PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A Hydrogen-Bond Network at the Active Site of Subtilisin BPN<latex>$\prime$</latex>

R. A. Alden, C. S. Wright and J. Kraut

Phil. Trans. R. Soc. Lond. B 1970 **257**, 119-124
doi: 10.1098/rstb.1970.0014

Phil. Trans. Roy. Soc. Lond. B **257**, 119–124 (1970) [119]
Printed in Great Britain

A hydrogen-bond network at the active site of subtilisin BPN′

By R. A. Alden, C. S. Wright and J. Kraut
Department of Chemistry, University of California, San Diego, La Jolla, California, 92037, *U.S.A.*

[Plates 56 and 57]

Further examination of the active site region in our X-ray crystallographic model of subtilisin BPN′ reveals a hydrogen-bond network that bears a remarkable resemblance to the one found in α-chymotrypsin. It involves the side chains of the reactive Ser-221, His-64, Asp-32 and Ser-33. Otherwise the two enzymes have entirely different three-dimensional structures. This observation suggests that the common hydrogen bond network plays some essential role in the catalytic mechanism of serine proteases generally.

Introduction

We have recently completed a three-dimensional X-ray crystallographic structure analysis of subtilisin BPN′ at 0.25 nm resolution by the method of multiple isomorphous replacement. The resulting electron density map was of sufficiently good quality to enable the entire backbone chain of 275 residues to be followed unambiguously and almost all of the side chains to be identified and positioned with the aid of the amino acid sequence previously determined by Smith, DeLange, Evans, Landon & Markland (1968). Based on our interpretation of this map, a Kendrew–Watson skeletal model representing the entire molecule has been constructed on a scale of 2 cm to 0.1 nm. A preliminary description of the structure, together with an outline of the experimental details, has recently appeared (Wright, Alden & Kraut 1969).

Although subtilisin BPN′ and α-chymotrypsin show a considerable degree of similarity in their respective enzymic mechanisms, and some degree of similarity with regard to specificity toward small-molecule substrates and inhibitors, the two amino acid sequences are entirely unrelated. Not surprisingly, therefore, crystallographic analysis has revealed that the overall three-dimensional structures of the two molecules are also quite different. Nevertheless, it has now become apparent that there are some unexpected resemblances between subtilisin BPN′ and α-chymotrypsin when one compares the hydrogen-bond network connecting various side chains in the vicinity of their respective active serine residues, and it seems very likely that such a comparison will provide an important clue to the possible existence of any common structural feature that may be essential to the functioning of the serine proteases. Indeed, we are probably justified in assuming that the two molecules evolved independently, and thus may represent a case of convergent evolution at the level of three-dimensional molecular structure. For these reasons, we would like to take this opportunity to present here a detailed description of the active site region of subtilisin BPN′ as it appears in our current model of the molecule, and to compare it qualitatively with α-chymotrypsin.

Enzymic properties of subtilisin

Subtilisin BPN′ is an extracellular protease of the sporeforming organism *Bacillus amyloliquefaciens*. The particular strain producing this enzyme was formerly designated the N′ strain of the more familiar *B. subtilis* (Hagihara 1960), but Welker & Campbell (1967) have shown that it is more properly classified as a strain of the unrelated *B. amyloliquefaciens*. The enzyme consists of

a single polypeptide chain of 275 residues, with no disulphide bridges or SH groups. Subtilisin Novo is apparently identical with BPN′ (Smith 1967). A closely similar enzyme, designated subtilisin Carlsberg, is presumably produced by *B. subtilis* itself, and differs from BPN′ in 84 amino acid replacements and one deletion. The amino acid sequences of both subtilisins have been determined by Smith *et al.* (1968). The amino acid sequence of subtilisin BPN′ is shown in figure 1. A third member of this class of enzymes, containing roughly 40 to 50 residue differences from BPN′, is produced by *B. subtilis* var. amylosacchariticus (F. S. Markland 1968, private communication). Indeed, subtilisin-type proteases appear to be widely distributed among the bacteria and molds (Hagihara 1960; Morihara 1967).

Considering their respective enzymic activities, the subtilisins show considerable similarity to chymotrypsin and the other serine proteases of animal origin. Both classes of enzyme contain a single peculiarly reactive serine residue that can be specifically phosphorylated by diisopropylfluorophosphate (Sanger & Shaw 1960), sulphonylated by phenylmethanesulphonyl fluoride (Neet & Koshland 1966; Polgar & Bender 1967) or acylated by a variety of reagents. What is more, Bernhard, Lau & Noller (1965) have found that both α-chymotrypsin and subtilisin Novo show the same characteristic shift in the absorption maximum of the acylating chromophoric group when the enzymes react with certain β-arylacryloylimidazoles. This effect was subsequently interpreted by Charney & Bernhard (1967) as resulting from conversion to the s-*cis* configuration at the acryloyl single bond in the native acylated enzyme. Further, Keizer & Bernhard (1966) obtained parallel effects on the respective protonic equilibria exhibited by both α-chymotrypsin and subtilisin Novo when the enzymes were acylated with *N*-β-(3-indole)-acryloylimidazole.

Another similarity between the enzymic properties of subtilisin and the other serine proteases is the involvement of a histidine residue. Considerable evidence for this is scattered throughout the literature. Polgar & Bender (1967) find the k_{cat} for hydrolysis of *p*-nitrophenyl acetate by BPN′ depends upon a group with pK about 7.2, presumably a histidine. Glazer (1967) has shown that the pH–velocity profile for hydrolysis of *N*-benzoyl-L-arginine ethyl ester by subtilisin Carlsberg, and its temperature dependence, are consistent with requirement of a non-protonated imidazole group in the active enzyme. Recently, Glazer (1968) has studied the inhibition of chymotrypsin and subtilisin by various phenylarsonates, and has found that binding of the inhibitors for both enzymes requires protonation of a group of pK' = 7.15, although the kinetics of the process for the two enzymes differed. Again, the group is presumed to be a histidine residue.

