

## Crystallographic Study of Turkey Egg-white Lysozyme and its Complex with a Disaccharide

R. SARMA† AND R. BOTT‡

*Department of Biochemistry  
State University of New York at Stony Brook  
N.Y. 11794, U.S.A.*

*(Received 1 February 1977)*

The crystal structure of turkey egg-white lysozyme, determined by the molecular replacement method at 5 Å resolution (Bott & Sarma, 1976) has now been refined to 2.8 Å resolution and a model has been built to fit the electron density. A comparison of the co-ordinates with those of hen lysozyme indicate a root-mean-square deviation of 1.6 Å for all the main-chain and side-chain atoms. A significant difference is observed in the region of residues 98 to 115 of the structure. The molecules are packed in this crystal form with the entire length of the active cleft positioned in the vicinity of the crystallographic 6-fold axis and is not blocked by neighboring molecules. A difference electron density map calculated between crystals of turkey lysozyme soaked in a disaccharide of *N*-acetyl glucosamine—*N*-acetyl muramic acid and the native crystals showed a strong positive peak at subsite C, a weak positive peak at subsite D and two strong peaks that correspond to the subsite E and a new subsite F'. This new site F' is different from the subsite F predicted for the sixth saccharide from model building in hen lysozyme. The interactions between the saccharides bound at subsites E and F' and the enzyme molecules are discussed.

### 1. Introduction

The structure of hen egg-white lysozyme in the tetragonal form was solved more than a decade ago. (Blake *et al.*, 1965, 1967*a*). A detailed description of the molecule and several mono-, di- and trisaccharide inhibitors in the tetragonal crystals have led to a complete elucidation of the specificity of the enzyme and a possible mechanism of hydrolysis of an oligosaccharide by lysozyme (Blake *et al.*, 1967*b*; Phillips 1966, 1967). The mechanism is based on the crystallographically observed binding of a trisaccharide at three subsites, termed A, B and C along a cleft of the enzyme molecule and an elegant model building of three additional saccharide units at three more subsites, termed D, E and F. The bond cleaved by the enzyme is the glycosidic linkage between the saccharides D and E. Considerable evidence is available for both the proposed mode of binding of the substrate and the bond catalyzed, from studies of lysozyme in solution (Rupley, 1967; Imoto *et al.*, 1972). However, crystallographic evidence of saccharide binding at sites E and F is not available either in the tetragonal form of hen lysozyme or in any other crystal form of lysozyme studied. (Blake & Swan, 1971; Fenna & Matthews, 1975; Kurachi *et al.*, 1976).

† To whom all communications should be sent.

‡ Present address: Department of Biochemistry, University of Wisconsin, Madison, Wisconsin, U.S.A.

We report here the structure determination of a hexagonal form of turkey egg-white lysozyme performed at a resolution of 2.8 Å. The packing of the molecules in the hexagonal unit cell indicates that the entire length of active cleft of the enzyme molecule is in the vicinity of the 6-fold screw axis and is not blocked by any neighboring molecules. We also present results of binding of the disaccharide *N*-acetyl glucosamine—*N*-acetyl muramic acid to the hexagonal crystals of the turkey lysozyme. The disaccharide appears to bind at subsites E and F', a new site close to F. This is the first time such a binding has been observed crystallographically, and preliminary model building suggests a network of hydrogen bonds between the disaccharide and the enzyme molecule.

## 2. Experimental Procedure

In a previous paper we reported the experimental details of purification of turkey lysozyme, crystallization and data collection (Bott & Sarma, 1976). We also reported the techniques used in determining the rotational and translational parameters relating the known structure of hen egg-white lysozyme to the unknown structure of turkey lysozyme. The rotational parameters were determined using the rotation function of Rossmann & Blow (1962). The translational parameters relating the local origin of the properly rotated known lysozyme molecule to the symmetry elements of the hexagonal unit cell were determined by a packing analysis technique that minimized the number of intermolecular contacts between symmetry related molecules. Surprisingly, there were only a small number of allowed translations, resulting in fewer than 5 such contacts closer than 5 Å. An additional criterion of a low *R* factor between the observed and the calculated structure amplitudes helped to resolve the few translations that gave an acceptable number of contacts. The program used to minimize the *R* factor and the number of contacts is described in our earlier report (Bott & Sarma, 1976). The same method was used extensively with more observed data to 2.8 Å resolution to refine the rotational and translational parameters.

