedure for purification of the zymogen (Kunitz & Northrop 1935), the relative stability of this form of the enzyme, and the abundant supply of the starting materials at large slaughterhouses, it has been cheaply available for many years. The amino acid sequence was among the first to be determined for any enzyme (Hartley 1964), and has been confirmed by an independent study (Meloun, Kluh, Kestka, Morávek, Prusík, Vaněček, Keil & Šorm 1966). The chemical steps leading to the formation of the α form of the enzyme from the zymogen were determined in detail (Rovery, Poilroux, Yoshida & Desnuelle 1957). The enzyme is inactivated by stoichiometric quantities of organophosphorous compounds (Jansen, Nutting & Balls 1949), resulting in the acylation of a specific serine residue, Ser-195 (Shaffer, May & Summerson 1953). Other chemical modification studies showed that photo-oxidation or alkylation of a specific histidine residue (His-57) completely destroys the enzyme activity (Ray & Koshland 1960; Ong, Shaw & Schoellman 1964); as does the acetylation of the α-amino group (of Ile-16) which is formed at the activation step (Oppenheimer, Labouesse & Hess 1966). Oxidation of a reactive methionine residue (Met-192) affects binding of the substrate, but not the rate of catalysis (Koshland, Strumeyer & Ray 1962). The residue Tyr-146, which is an exposed and reactive C-terminal residue in the α form of the enzyme, is not essential for activity (Gladner & Neurath 1954).

Trypsin and chymotrypsin were found to have very similar catalytic properties, but very different specificities, chymotrypsin favouring large hydrophobic side chains and trypsin favouring basic side chains (Neurath & Hartley 1959). Niemann and his co-workers, by extensive classical kinetic studies, attempted to outline the complete binding site of the enzyme (see especially Hein & Niemann 1962). The binding specificity has been more closely defined by Cohen and his colleagues (see Cohen, Milovanović, Schultz & Weinstein 1969). One result of these studies was the discovery of a series of 'locked' substrate molecules in which rotation at the α- and β-carbon atoms of the side chain is prevented, but which are rapidly hydrolysed by the enzyme, thus defining a correct conformation of the substrate molecule. In equilibrium binding studies, Johnson & Knowles (1966) found evidence that the binding site contained a group with pK_a of 7, carrying a negative charge at high pH.

It was shown that the enzymic reaction proceeded in several steps (Gutfreund & Sturtevant 1956; Bernhard & Gutfreund, this volume, p. 105), and that at least in several cases amenable to study, an acyl enzyme intermediate was formed (see Kézdy, Clement & Bender 1964). Optical and thermodynamic evidence was found for the existence of several conformational states of the enzyme (Biltonen, Lumry, Madison & Parker 1965; Brandts 1964). Much experimental and theoretical effort has been devoted to the determination of the chemical steps of the enzymic process. Some of this work has been reviewed by Bender & Kézdy (1964), Cunningham (1965) and Bruice & Benković (1966).

This intensive study, outlined very crudely above, made the chymotrypsins an important subject for crystallographic study. Despite crystallographic difficulties noted by Bernal, Fankuchen & Perutz (1938), workers in our laboratory chose to study the α form of the enzyme (Blow, Rossmann & Jeffery 1964; Sigler, Jeffery, Matthews & Blow 1966). Low resolution electron density maps of chymotrypsinogen (Kraut, Sieker, High & Freer 1962), δ-chymotrypsin (Kraut, Wright, Kellermann & Freer 1967) and γ-chymotrypsin (Matthews, Cohen, Silverton, Braxton & Davies 1968) have been obtained elsewhere.

During 1967 a high resolution electron density map of α-chymotrypsin was obtained whose general features could be interpreted (Matthews, Sigler, Henderson & Blow 1967; Sigler, Blow, Matthews & Henderson 1968). A number of questions remained unsolved about the detailed conformation of the polypeptide chains, and about the hydrogen bonding scheme, which clearly needed a more accurate electron density map. A map has been calculated recently which shows very much more detail, which can be interpreted in terms of a more accurate model, and which leads to some new conclusions. The improvement in the map is due to improved interpretation of the data which already existed, and further improvement in the model is due to more accurate techniques of model building, following a technique recently introduced by Richards (1968).*

In this paper, the conformation of the molecule is first described in relation to the hydrogen bonding scheme; secondly, the conformation of some of the residues is related to a possible catalytic mechanism; and finally the substrate binding properties of the crystalline enzyme are discussed.

2. The conformation of the molecule

The crucial difference between the new electron density map and the one calculated previously is that the enhancement of electron density of the carbonyl groups of the peptide bonds is clearly visible in almost every case. From these the orientation of the planar peptide unit can be inferred, and this leads to a reasonable degree of certainty about the existence of main chain hydrogen bonds. In several parts of the earlier map, the density faded out to the point where the density of the polypeptide chain could not be followed. With a small exception this uncertainty has now been removed. The side chains are clear, and where an aromatic side chain exists, the side chain is flattened. There is a tendency towards dimpling of the electron density at the centres of six membered rings. Valine, leucine, isoleucine and threonine residues generally appear clearly branched. The electron density of the longer side chains, especially where lysine and arginine extend into the solvent, is less clear.

The interpretation of the electron density map is based on the determination of the sequence of the 245 amino acids which compose chymotrypsinogen by Hartley (1964; Hartley & Kauffmann 1966), and by Keil & Šorm's group in Prague (Meloun *et al.* 1966). In α-chymotrypsin,

* Preliminary coordinates taken from this model have been published by Birktoft, Matthews & Blow (1969).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

peptide bond cleavages due to activation and autolysis have removed two dipeptides, leaving three polypeptide chains joined by disulphide bridges (Rovery *et al.* 1957). The A chain is a short chain of 13 residues, while the B and C chains contain 131 and 97 residues respectively. Amino acid residues are numbered according to their position in the chymotrypsinogen sequence.

Figure 1, plate 51, is a stereoscopic view of a space-filling model of the molecule. The parts of the molecule with which we had difficulty in the earlier map are at the top and bottom as viewed here. At the top there is still a small region Gly-74-Ser-75-Ser-76-Ser-77 which is very hard to interpret. The difficult region at the bottom of the molecule has now been satisfactorily interpreted. It turns out to have a short section of α-helical conformation, which had not previously been identified, from residues 164 to 170. A third region of difficulty was associated with Phe-39. This residue appears in our latest map only as far as the β-carbon, with very little density for the rest of the side chain, which is perhaps in free rotation. This seems unusual for a surface aromatic group, but there is no reason to doubt that the amino acid sequences of Hartley and of the Prague group are correct at this point. Apart from the density of this side chain, the interpretation of this part of the molecule is perfectly satisfactory. A final point of difficulty which remains is the conformation of the last three residues of the A chain (residues 11, 12 and 13). This is probably because they take up different conformations in the two crystallographically distinct views of the molecule which this particular crystal structure gives us (Blow, Rossmann & Jeffery 1964; Sigler *et al.* 1968).

Figure 2 gives a summary of the hydrogen bonds between atoms in the main polypeptide chain, by showing each as an arrow from the amido to the carbonyl group. Uncertain hydrogen bonds, as well as hydrogen bonds involving side chain atoms, have been omitted. Adjacent chains in the diagram are always antiparallel, and it is evident that the hydrogen bonding between adjacent chains is predominantly of the antiparallel pleated sheet type. However, to say that the structure consisted to any large extent of antiparallel chain pleated sheet would be a gross oversimplification. The form in which this diagram is presented emphasizes a curious relation first noted in a personal communication by Dr B. W. Matthews. After the A chain, which is the very top line, the next six lines of the diagram are drawn according to a particular pattern. Lower down the diagram the same pattern is repeated again. Within each of these patterns, there is a further series of antiparallel interactions between the last and the first line. On this basis, the overall structure of the molecule could be described as two cylinders of antiparallel chain pleated sheet. While it is true that this relation describes the topological properties of the main chain hydrogen bonding pattern, it is only a very partial description of the molecule, and the two cylinders described are not immediately obvious on looking at a complete model of the structure. There are only very short regions where the full, classical pleated sheet structure of Pauling is found.