Regarding substrate specificities, the general picture that emerges when comparing the animal serine proteases with subtilisin is that the latter is relatively poor as an esterase, but is more effective as a protease (Ottesen & Spector 1960; Hill 1965). This may simply be a reflection of subtilisin's less stringent specificity requirements. Nevertheless, the best small-molecule substrates for chymotrypsin, like *N*-acetyl-L-tyrosine ethyl ester, or *N*-benzoyl-L-tyrosine ethyl ester are also good substrates for subtilisin (Barel & Glazer, 1968; Johansen, Ottesen & Svendsen 1967). Similarly, a number of aromatic compounds, like indole, phenol and hydrocinnamate are competititive inhibitors for both enzymes (Glazer 1967). Finally, both enzymes hydrolyse amide substrates several orders of magnitude more slowly than esters (Glazer 1967).

Although in most of the above we have emphasized the similarities in enzymic properties between subtilisin and chymotrypsin, it should not be overlooked that there are distinct differences. For example, proflavine, a good competititive inhibitor for chymotrypsin, is ineffective

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

against subtilisin (Glazer 1967), while conversely, 4-(4′-aminophenylazo)-phenylarsonic acid is a strong subtilisin inhibitor, but does not interact with chymotrypsin (Glazer 1968). Then, too, despite the involvement of a histidine residue in the active site of subtilisin, the specific chloromethylketone inhibitors for chymotrypsin and trypsin, known to function by alkylating the

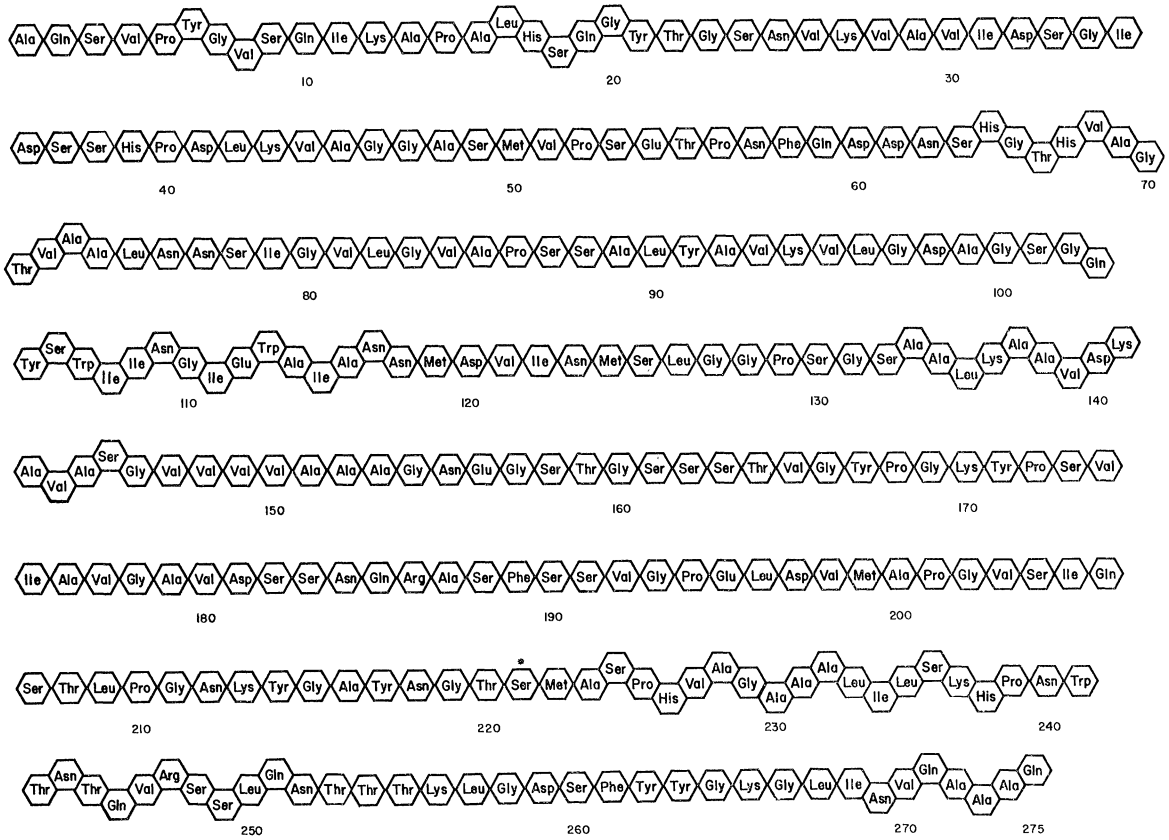


FIGURE 1. Amino acid sequence of subtilisin BPN′, according to Markland & Smith (1967). Segments of α-helix are indicated by wavy portions.

active site histidine of those enzymes, do not react with subtilisin (Smith 1967). These differences are hardly surprising when the amino acid sequences of subtilisin BPN′ (figure 1) or Carlsberg (Smith *et al.* 1968) are compared with the sequences for the animal serine proteases (Smilie Furka, Nagabhushan, Stevenson & Parkes 1968). Rather, it is the similarities that seem to require explanation. For example, subtilisin contains no disulphide bridges, while chymotrypsin has five, and trypsin six. Perhaps more significantly, even in the immediate vicinity of the active serine residues (Ser-221 in subtilisin), the two classes of enzymes show no resemblance at all, as far as sequence is concerned.

DESCRIPTION OF THE MOLECULE

In this section we shall briefly describe the overall three-dimensional structure of the subtilisin BPN′ molecule as revealed by our crystallographic studies. Figure 2, plate 56, is a stereoscopic view of the entire skeletal model, including all the side chains. A length of knitting yarn has been threaded through the model to assist in following the course of the polypeptide backbone. A

smaller model showing only the backbone chain was also constructed, to a scale of 5 cm to 1 nm by soldering together short lengths of 3 mm copper tubing connecting α-carbon atom positions. Figure 3, plate 57, is a stereoscopic view of this backbone-chain model, taken from the same vantage point as figure 2. In both cases, the view is directly toward the active site, perpendicular to the plane of the imidazole ring of His-64 as it is oriented in the active enzyme molecule.