The experimental conditions used for soaking the lysozyme crystals in the disaccharide are also described in our earlier paper.

## 3. Results

### (a) *Refinement of the rotational and translational parameters*

In all stages of the structure analysis of turkey lysozyme, the progress of the refinement of the rotational and translational parameters was measured by the minimization of two quantities: (1) the number of close contacts between the alpha carbon atoms of each molecule and its neighbors and (2) the conventional *R* factor between the observed diffraction pattern and the pattern calculated using the co-ordinates of the properly rotated and translated hen lysozyme molecule.

The rotational and translational parameters described earlier (Bott & Sarma, 1976) resulted in seven intermolecular contacts between the  $\alpha$  carbon atoms of neighboring molecules less than 5 Å and an *R* factor of 46.7% for reflections lying between 10 and 6 Å and 47.1% for data between 10 and 4 Å.

Before extending the resolution of the structure analysis to 2.8 Å, the rotational and translational parameters were varied by small increments around their previous values to determine the set of parameters giving a minimum residual value for reflections between 10 and 2.8 Å with  $F_{\text{obs}} \geq 2\sigma$  (in  $F_{\text{obs}}$ ),  $\sigma$  being the standard deviation determined from measurement statistics. The translational parameters were incremented in intervals of 0.01 of a cell edge, which corresponds to approximately 0.75 Å in each direction. This resulted in a new orientation matrix and a set of translational parameters with 45.8% *R* value, for data between 10 and 4 Å and

46.7%, for data between 10 and 2.8 Å. There were five intermolecular contacts between the  $\alpha$  carbon atoms closer than 4 Å for these parameters. There were two severe short contacts between the  $\alpha$  carbon atoms of residues 113 and 122 and the same residues in a neighboring molecule.

(b) *Electron density map at 2.8 Å resolution*

A preliminary electron density map was calculated with the observed structure amplitudes of turkey lysozyme and phases calculated with the main-chain atoms and the beta carbon of the hen lysozyme molecule rotated and translated using the parameters of the final refinement described in the previous section. In view of the short contacts observed for residues 113 and 122, these were completely omitted from this phase calculation. This electron density map showed density corresponding to several of the side chains that had not been included in the phase calculation, including two of the four —S—S— bridges. Those side chains whose electron density coincided with the corresponding side-chain co-ordinates of hen lysozyme were included in a second phase calculation. These new phases are used in a second 2.8 Å electron density map, two typical sections of which are illustrated in Figure 1(a) and (b).

Each of these is a stack of seven *YZ* sections of the electron density map. The dark contours in Figure 1 (a) and (b) correspond to the native electron density map and the grey contours correspond to the positive regions of a difference electron density map with  $(|F_o| - |F_c|)$  and calculated phase angles.

Figure 1(a) shows the electron density for the main-chain atoms of residue 122 (arrow - 1), a residue that had been omitted from the phase calculation. The Figure also shows the —S—S— bridge between residues 6 and 127 (arrow-2). In Figure 1(b) may be seen electron density corresponding to side chains of leucine 56, leucine 75, isoleucine 78 and proline 79. Electron density for a second —S—S— bridge omitted in the phase calculation may also be seen in Figure 1(b) (arrow-3) as grey contours. A skeletal model has been used to measure co-ordinates of 878 of the 998 non-hydrogen atoms identified in the electron density map. These co-ordinates were idealized using a program written by N. W. Isaacs (Dodson *et al.*, 1976) and were used to calculate a new set of structure amplitudes and phases, and this gives an *R* value of 45.2% for data between 10 and 2.8 Å. There remain four intermolecular contacts less than 4 Å between the alpha carbon atoms of one molecule and its neighbors for this final model; these are given in Table 1.