In figure 2 helical regions have been indicated by a zigzag line. The C-terminal α-helix ends with a hydrogen bond which is not of the normal α-helical type. Possibly it belongs to the α_{II}-helical type (Némethy, Phillips, Leach & Scheraga 1967) but more probably to the 3_{10} type. A similar termination is observed to the other turn of α-helix, but this is confused by further 3_{10} and α-type bonds. Outside of these helical regions, probably only one hydrogen bond exists which is from a residue to its fourth neighbour backwards along the chain. On the other hand, it will be noted that hydrogen bonds to the third nearest neighbour back along the chain are very common. Recently, Venkatachalam (1968) has discussed the conditions for formation of such hydrogen bonds. The only conformations which permit formation of such hydrogen bonds

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

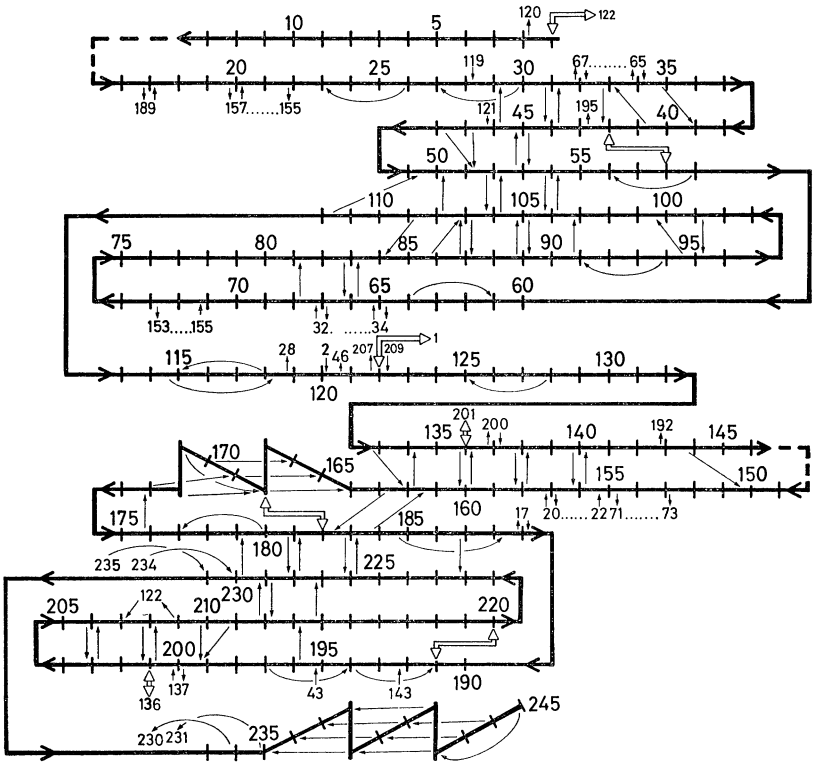


FIGURE 2. Summary of hydrogen bonds between peptide groups. Only the hydrogen bonds which are reasonably certain have been included: several more may exist, and there are many others between peptide groups and side chain atoms, which are not shown. The arrows are drawn from amido to carbonyl group. The double headed, open arrows represent disulphide bridges, and the dotted lines represent residues present in the zymogen which are removed from α-chymotrypsin by autolysis.

DESCRIPTION OF PLATE 51*

FIGURE 1. Space-filling model of α-chymotrypsin. The model has been constructed of CPK components (Ealing Scientific). Oxygen atoms (which are red in the original) can be distinguished from the black carbon atoms by their lighter colour and by the slots in the surface of those oxygen atoms which would be capable of accepting hydrogen bonds. Those hydrogen atoms (white) capable of being donor protons in hydrogen bonds are equipped with hooks. Nitrogen atoms (blue) appear as a medium grey. The two large pale coloured atoms prominent in the centre of this view are sulphur (yellow). That on the left belongs to Cys-42, and that on the right to Met-192. The residue number written on some of the atoms may be legible. The active centre region, which lies immediately below these two residues, can be seen in more detail in figure 9.

FIGURE 3. Detail of the end of the serine loop, residues 203–206, constructed from skeletal model parts (Cambridge Repetition Engineers, Greens Road, Cambridge). Starting from the lysine residue whose side chain hangs down to the bottom edge of the illustration, the sequence ...Lys-203-Asn-204-Gly-205-Ala-206-... forms the end of the loop. A white disk is attached to each α-carbon atom; the carbonyl groups of the peptide bond can be recognized by the longer stem; hydrogen bonds are represented by plastic tubing. The conformation represented here is called type II (Venkatachalam 1968) and is only possible when glycine is in position 3.

FIGURE 4. Electron density map of the active centre region of tosyl-α-chymotrypsin, with an interpretation superimposed. This is a view of stacked Perspex sheets for sections $x = \frac{17}{64}$ to $x = \frac{24}{64}$. Some residues clearly visible are identified by number, namely Ile-16, His-57, Asp-102 (only the density of the carboxyl group in the sections), Arg-145, Tyr-146, Cys-191-220, Met-192, Asp-194, tosyl-Ser-195, Gly-196. W indicates a water molecule which mediates hydrogen bonding between His-57 and S, the sulphonyl group of tosyl-Ser-195. This view is parallel to a local twofold axis which brings the active centres of two molecules close together. This twofold axis passes close to the bottom of the Perspex sheets, in the centre. Contours at the left bottom show part of the density for Met-192, Cys-191-220, and Ala-195; to the left of W is part of the density of Tyr-146 and the carboxyl terminus of the B chain.

* Colour transparencies corresponding to figures 1, 3, 4, 5, 7, 8, 9, and to several other figures in this volume, in either standard or stereoscopic mounts, may be purchased from B. K. Harvey, 7 Perse Way, Cambridge.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