The overall molecular shape is approximately spherical, with a diameter of about 4.2 nm. Although something of a depression exists at the active site, it is not especially obvious in comparison with other surface irregularities elsewhere on the molecule. As has so far been the case for all other known protein structures, the core of the subtilisin BPN′ molecule is also composed largely of packed non-polar side chains. Side chains projecting into the surrounding medium or lying in surface crevices may be either polar or non-polar.

Although the molecule consists of a single continuous polypeptide chain, the backbone-chain model shown in figure 3 is painted in three colours. The section from Ala-1 to Gly-100 appears as black, Gly-100 to Ala-176 as white, and Ala-176 to C-terminal Gln-275 as grey in the photograph. It is immediately obvious that the backbone chain has folded into three distinct and potentially separable pieces. Most interestingly, the active site occurs at one of the two places on the molecular surface where the three pieces come together.

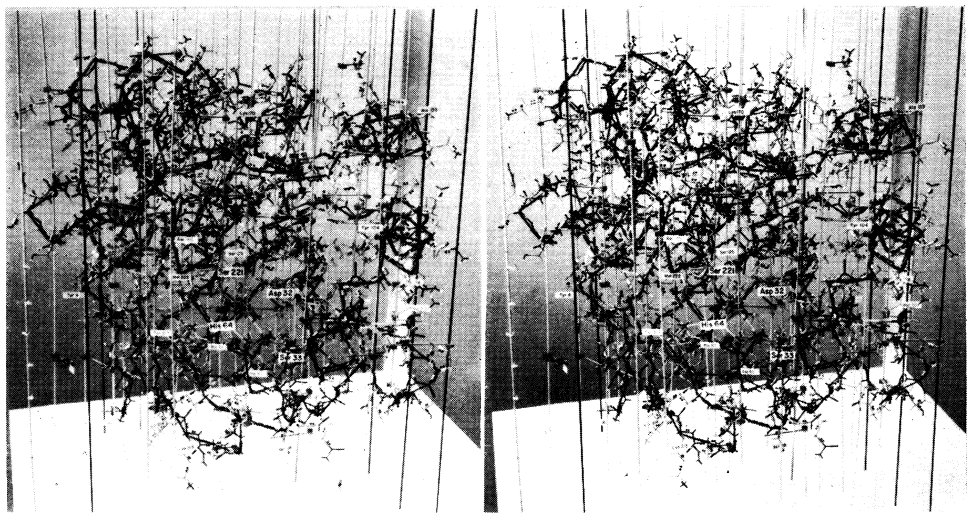
The molecule contains eight right-handed α-helical segments, from 6 to 16 residues long. These helical regions are shown as wavy sections of sequence in figure 1. They comprise about 31% of the 275 residues in the chain. Many distortions from the standard α_I-helix conformation are observed. The longest helix, consisting of residues Ala-223 to His-238, runs roughly through the centre of the molecule, and six of the remaining seven helical segments lie approximately parallel to it.

Another notable feature of the chain conformation is a twisted parallel-chain pleated sheet composed of five segments of extended chain. The angle of twist between the top and bottom segments is approximately 45°. The side chains of this pleated sheet, together with those of the central helix, form a large part of the non-polar core of the molecule.

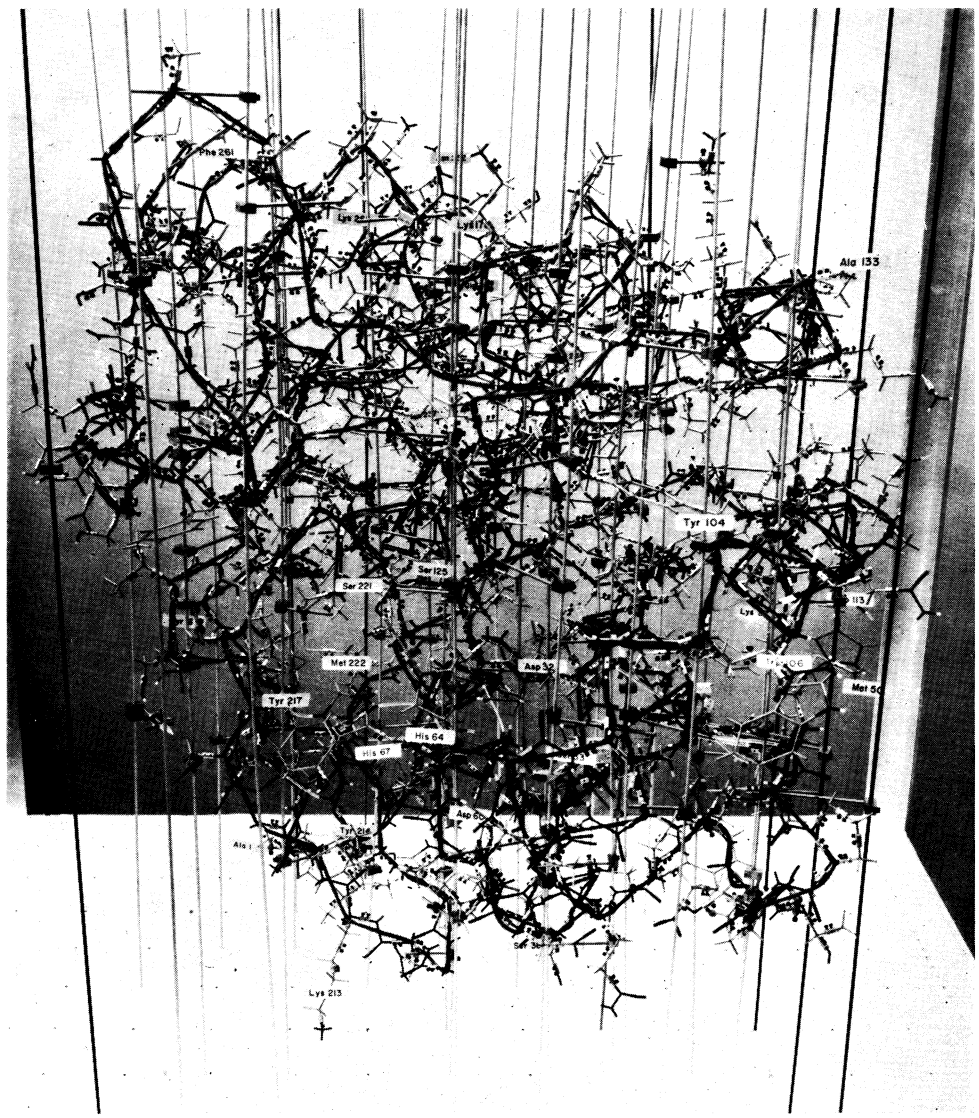
The active site

For technical reasons, the parent protein crystal whose phases were estimated by isomorphous replacement, and for which the electron-density map was interpreted to give the picture of the molecule just described, was in fact not the active enzyme, but the PMS inhibited enzyme. Fortunately, crystals of the active enzyme itself are isomorphous with the inhibited crystals, and so we were able to calculate a difference electron density map between the two. As detailed in our earlier publication (Wright *et al.* 1969), this difference map clearly showed the PMS group and thus confirmed the location of reactive Ser-221. It also revealed a 0.4 nm movement of the side chain of His-64 upon binding of the PMS group to Ser-221, as well as a slight (*ca.* 0.1 nm) movement of the side chain of Met-222. We have thus been able to reconstruct the region of the active site as it appears in the active enzyme. A stereoscopic close-up of this region is shown in figure 4, plate 57. In effect, figure 4 is a 'zoom-in' toward the active site from the vantage point of figures 2 and 3. Similarly, figure 5, plate 57, is a close-up of the same region as it appears in the PMS inhibited enzyme.

From figure 4 it will be apparent that the side chains of several residues in the vicinity of the



(*a*)



(*b*)

FIGURE 2. (*a*) Stereoscopic view of subtilisin BPN′ skeletal model, looking directly toward the active site. Viewing direction is approximately perpendicular to the imidazole ring of His-64 in the active conformation. The same viewing direction is maintained in figures 3 to 5 as well. Knitting yarn follows the polypeptide backbone chain. (*b*) This is a larger photograph of the model.

(*Facing p.* 122)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

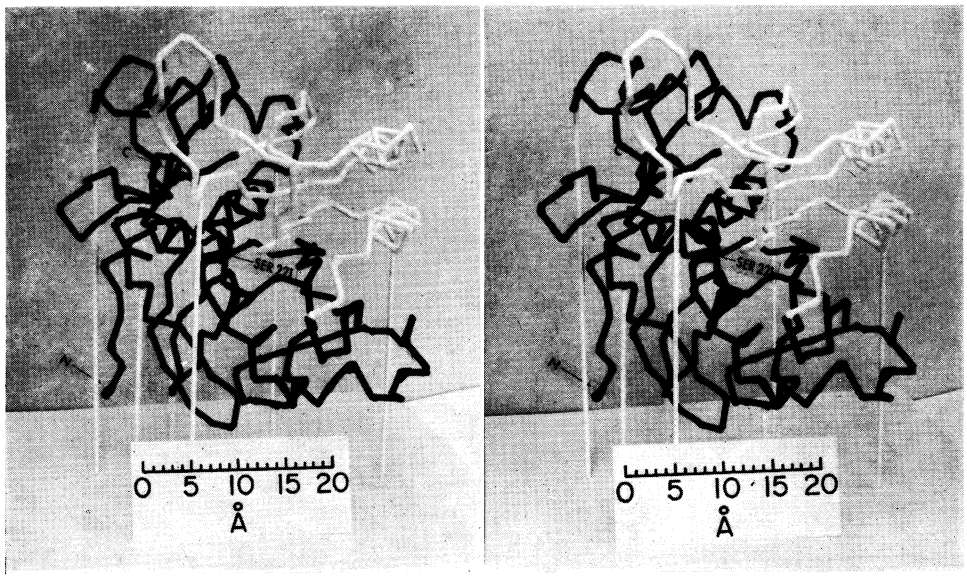


FIGURE 3. Backbone chain model, showing folding of single continuous polypeptide chain into three potentially separable pieces, at the conjunction of which is the active site.

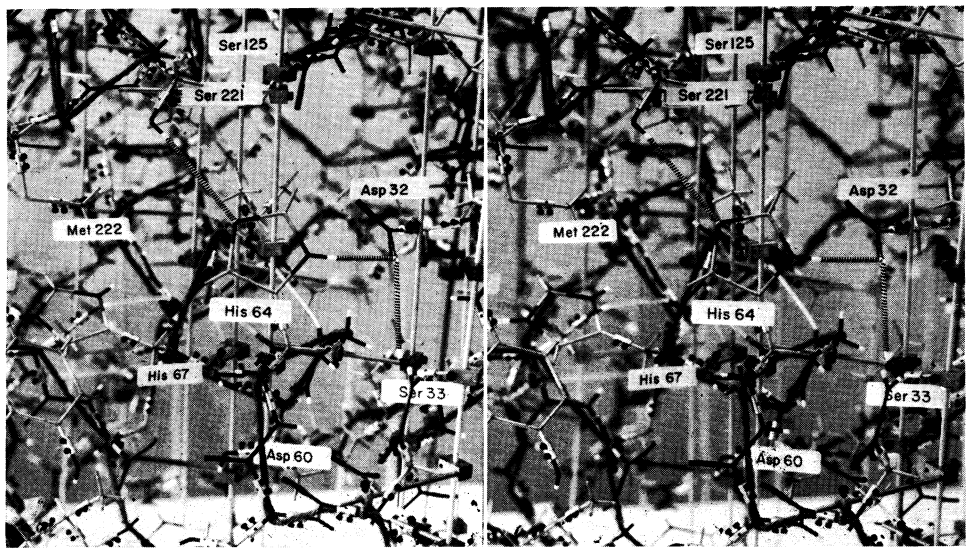


FIGURE 4. Close-up of the active site region as it appears in the active enzyme. Hydrogen bonds that are analogous to those found in α-chymotrypsin are emphasized by insertion of dotted connectors.