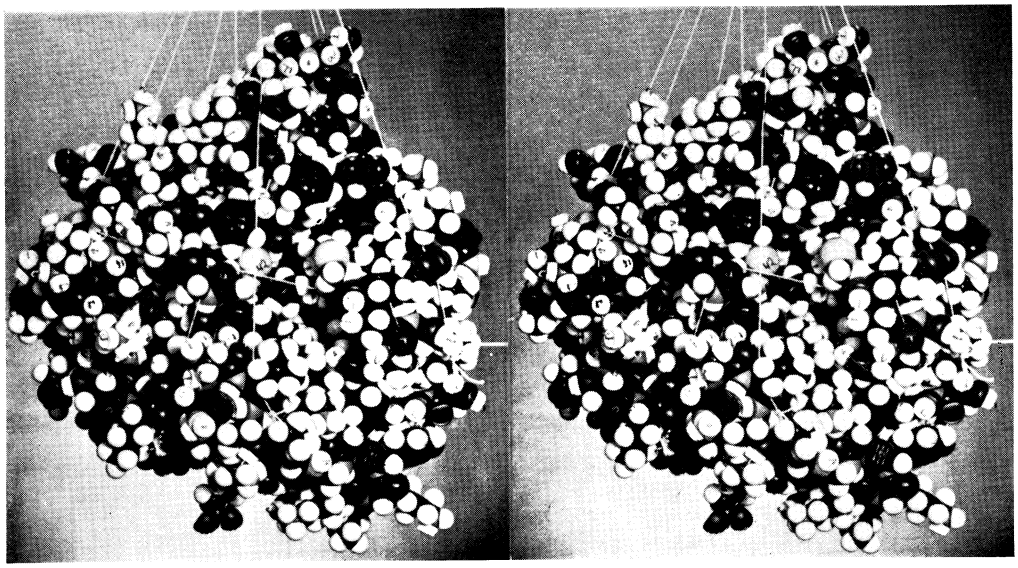


FIGURE 1

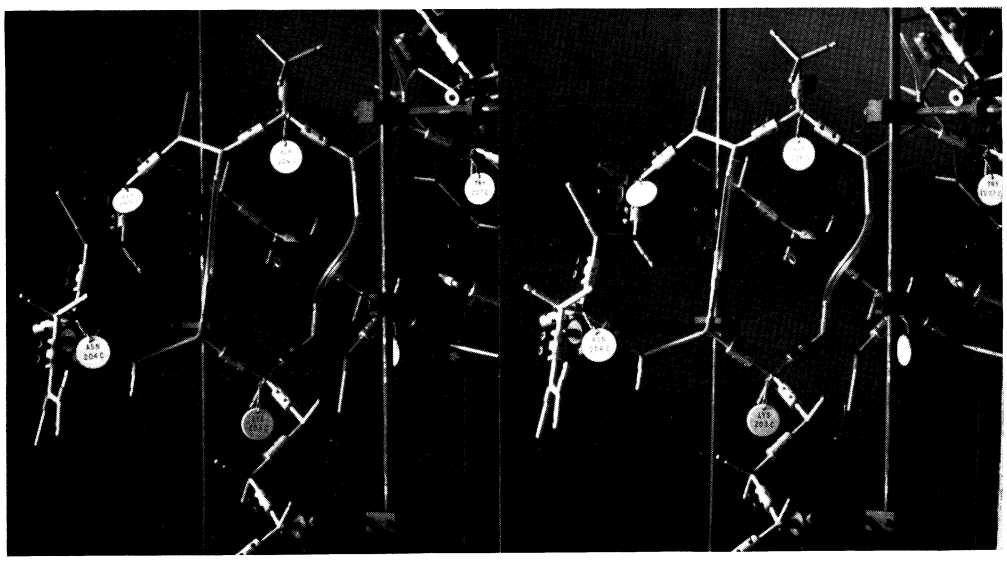


FIGURE 3

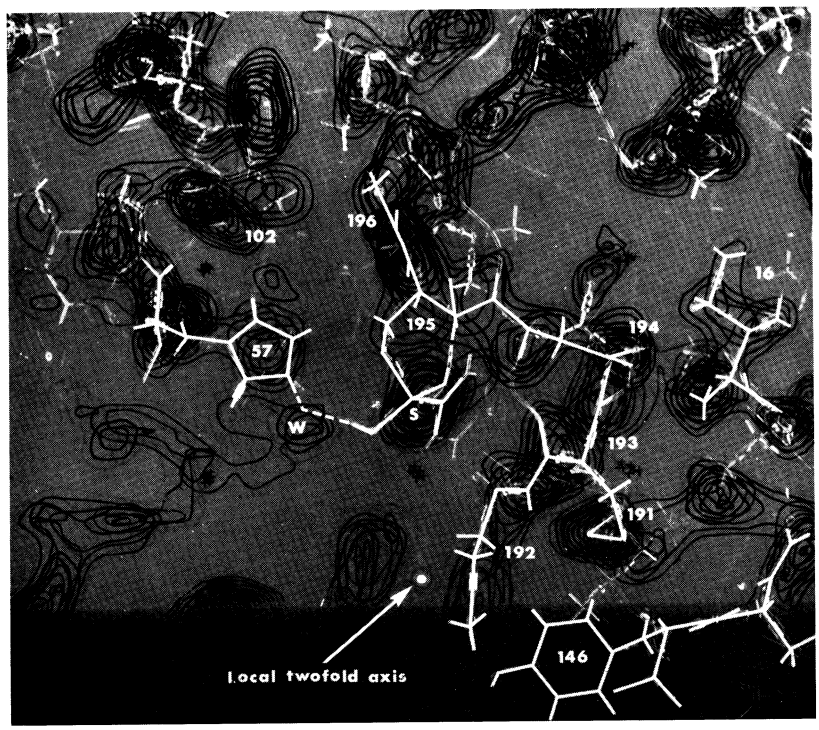
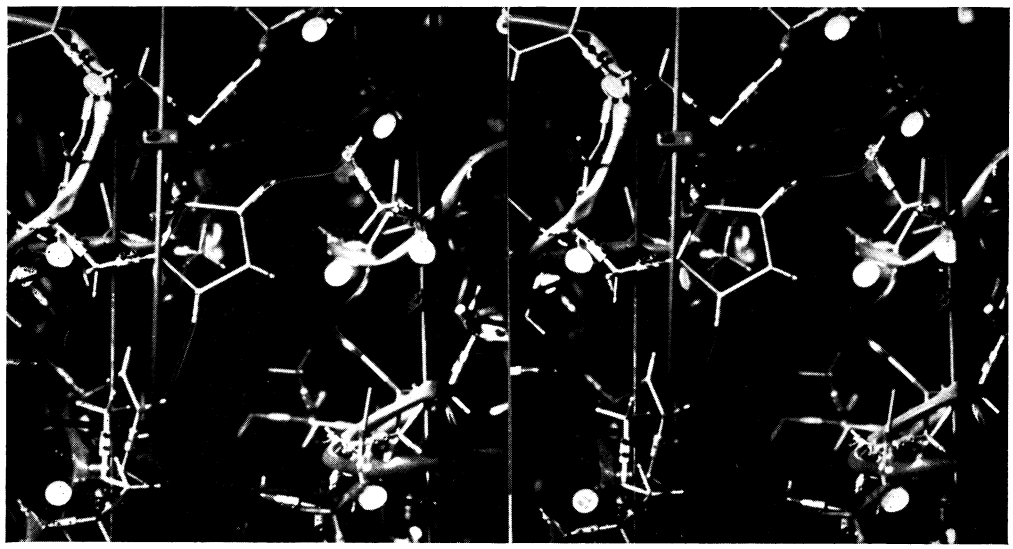


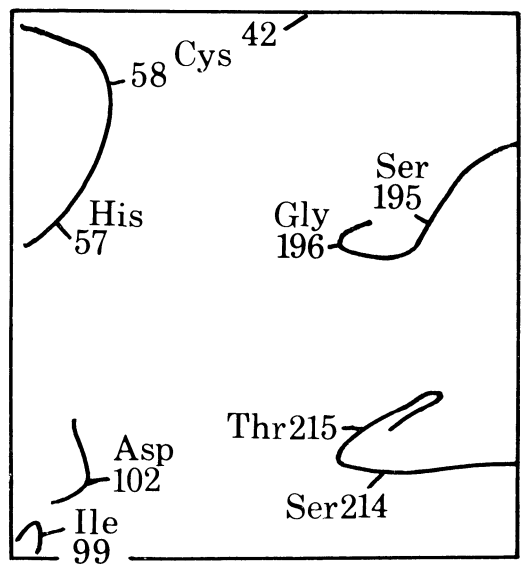
FIGURE 4 (for legends see facing page)

(*Facing p.* 70)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES



(a) (b)



(c)

FIGURE 5

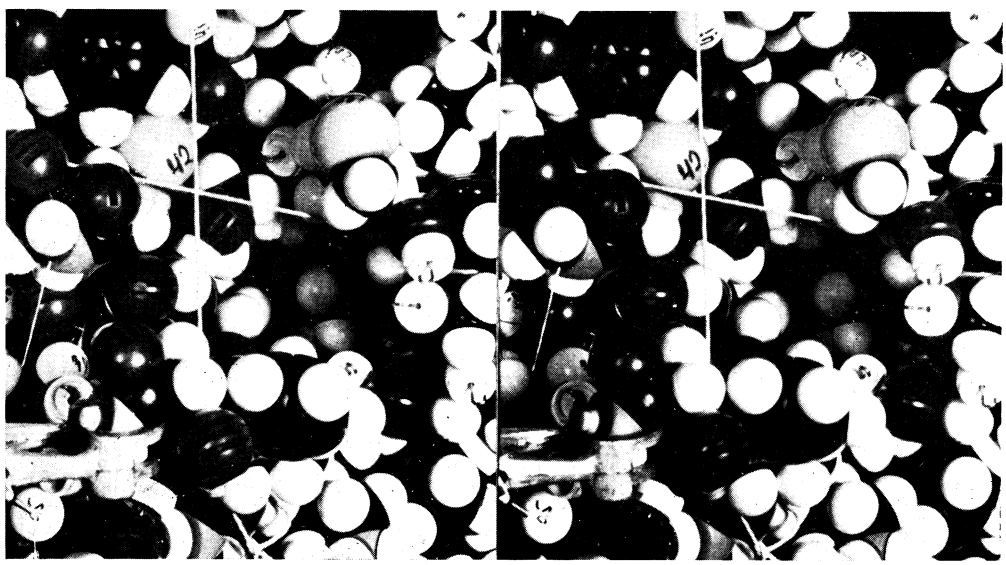
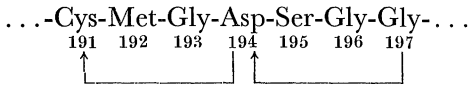


FIGURE 9 (for legends see facing page)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

result in the reversal of the direction of the third peptide bond, in relation to the first, resulting in a 'hairpin bend' in the polypeptide chain. An example is shown in figure 3, plate 51.

Venkatachalam has described two types of conformation which permit this hydrogen bonding system. In type I, each of the three peptide bonds are in a roughly similar relation to each other, so that the conformation of the three residues is approximately helical. The 3_{10} helix is a special case of this type of conformation. In type II the orientation of the second of the three peptide bonds is reversed. For L-amino acids this conformation is permitted only if glycine is in position 3. Of the 17 cases where such a type of hydrogen bond pattern has been observed, 14 are of type I. The other three are of type II, and glycine is in position 3 in each case. Two of these cases are associated with the active centre sequence



This hydrogen bond pattern provides rigid structure in the peptide backbone around the active site, helping to hold Ser-195 and Asp-194 in an exact orientation. Glycine residues 193 and 196 are invariant in all known sequences of vertebrate serine proteinases (Hartley, this volume, p. 77). It follows that serine enzymes which do not have glycine in these positions cannot have the same polypeptide conformation around their active serine.

The whole structure is made up of a number of loops in which the polypeptide chain turns back on itself. A number of antiparallel-type hydrogen bonds are used in stabilizing each loop. Some of the loops are further stabilized by the existence of disulphide bridges. The histidine loop (residues 42–58), containing the active centre histidine, the methionine loop (residues 165–182), which contains one of the two methionines in the molecule, and the serine loop (residues 190–220), which contains the active serine, come in to this category. Other prominent loops are not stabilized by disulphide bridges. One (residues 84–110) is called the aspartate loop because it contains Asp-102, which will be referred to later; another (residues 133–164) is called the autolysis loop because it contains the dipeptide 147–148 in the zymogen, which is removed in the α form of the enzyme. A prominent loop across the top of the molecule includes residues 66–83.

We have recently put together a space filling model of the structure. This was done by first assembling the atoms of each individual loop. Some of these loops exhibit the tendency of the

Description of plate 52

Figure 5. The charge relay system in α-chymotrypsin. Plastic tubing has been attached to the main polypeptide chains, to help emphasize them. The prominent residues can be identified by reference to the sketch. This photograph, taken before residue 102 was confirmed as aspartic acid, shows it as asparagine. The carbonyl group of Ser-214 which may be involved in a hydrogen bond to the amido group of a substrate, is also clearly visible in this photograph.

Figure 9. Enlarged view of the space filling model of the active centre region of α-chymotrypsin. This view is taken from almost the same direction as figure 1, and the sulphur atoms of Cys-42 and Met-192 are very prominent. The —CH_2OH side chain of Ser-195 can be seen immediately to the right of and below Cys-42, the hydroxyl proton being involved in a hydrogen bond to His-57, as in figure 6*b*. The 'tosyl hole' is to the immediate right of these residues, and below them an *N*-formyl-tryptophanyl group is held in a clamp in the correct orientation to enter the hole. The —CH_2— group of Ser-189 is clearly visible at the back of the hole, and a hydrogen atom attached to the α-carbon of Gly-216 is seen directly above the centre of the six-membered ring of the tryptophanyl residue. This hydrogen atom would be replaced by an isopropyl group if residue 216 were valine, as it is in elastase. The importance of these residues is considered elsewhere in this Discussion by Hartley (p. 77) and by Shotton & Watson (p. 111).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

two antiparallel chains to spiral round each other. This seems to indicate the existence of similar packing forces to those which cause the large area of twisted pleated sheet in carboxypeptidase (Reeke, Hartsuck, Ludwig, Quiocho, Steitz & Lipscomb 1967), causing a distortion of the planar system of the pleated sheet.