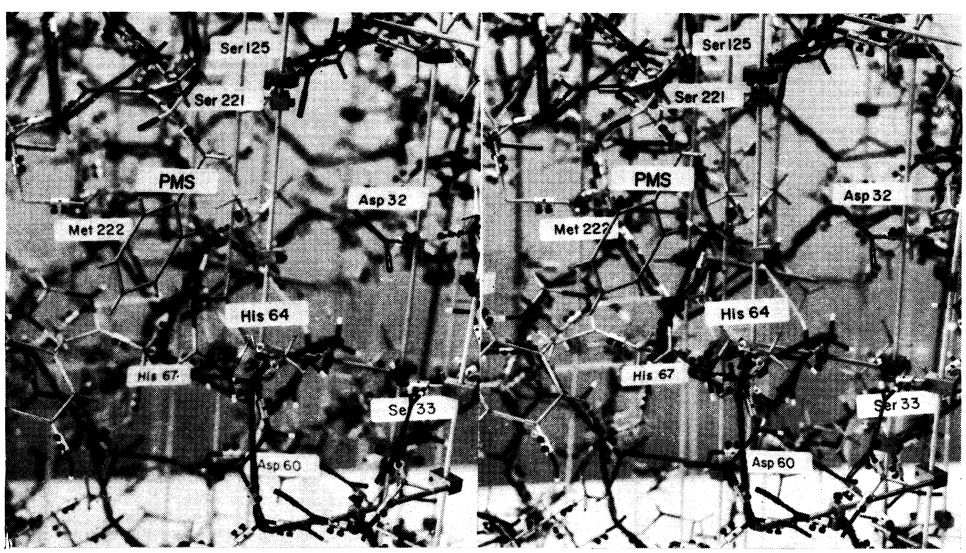


FIGURE 5. Close-up of active site region with phenylmethanesulphonyl inhibitor group covalently bonded to Ser-221. Note movement of His-64 side chain.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

active site are in positions that strongly suggest that they are connected by a rather complex hydrogen bond network. We shall direct attention first to those features of the hydrogen bond network that are now thought to be common to the active conformation of both subtilisin BPN′ and α-chymotrypsin (D. M. Blow 1968, personal communication). They appear as dotted connectors in figure 4.

(1) A hydrogen bond from the OH group of reactive Ser-221 to $N^{\epsilon 2}$ of His-64 (atom numbering system is that proposed by Edsall, Flory, Kendrew, Liquori, Némethy & Ramachandran 1966). The O^{γ} to $N^{\epsilon 2}$ distance is 0.35 nm in our present model—rather long for a standard hydrogen bond but probably within allowable tolerances for the present unrefined 0.25 nm model.

(2) A hydrogen bond from $N^{\delta 1}$ of His-64 to the side chain $O^{\delta 2}$ of Asp-32. The N...O distance is 0.22 nm.

(3) A second hydrogen bond to $O^{\delta 2}$ of Asp-32 from the side chain OH group of Ser-33. The O...O distance is 0.31 nm.

The five atoms involved in this part of the hydrogen bond network all lie approximately within the plane of the imidazole ring. Their mean deviation from a least-squares plane is 0.01 nm, and the greatest distance from this plane is 0.02 nm, for Oγ of Ser-33. The same hydrogen bond system in α-chymotrypsin involves the side chains of Ser-195, His-57, Asp-102 and Ser-214 (D. M. Blow 1968, personal communication). The close similarity in this regard between α-chymotrypsin and subtilisin BPN′ strongly suggest some fundamental role for this hydrogen bond network in the mechanism of serine protease activity.

There are also other apparent hydrogen bonds in the region of the active site of subtilisin BPN′, but these do not have exact counterparts in α-chymotrypsin (Blow 1968).

(1) Three other potentially hydrogen bonding groups are crowded closely together in the neighbourhood of the side chain of Ser-33. These are the backbone NH groups of His-64 and Gly-65 and the side chain of Asp-60. Although the distances are sufficiently short to suggest hydrogen bond formation, the angles are rather unconvential in the present model. In α-chymotrypsin, the backbone NH group of His-57 is hydrogen bonded to $O^{\delta 1}$ of Asp-102, i.e. to the oxygen atom *not* involved in hydrogen bond formation with the side chain of His-57.

(2) There is a close approach, 0.29 nm, between $C^{\epsilon 1}$ of His-64 and the backbone carbonyl of Ser-125.

(3) The side chain OH group of Ser-125 forms a hydrogen bond to the backbone carbonyl of Ser-221. In α-chymotrypsin, the backbone carbonyl of Ser-195 is also hydrogen bonded, but the donor group is the backbone NH group of Gly-43.

Whether or not these additional features of the active site hydrogen bond network have any special significance is, of course, impossible to say at this stage, but, as pointed out above, there are certain resemblances to α-chymotrypsin here as well.

Finally, it should be noted that all the components of the active site hydrogen bond network are also present in subtilisin Carlsberg. That is to say, Asp-32, Asp-60, His-64, Ser-125 and, of course, the reactive Ser-221 are unchanged. Ser-33 is changed to a threonine, but the latter can easily be accommodated in the structure to retain an OH group in hydrogen bonding position. A detailed examination of the geometrical distribution of the 85 amino acid differences between subtilisins BPN′ and Carlsberg (Wright *et al.* 1969) suggests that the two structures are very similar.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

This research was supported by grants from the U.S. Public Health Service and the National Science Foundation. C. S. Wright is a predoctoral fellow of the U.S. National Institutes of Health.