3. The stereochemistry of the active centre

Figure 4, plate 51, shows the electron density map of tosyl-α-chymotrypsin in the region of the residues known to be implicated in the active centre, with the interpretation superimposed. A small but significant change has been made since our preliminary interpretation (Matthews *et al.* 1967; Sigler *et al.* 1968), which consists of turning the side chain of His-57 over in such a way that the δ-nitrogen atom is pointing internally rather than to the solvent. This map is an electron density map of an inhibited derivative of chymotrypsin in which a toluene sulphonyl group is acylating the serine of the active centre. One of the oxygen atoms of the sulphonyl group is involved in a hydrogen bond with the amide group of Ser-195. The other oxygen atom, which we originally thought was directly hydrogen-bonded to a nitrogen of His-57 seems to be slightly too far away for that to occur and we now believe there to be a pair of hydrogen bonds, through a water molecule, joining these atoms. The δ-nitrogen of His-57 forms a hydrogen bond to residue 102, which in the sequence as originally determined by Hartley (1964) and Meloun *et al.* (1966) appeared as asparagine. Asp-194 occupies an interior position with its carboxyl group in close proximity to another charged group, the α-amino group of Ile-16. One of the most important deductions made from the earlier electron density map was that the existence of this ion pair was a result of the activation step in which a peptide bond is cleaved by trypsin, so as to create an α-amino group at Ile-16. It was presumed that the existence of this ion pair buried deeply within the molecule exerted a force which held the active centre of the enzyme in the correct stereochemical conformation (Sigler *et al.* 1968; Hess, McConn, Ku & McConkey, this volume, p. 89). The more accurate model now available shows that the carboxyl group could be brought into contact with the solvent simply by rotation about the α-β bond of Asp-194, and that in this position the carboxyl group could readily form a hydrogen bond to Ser-195. Whether this speculation has any factual basis will only be known when the structure of chymotrypsinogen becomes available.

A comparison of the sulphonylated and native enzymes showed that sulphonylation causes a movement of the γ-oxygen of Ser-195 together with a movement of the His-57 side chain, and of Met-192 (Sigler *et al.* 1968). The movement of the serine oxygen is consistent with a simple rotation about the α-β bond. In the native enzyme the serine and the histidine are slightly closer, so that a direct hydrogen bond can now be formed between the γ-oxygen of Ser-195 and the ε-nitrogen of His-57. The environment of His-57 in the native form is shown in figure 5, plate 52. There is seen to be a hydrogen bond from the δ-nitrogen to residue 102. This observation focused interest on residue 102, which had already invited attention because, although it had been identified as asparagine in chymotrypsinogen A, the sequences of several homologous enzymes, including chymotrypsinogen B, were found to have aspartic acid at this point (Smillie, Furka, Nagabhushan, Stevenson & Parks 1968; Hartley, this volume, p. 77). Because of this Hartley has re-examined this sequence, and was able to inform us that this residue was also aspartic acid in chymotrypsinogen A (Blow *et al.* 1968). Asp-102 is also in an internal position. There is a group of three side chains below it: Ile-99, Ser-214 and Trp-215 which shield it from

solvent, and the small space between them would be too small to allow access of a hydronium ion, without some disruption of the tertiary structure.

At pH 4, the pH of the crystallographic study, we may expect that the aspartate group will carry a negative charge and the histidine side chain a positive charge, thus forming an internal ion pair (figure 6*a*). We cannot strictly speak of the p*K* of the histidine residue by itself any more; instead we have to speak of the p*K* of this hydrogen bonded system. When a proton is removed from it, the histidine will be uncharged, and the overall hydrogen bonded system comprising

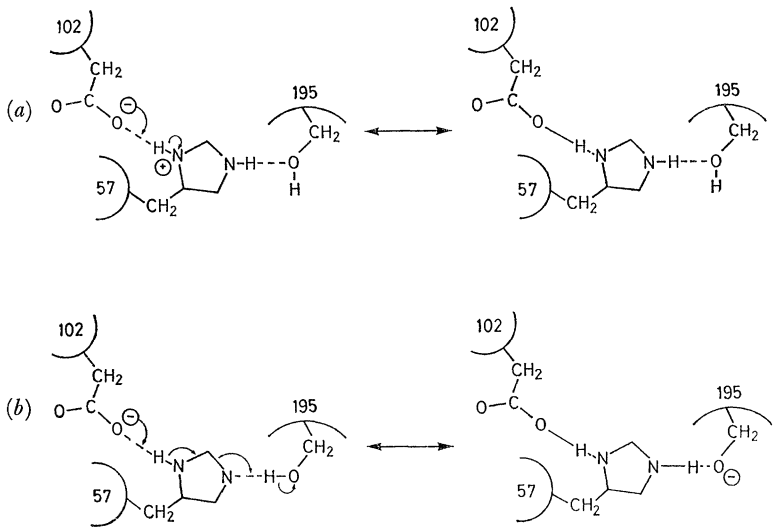


FIGURE 6. Canonical forms of the active centre (*a*) as observed at pH 4; (*b*) assumed at pH 8. The forms are indicated as being related by purely electronic rearrangements. The protons in hydrogen bonds are represented schematically as being near to the atom to which they are covalently bonded in the more 'conventional' of the canonical forms, namely the forms on the left. The possibility of proton transfer is discussed in the text.

Asp-102, His-57 and Ser-195 carries a net negative charge. The residues surrounding the hydrogen bond between Asp-102 and His-57 are non-polar, apart from the polarizable side chain of histidine. Because of the positive charge which will be induced on the side of this residue adjacent to the carboxyl group, and because of the difficulty of access to the solvent in any case, we would expect the proton to be localized on this side of the histidine residue, leaving the ε-nitrogen deprotonated. Under these circumstances the proton of the serine can be brought into such a position as to make a hydrogen bond with the ε-nitrogen of the histidine, simply by a rotation about the β–γ bond of the serine. The possibility of electron transfer by rearrangement of the hydrogen bonds (figure 6*b*) gives rise to a 'charge relay system' in which the negative charge originating from the buried aspartate group may be conveyed through a conduction channel, surrounded by a non-polarizable environment, to the surface of the molecule. In one of the canonical forms of the system, Ser-195 exists as a highly nucleophilic alkoxide group.

Complete rearrangement of the hydrogen bonds would involve transfer of protons along the hydrogen bonds, similar to the proton tunnelling observed in ice. While such a process would be much slower than a purely electronic rearrangement, and is not a necessary part of the above proposal to explain the nucleophilic character of Ser-195, it would be very much faster than the rate of chymotrypsin catalysis (Eigen 1964). It bears an interesting relation to the work of Inagami, York & Patchornik (1965) who found evidence from the chymotrypsin catalysed hydrolysis of amide substrates that a proton was transferred from an acid group on the enzyme to

the substrate during the acylation step. Wang & Parker (1967) suggested that for slowly hydrolysed substrates a pretransition equilibrium takes place between the normal form of the enzyme and an activated species in which a proton is transferred from Ser-195 via His-57 to the amido group of the substrate. The activated species thus contains an alkoxide ion at Ser-195 and a protonated amido group, before the acylation step. This proposal seemed weakened by its quantitative inability to account for the rates observed for fast substrates, due to the small proportion of activated species predicted (Wang 1968). The influence of Asp-102 in helping to draw the proton away from Ser-195 might well be sufficient to overcome this weakness.

4. The binding of substrate molecules to the enzyme

The position of the tosyl group in tosyl-α-chymotrypsin (figure 7, plate 53) seemed to lead to an obvious deduction about the position of binding of the aromatic side chain of a typical 'good' substrate for chymotrypsin. Partly due to the inaccuracies of our original model, which seemed to be extremely crowded in this region, and perhaps partly due to a desire not to accept the obvious uncritically, we refused to become committed to this idea without experimental evidence. During the past twelve months a great deal of effort has been devoted to studying the binding of substrates and substrate analogues to crystalline α-chymotrypsin, both by studying the catalytic properties of the crystals (in collaboration with Mrs S. Banks) and by X-ray diffraction experiments, based on soaking crystals with the various substituents. Until recently, these experiments seemed to agree that substrate analogues would not readily enter the 'tosyl hole'. This type of experiment is fraught with difficulty due to a large number of possible forms of non-specific binding away from the active site and the crowded conditions around the active site due to the close proximity of a dyad related molecule in the crystal.

Recent evidence has confirmed that the 'tosyl hole' is, after all, the substrate binding site. A low resolution difference map prepared from data on crystals soaked in solutions of *N*-formyl-L-tryptophan showed evidence of binding in this position. Also, crystals soaked in a saturated indole solution produced a difference map which included a peak at the 'tosyl site'. A difference map was accordingly calculated at 0.25 nm resolution, using data from crystals soaked in *N*-formyl-L-tryptophan. This map contained a flattened peak in the tosyl hole, extended towards Ser-195. Figure 8, plate 53, shows *N*-formyl-L-tryptophan placed in the model of chymotrypsin in a position consistent with the difference map. While the orientation of the carboxyl group of the inhibitor is not well defined in the difference map, it is probably in contact with

---

Description of plate 53

Figure 7. Skeletal model, showing residues at the active centre of tosyl-α-chymotrypsin. Oxygen, nitrogen and sulphur atoms are indicated by coloured sleeves in red, blue and green respectively. The course of the polypeptide chains is marked by coloured tubing, in blue for the B chain and green for the C chain. The internal ion pair between Asp-194 and Ile-16 is marked by red and blue shapes attached to the charged groups. As in figure 5, the residue marked Asx-102 is indicated as asparagine. This view looks 'horizontally' in towards the active centre: figure 5 was taken from a lower viewpoint, with the camera looking upwards. In this figure Ser-195 is tosylated, and the toluene-sulphonyl group points down and to the right of the serine oxygen, in the space known as the 'tosyl hole'. As described in the text, His-57 and Ser-195 are linked by hydrogen bonding through a water molecule which is not illustrated.

Figure 8. Binding of *N*-formyl-tryptophan to α-chymotrypsin. The *N*-formyl-tryptophan residue, coloured pink, has been placed in a position to give the best fit to the observed difference map. Met-192 and Tyr-146 (extreme right) have been moved slightly to allow accommodation of the tryptophanyl group.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

FIGURE 7

FIGURE 8 (for legends see facing page)

(*Facing p.* 74)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Ser-195 or His-57. The inhibitor amido and carboxyl groups can be orientated to point towards the carbonyl group of Ser-214 and the amido group of Gly-193 respectively, but it is not yet clear whether hydrogen bonds exist.

If the acyl amido group of the substrate is larger than a formyl group, it is probably not possible for it to be accommodated in the crystal. For example, Miss B. A. Jeffery has found that the reagent for His-57, *N*-tosylphenylalanyl chloromethylketone (Ong *et al.* 1964) does not react with the crystals, and that the product of the reaction in solution could be not crystallized (unpublished observations). On the other hand, *N*-formyl-phenylalanyl bromomethylketone (kindly sent to us by Dr Elliott Shaw) reacts readily with crystals and preliminary results indicate that it substitutes adjacent to His-57.

The construction of the more accurate model of chymotrypsin made the shape of the 'tosyl hole' a very much more striking feature of the structure and thus more reasonable as a side chain binding site than it had originally appeared. When a space filling model was assembled, this impression was very much strengthened. The model shows quite clearly the existence of a pocket, suitably shaped for the accommodation of a side chain as large as that of tryptophan, in an orientation convenient to the active centre. The illustration of the space filling model in figure 1 gives an overall impression of its shape, and a groove is seen, rather exaggerated by the stereoscopic view, which runs over the surface of the molecule from lower left to upper right. It seems quite evident from model building that in solution a polypeptide chain could easily be accommodated along the length of the groove. A close-up view of the model (figure 9) shows the tosyl hole, close to the active centre, and indicates its flattened shape. The 'locked substrate' of Hein & Niemann (1962) can be accommodated with a conformation consistent with that shown for *N*-formyl-tryptophan in figure 8, plate 53. Another type of locked substrate recently reported by Belleau & Chevalier (1968) has exactly the conformation shown.

A most important requirement for a side chain specificity site in chymotrypsin is that it must account for the difference in substrate specificity between chymotrypsin and trypsin. Model building studies with α-chymotrypsin indicated that an arginine or lysine residue in the binding site would have its basic group close to residue 189, which is serine in chymotrypsin. The homologous residue in trypsin was indicated as asparagine by Walsh, Kauffman, Kumar & Neurath (1964) but these considerations led Hartley (see this volume, p. 77) to reinvestigate this part of the sequence. He found the residue concerned to be aspartic acid. This provided the strongest evidence that we have correctly identified the substrate binding site in chymotrypsin.

It remains to study the binding of much larger substrates to the molecule. It appears possible to accommodate a polypeptide substrate, or substrate analogue, continued beyond the carboxyl group of *N*-formyl-tryptophan, in crystalline α-chymotrypsin. A substrate extended beyond the *N*-formyl group could not be accommodated, due to intermolecular interactions.

J. J. Birktoft is a postdoctoral research Fellow from Odense University, Denmark, R. Henderson is a Medical Research Council Scholar, and T. A. Steitz is a Fellow of the Jane Coffin Childs memorial fund for medical research.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References (Birktoft *et al.*)

Belleau, B. & Chevalier, R. 1968 *J. Am. chem. Soc.* **90**, 6864.
Bender, M. L. & Kézdy, F. J. 1964 *J. Am. chem. Soc.* **86**, 3704.
Bernal, J. D., Fankuchen, I. & Perutz, M. F. 1938 *Nature, Lond.* **141**, 523.
Biltonen, R., Lumry, R., Madison, V. & Parker, H. 1965 *Proc. natn. Acad. Sci. U.S.A.* **54**, 1412.
Birktoft, J. J., Matthews, B. W. & Blow, D. M. 1969 *Biochem. biophys. Res. Commun.* **36**, 131.
Blow, D. M., Birktoft, J. J. & Hartley, B. S. 1969 *Nature, Lond.* (in the Press).
Blow, D. M., Rossmann, M. G. & Jeffery, B. A. 1964 *J. molec. Biol.* **8**, 65.
Brandts, J. 1964 *J. Am. chem. Soc.* **86**, 4302.
Bruice, T. C. & Benković, S. 1966 *Bioorganic mechanisms*, vol. 1. New York: Benjamin.
Cohen, S. G., Milovanović, A., Schultz, R. M. & Weinstein, S. Y. 1969 *J. biol. Chem.* **244**, 2664.
Cunningham, L. 1965 *Comprehensive Biochem.* **16**, 85 (ed. Florkin and Stotz). Amsterdam, London, New York: Elsevier.
Eigen, M. 1964 *Angew. Chem.* (int. edn) **3**, 1.
Gladner, J. A. & Neurath, H. 1954 *J. biol. Chem.* **206**, 911.
Gutfreund, H. & Sturtevant, J. M. 1956 *Proc. natn. Acad. Sci. U.S.A.* **42**, 719.
Hartley, B. S. 1964 *Nature, Lond.* **201**, 1284.
Hartley, B. S. & Kauffman, D. L. 1966 *Biochem. J.* **101**, 229.
Hein, G. E. & Niemann, C. 1962 *J. Am. chem. Soc.* **84**, 4487, 4495.
Inagami, T., York, S. S. & Patchornik, A. 1965 *J. Am. chem. Soc.* **87**, 126.
Jansen, E. F., Nutting, M. D. F. & Balls, A. K. 1949 *J. biol. Chem.* **179**, 201.
Johnson, C. H. & Knowles, J. R. 1966 *Biochem. J.* **101**, 56.
Kézdy, F. J., Clement, G. E. & Bender, M. L. 1964 *J. Am. chem. Soc.* **86**, 3690.
Koshland, D. E., Strumeyer, D. H. & Ray, W. J. 1962 *Brookhaven Symp. Biol.* **15**, 101.
Kraut, J., Sieker, L. C., High, D. F. & Freer, S. T. 1962 *Proc. natn. Acad. Sci. U.S.A.* **48**, 1417.
Kraut, J., Wright, H. T., Kellerman, M. & Freer, S. T. 1967 *Proc. natn. Acad. Sci. U.S.A.* **58**, 304.
Kunitz, M. & Northrop, J. H. 1935 *J. gen. Physiol.* **18**, 433.
Matthews, B. W., Cohen, G. H., Silverton, E. W., Braxton, H. & Davies, D. R. 1968 *J. molec. Biol.* **36**, 179.
Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. 1967 *Nature, Lond.* **214**, 652.
Meloun, B., Kluh, I., Kostka, V., Morávek, L., Prusík, Z., Vaněček, J., Keil, B. & Šorm, F. 1966 *Biochim. biophys. Acta* **130**, 543.
Némethy, G., Phillips, D. C., Leach, S. J. & Scheraga, H. A. 1967 *Nature, Lond.* **214**, 363.
Neurath, H. & Hartley, B. S. 1959 *J. cell. comp. Physiol.* **54** (Suppl. 1), 179.
Ong, E. B., Shaw, E. & Schoellman, G. 1964 *J. Am. chem. Soc.* **86**, 1271.
Oppenheimer, H. L., Labouesse, B. & Hess, G. P. 1966 *J. biol. Chem.* **241**, 2720.
Ray, W. J. & Koshland, D. E. 1960 *Brookhaven Symp. Biol.* **13**, 135.
Reeke, G. N., Hartsuck, J. A., Ludwig, M. L., Quiocho, F. A., Steitz, T. A. & Lipscomb, W. N. 1967 *Proc. natn. Acad. Sci. U.S.A.* **58**, 2220.
Richards, F. M. 1968 *J. molec. Biol.* **37**, 225.
Rovery, M., Poilroux, M., Yoshida, A. & Desnuelle, P. 1957 *Biochim. biophys. Acta* **23**, 608.
Shaffer, N. K., May, S. C. & Summerson, W. H. 1953 *J. biol. Chem.* **202**, 67.
Sigler, P. B., Blow, D. M., Matthews, B. W. & Henderson, R. 1968 *J. molec. Biol.* **35**, 143.
Sigler, P. B., Jeffery, B. A., Matthews, B. W. & Blow, D. M. 1966 *J. molec. Biol.* **15**, 175.
Smillie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J. & Parkes, C. O. 1968 *Nature, Lond.* **218**, 343.
Venkatachalam, C. M. 1968 *Biopolymers* **6**, 1425.
Walsh, K. A., Kauffman, D. L., Kumar, K. S. V. S. & Neurath, H. 1964 *Proc. natn. Acad. Sci. U.S.A.* **51**, 301.
Wang, J. H. 1968 *Science, N.Y.* **161**, 328.
Wang, J. H. & Parker, L. 1967 *Proc. natn. Acad. Sci. U.S.A.* **58**, 2451.
Zerner, B., Bond, R. P. M. & Bender, M. L. 1964 *J. Am. chem. Soc.* **86**, 3674.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