References (Alden *et al.*)

Barel, A. O. & Glazer, A. N. 1968 *J. biol. Chem.* **243**, 1344.
Bernhard, S. A., Lau, S. J. & Noller, H. 1965 *Biochemistry* **4**, 1108.
Charney, E. & Bernhard, S. A. 1967 *J. Am. chem. Soc.* **89**, 2726.
Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Némethy, G. & Ramachandran, G. N. 1966 *J. molec. Biol.* **15**, 399.
Glazer, A. N. 1967 *J. biol. Chem.* **242**, 433.
Glazer, A. N. 1968 *J. biol. Chem.* **243**, 3693.
Hagihara, B. 1960 Chapter 9 in *The enzymes*, vol. 4 (ed. P. D. Boyer, H. Lardy and K. Myrbäck). New York: Academic Press.
Hill, R. L. 1965 *Adv. Protein Chem.* **20**, 37.
Johansen, J. T., Ottesen, M. & Svendsen, I. 1967 *Biochim. biophys. Acta* **139**, 211.
Keizer, J. & Bernhard, S. A. 1966 *Biochemistry* **5**, 4127.
Markland, F. S. & Smith, E. L. 1967 *J. biol. Chem.* **242**, 5198.
Morihara, K. 1967 *Biochem. biophys. Res. Commun.* **26**, 656.
Neet, K. E. & Koshland, D. E. 1966 *Proc. natn. Acad. Sci. U.S.A.* **56**, 1606.
Ottesen, M. & Spector, A. 1960 *C. r. Trav. Lab. Carlsberg* **32**, 63.
Polgar, L. & Bender, M. L. 1967 *Biochemistry* **6**, 610.
Sanger, F. & Shaw, D. C. 1960 *Nature, Lond.* **187**, 872.
Smilie, L. B., Furka, A., Nagabhushan, N., Stevenson, K. J. & Parkes, C. O. 1968 *Nature, Lond.* **218**, 343.
Smith, E. L. 1967 *Abst. 7th Int. Congr. Biochem. Tokyo*, I-1, 5; p. 9.
Smith, E. L., DeLange, R. J., Evans, W. H., Landon, M. & Markland, F. S. 1968 *J. biol. Chem.* **243**, 2184.
Welker, N. E. & Campbell, L. L. 1967 *J. Bact.* **94**, 1124.
Wright, C. S., Alden, R. A. & Kraut, J. 1969 *Nature, Lond.* **221**, 235.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(*a*)

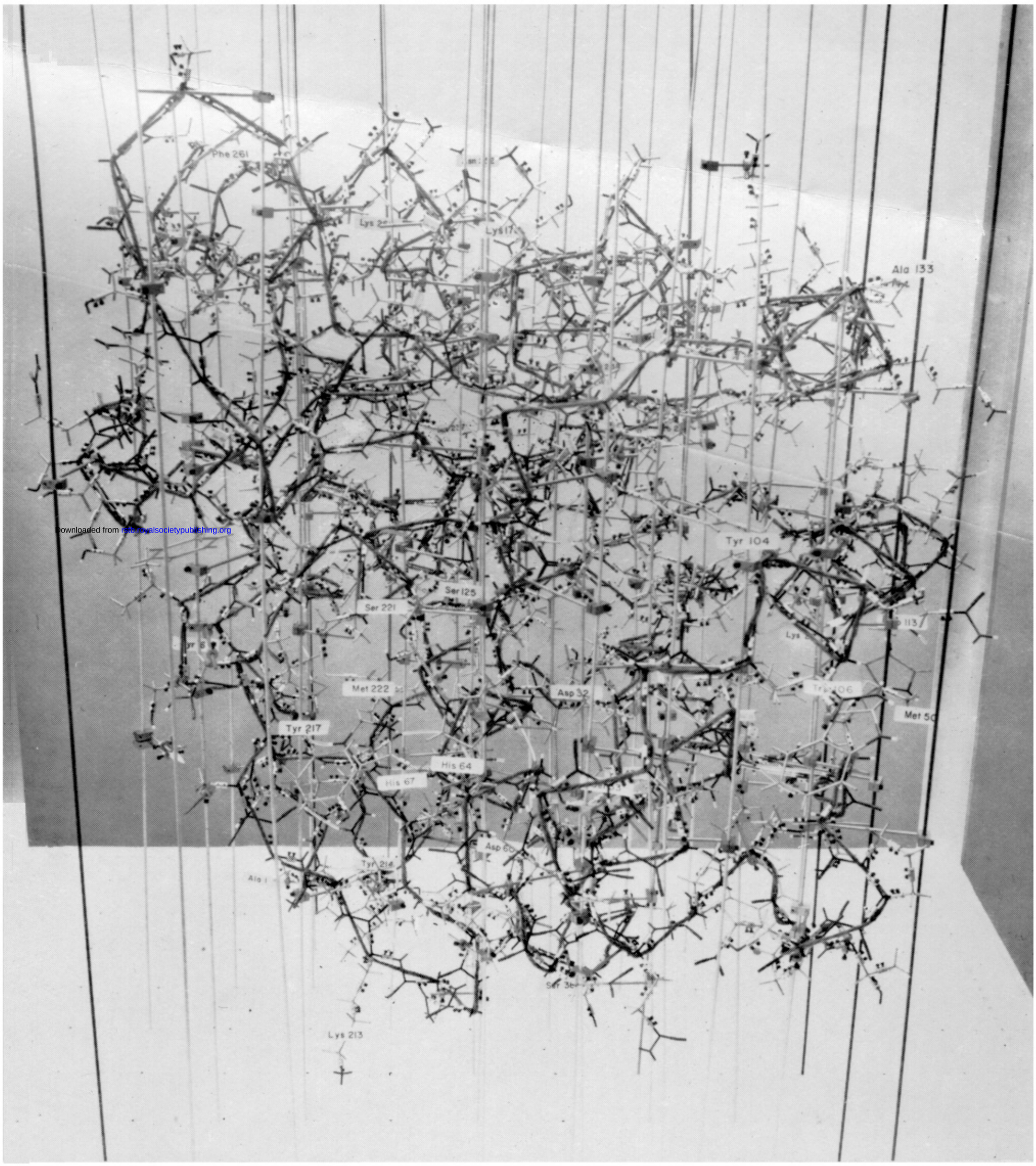


FIGURE 2. (*a*) Stereoscopic view of subtilisin BPN′ skeletal model, looking directly toward the active site. Viewing direction is approximately perpendicular to the imidazole ring of His-64 in the active conformation. The same viewing direction is maintained in figures 3 to 5 as well. Knitting yarn follows the polypeptide backbone chain. (*b*) This is a larger photograph of the model.