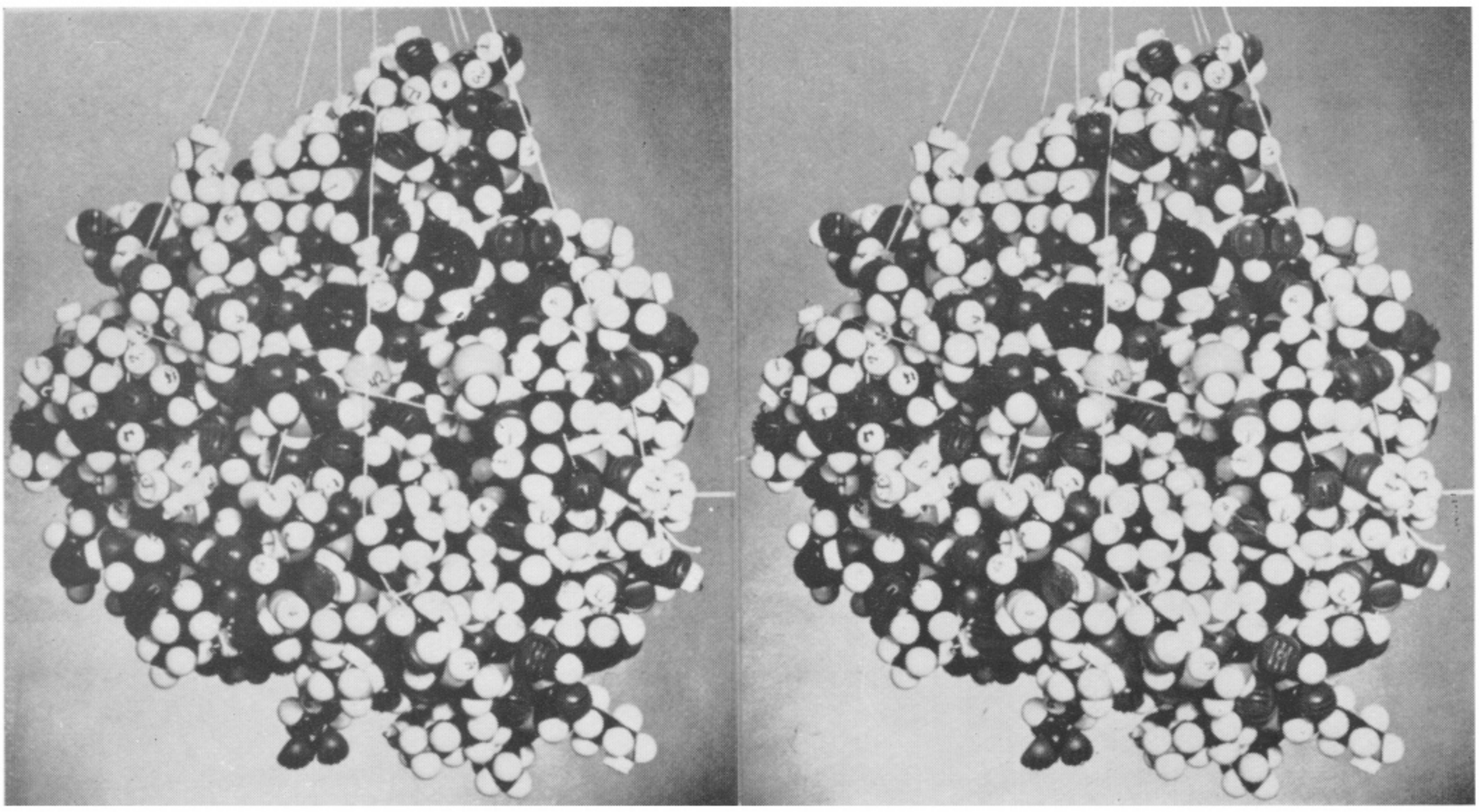


FIGURE 1

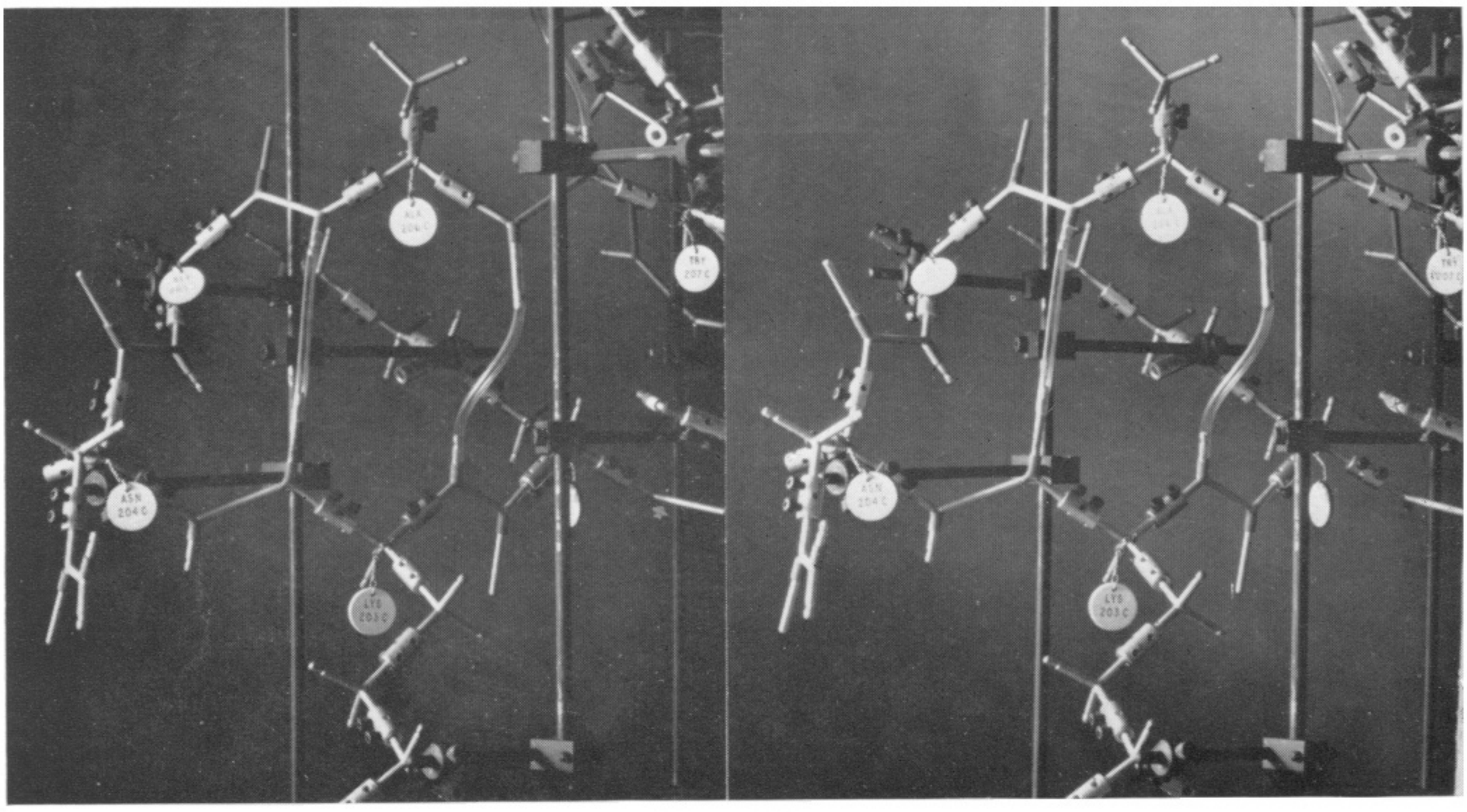


FIGURE 3

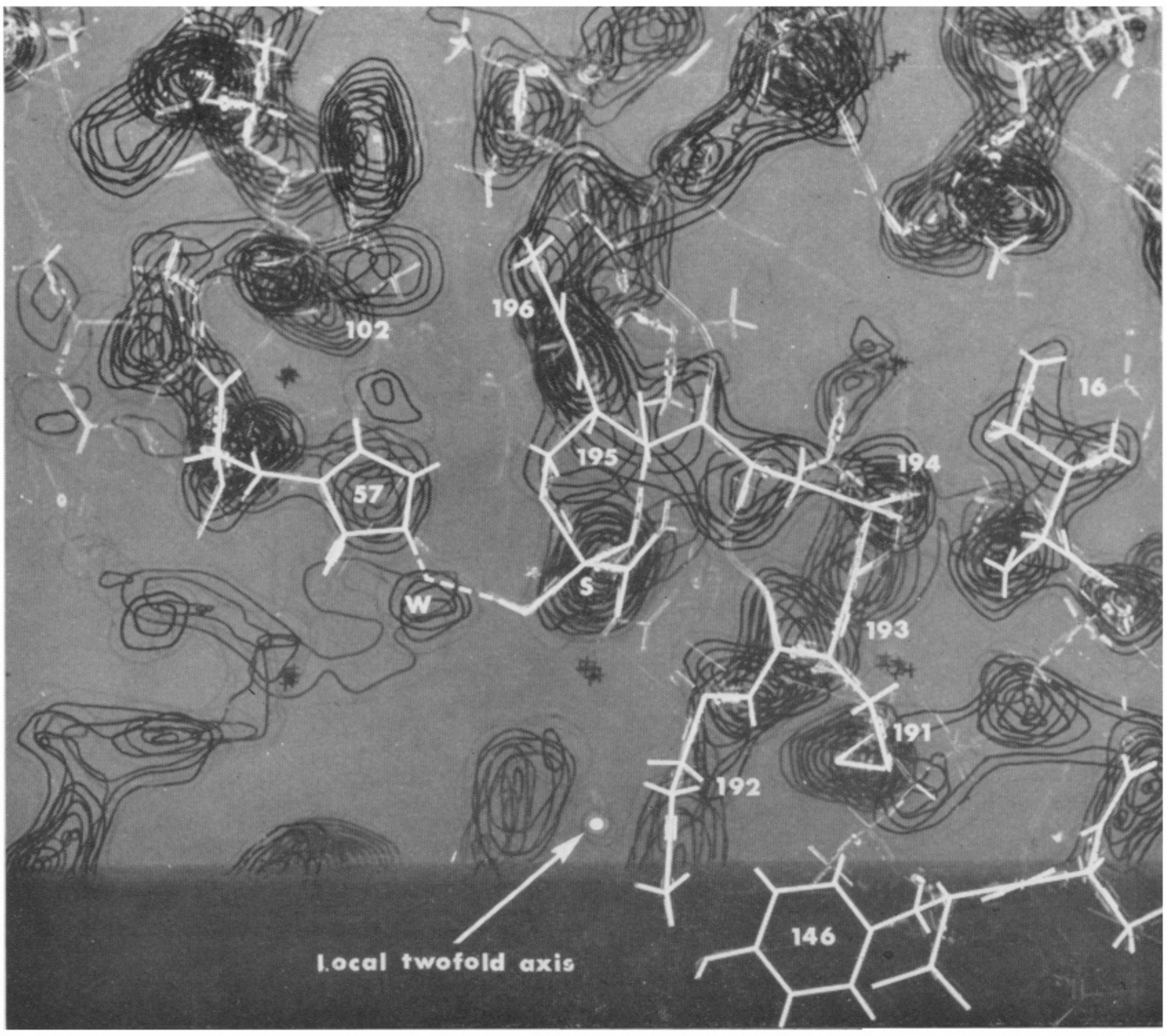
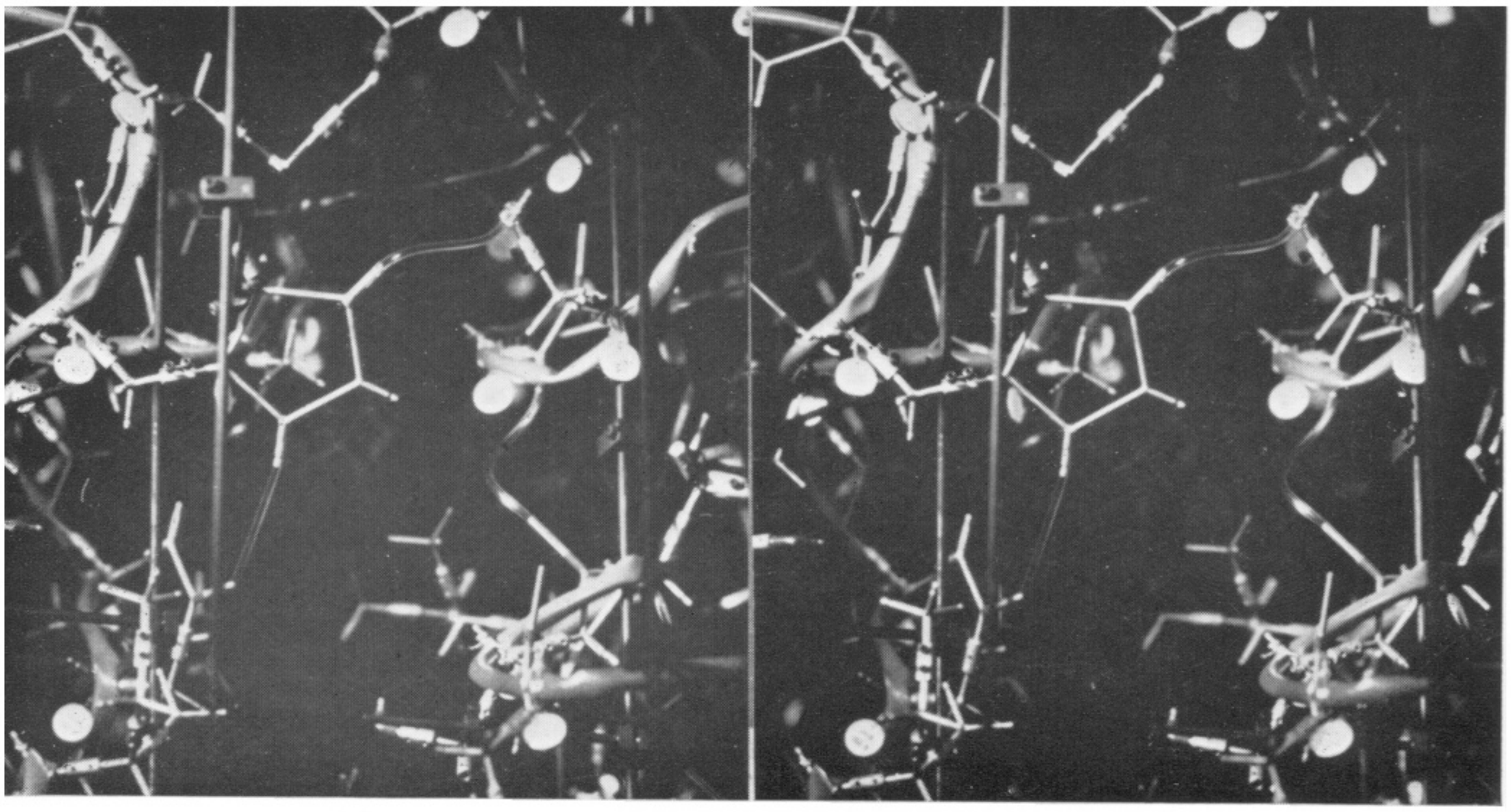
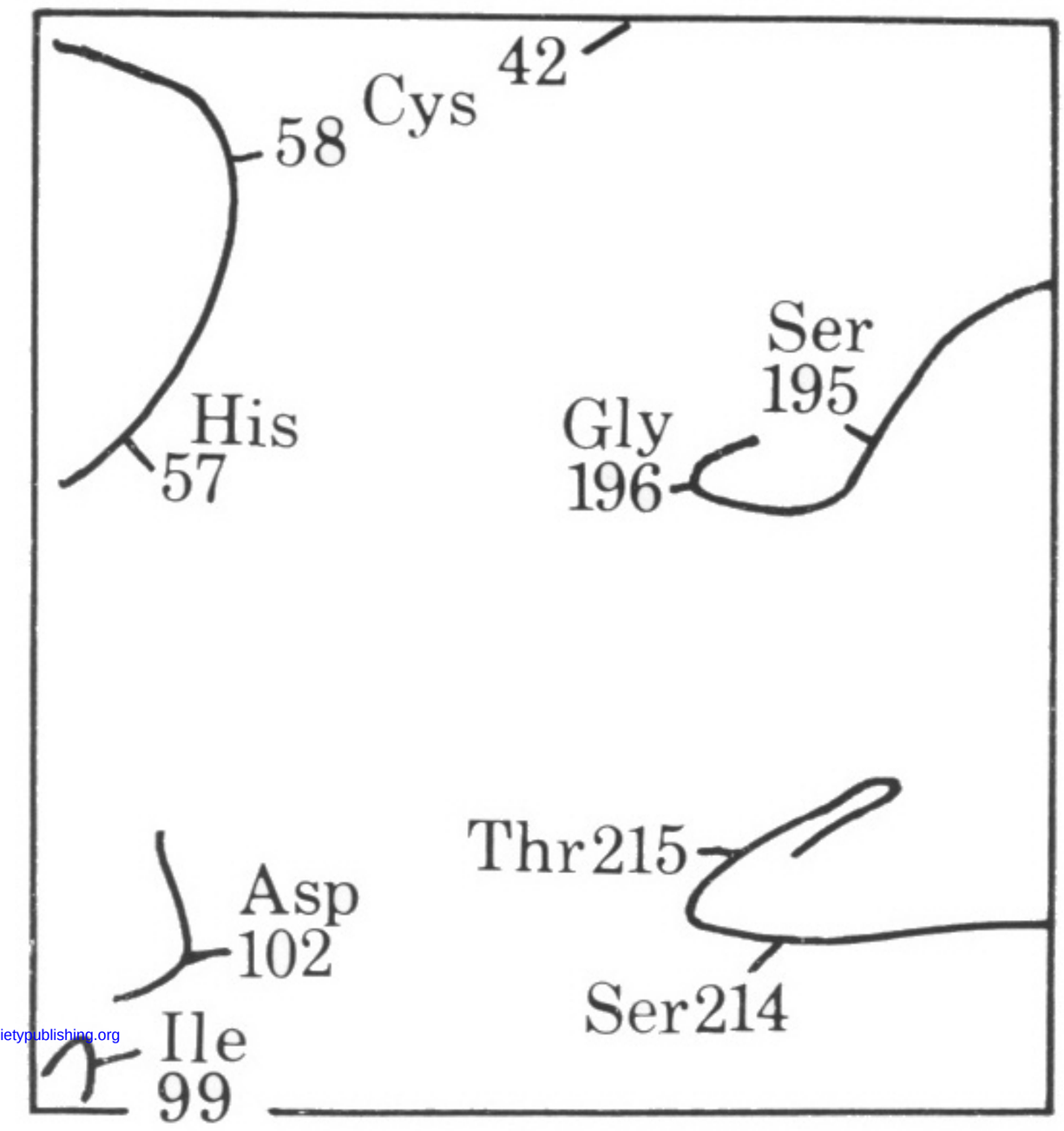