(*Facing p.* 122)

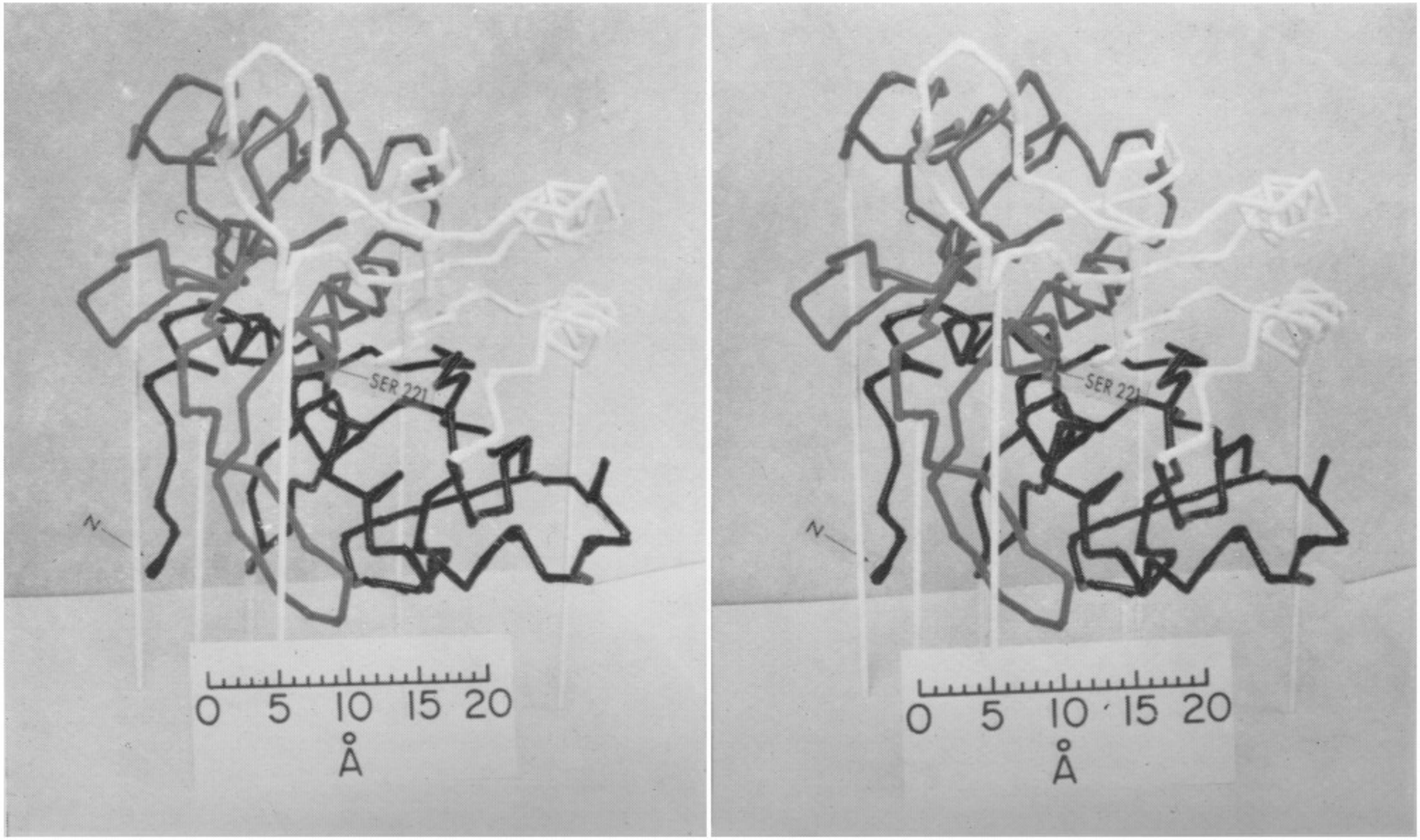


FIGURE 3. Backbone chain model, showing folding of single continuous polypeptide chain into three potentially separable pieces, at the conjunction of which is the active site.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