FIGURE 4 (for legends see facing page)



(*a*) (*b*)



(*c*)

FIGURE 5

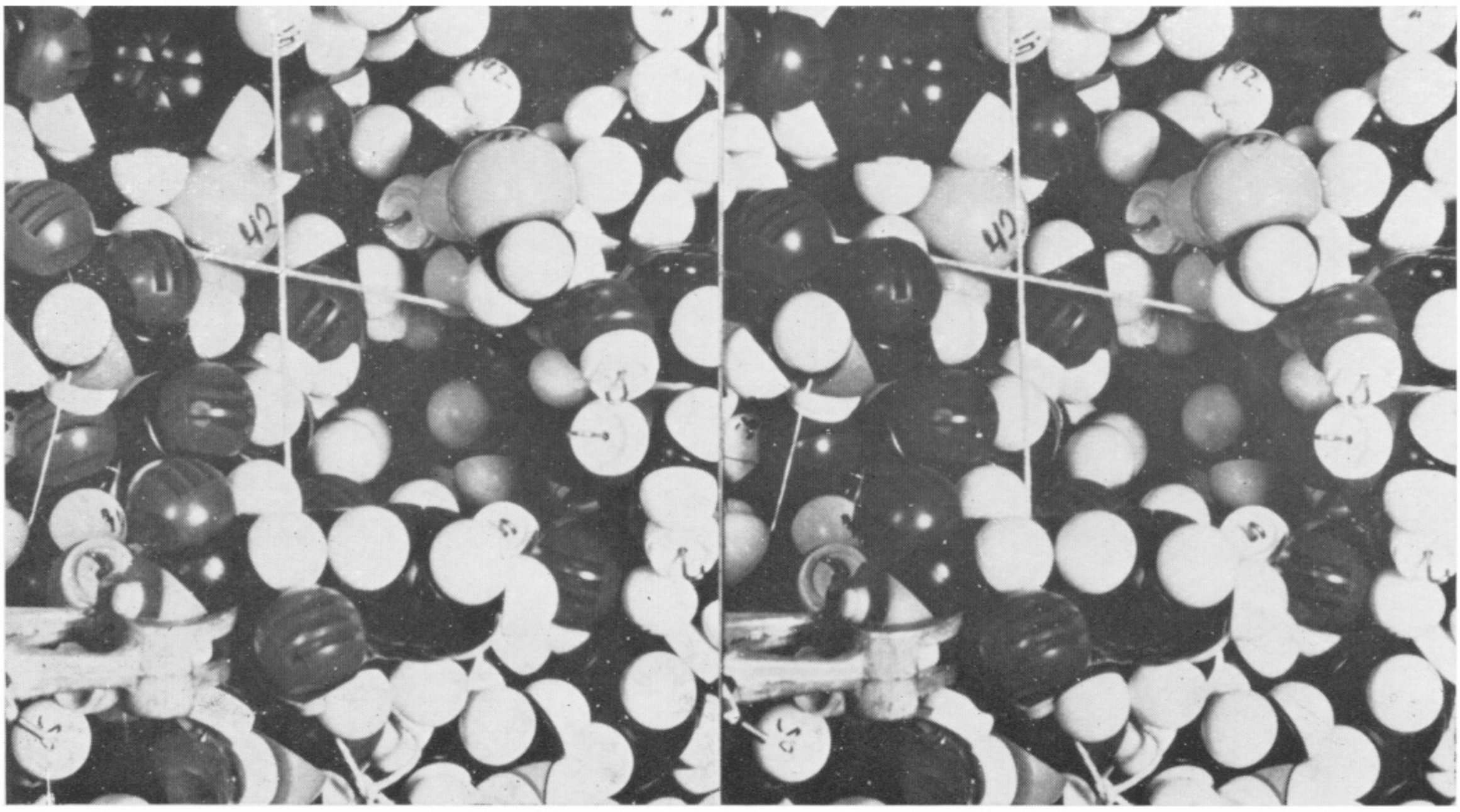


FIGURE 9 (for legends see facing page)

FIGURE 7

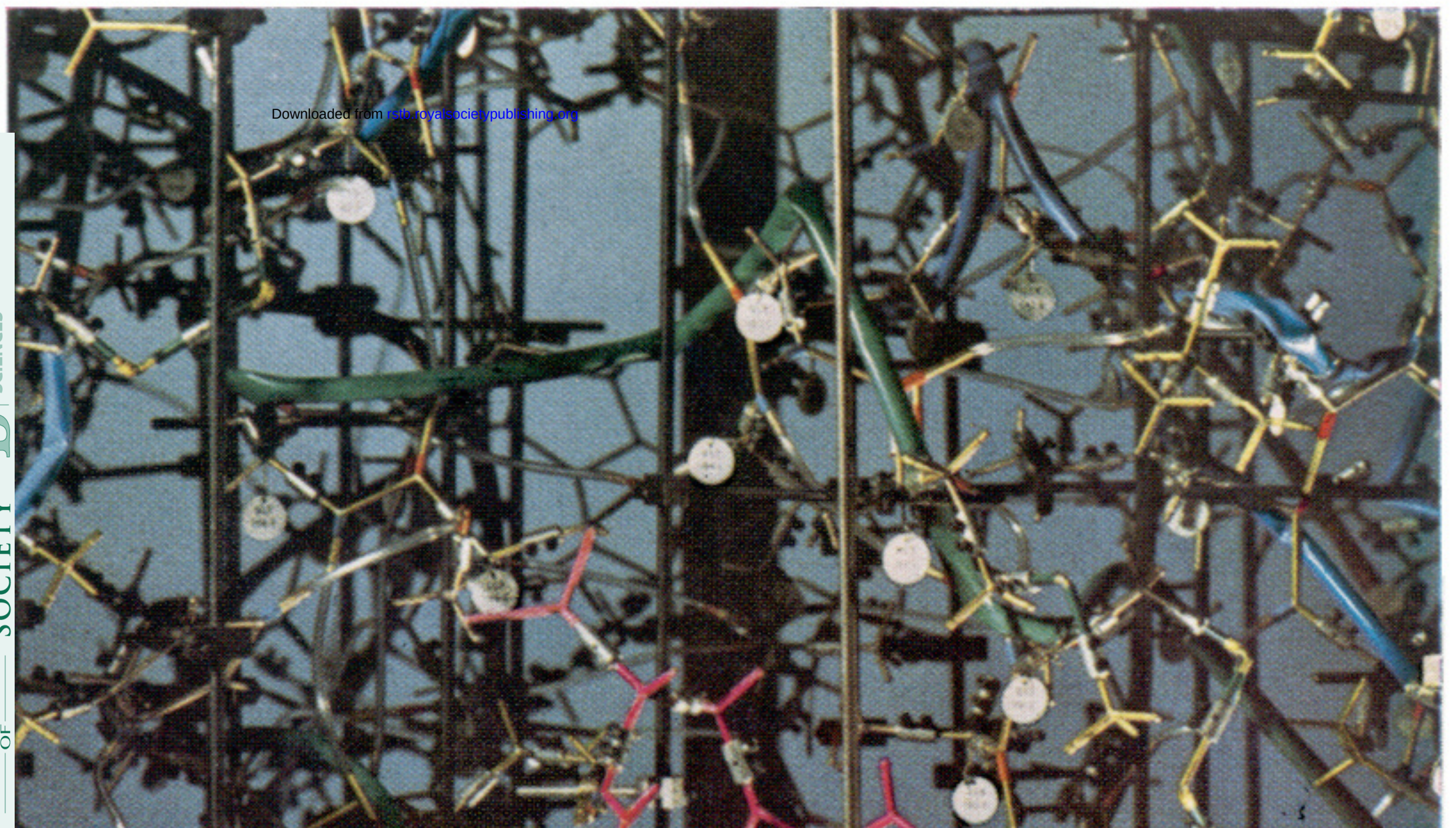


FIGURE 8 (for legends see facing page)