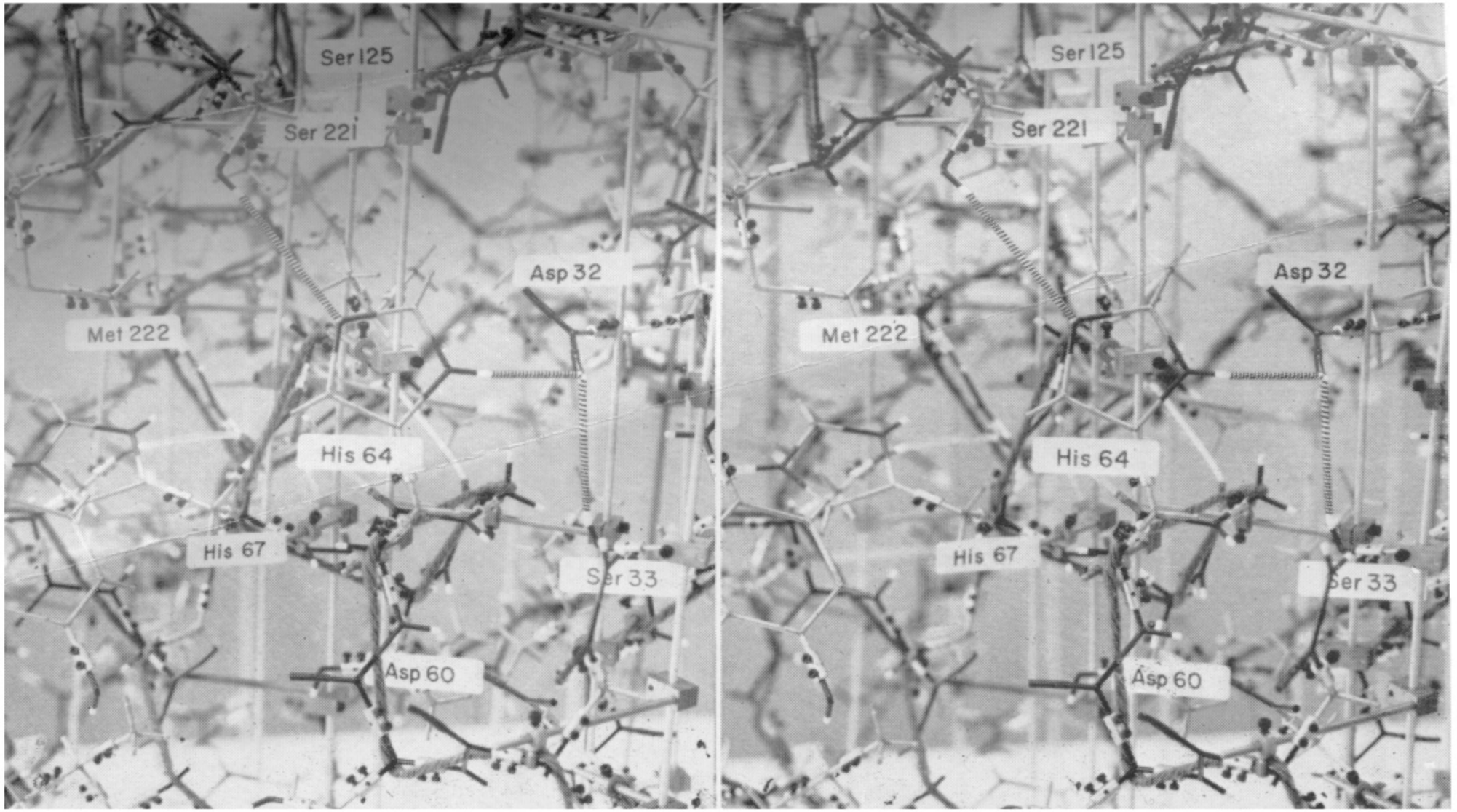


FIGURE 4. Close-up of the active site region as it appears in the active enzyme. Hydrogen bonds that are analogous to those found in α-chymotrypsin are emphasized by insertion of dotted connectors.

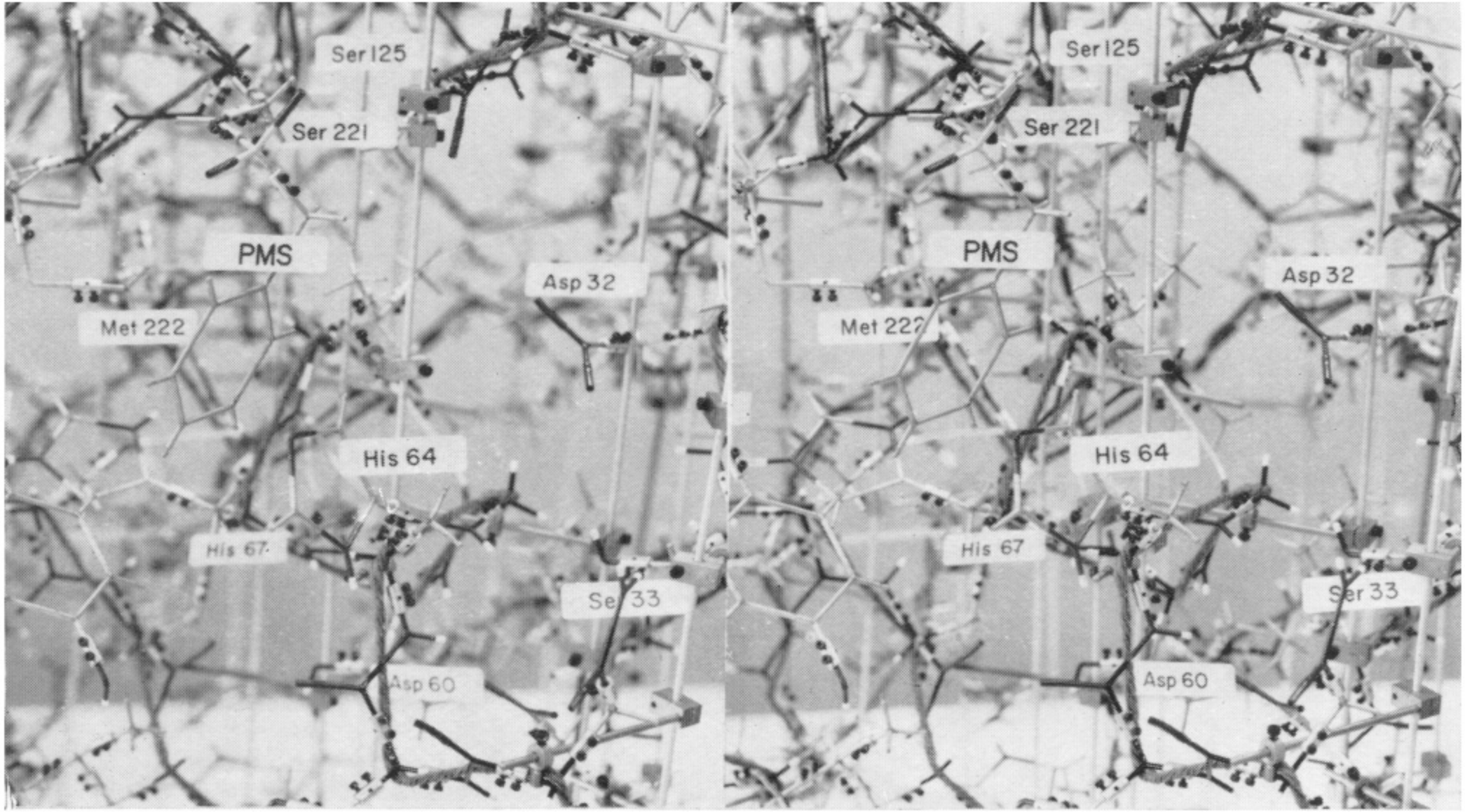


FIGURE 5. Close-up of active site region with phenylmethanesulphonyl inhibitor group covalently bonded to Ser-221. Note movement of His-64 side chain.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B