TABLE I

*Distance between carbon atoms closer than 4 Å for the final idealized model co-ordinates of turkey lysozyme*

| Residue | Equivalent point   | Residue | Equivalent point         | Distance (Å) |
|---------|--------------------|---------|--------------------------|--------------|
| Glu7    | ( <i>x, y, z</i> ) | Ser85   | ( $-y, -x, 5/6 - z$ )    | 3.78         |
| Ser85   | ( <i>x, y, z</i> ) | Glu7    | ( $-y, -x, 5/6 - z$ )    | 3.78         |
| Asn113  | ( <i>x, y, z</i> ) | Asn113  | ( <i>y, x, 1/3 - z</i> ) | 2.76         |
| Arg114  | ( <i>x, y, z</i> ) | Arg114  | ( <i>y, z, 1/3 - z</i> ) | 3.21         |

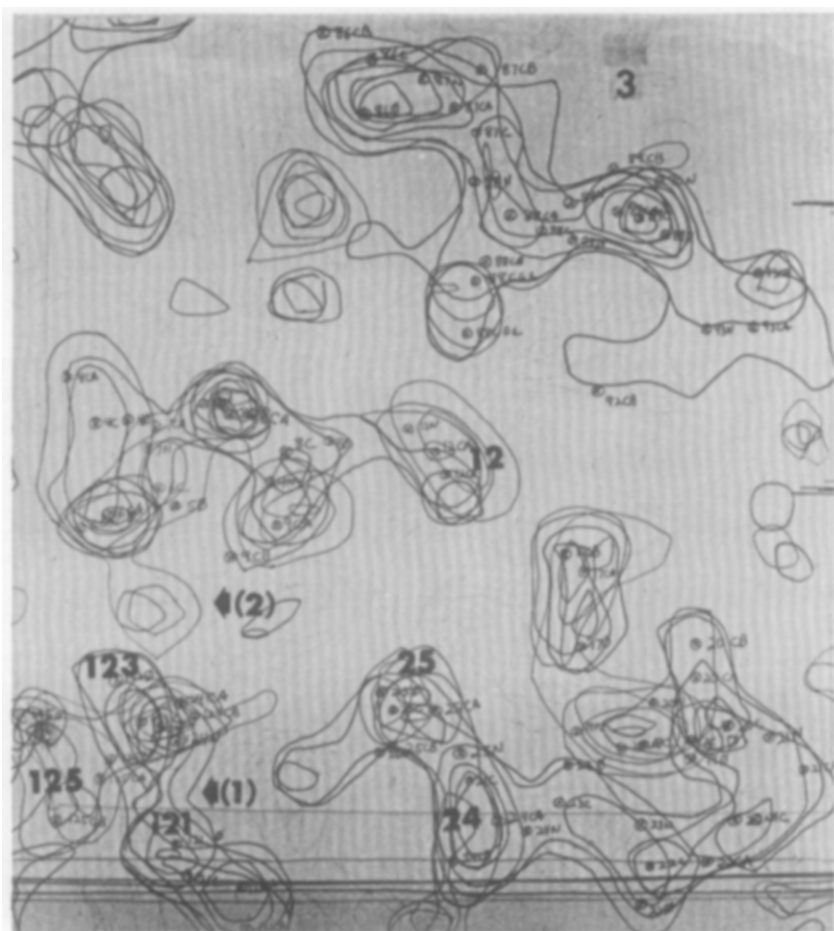


FIG. 1(a).

(c) *Molecular packing of turkey lysozyme in the hexagonal cell*

The above idealized co-ordinates have been used to generate the drawing in Figure 2 showing the packing of 6 of the 12 molecules in the hexagonal unit cell related by the 6-fold screw axis. Only the main-chain atoms are shown in this Figure. The results we reported from the low-resolution studies of turkey lysozyme (Bott & Sarma, 1976) may be clearly seen in this Figure, namely the orientation of the active cleft of the enzyme molecule in the vicinity of the 6-fold axis. The alpha carbon atoms of some of these residues involved in the binding of the substrate are marked in this Figure. The positions of some of the dyad axes present in this space group show that the active cleft in this form is not blocked by any neighboring molecule.

(d) *Binding of the disaccharide N-acetyl glucosamine—N-acetyl muramic acid to turkey lysozyme crystals*

A difference electron density map was calculated using the phase angles derived from the above model of turkey lysozyme and the difference in the observed structure amplitudes between the crystals of turkey lysozyme soaked in the disaccharide N-acetyl glucosamine—N-acetyl muramic acid and the native crystal. Figure 3(a)

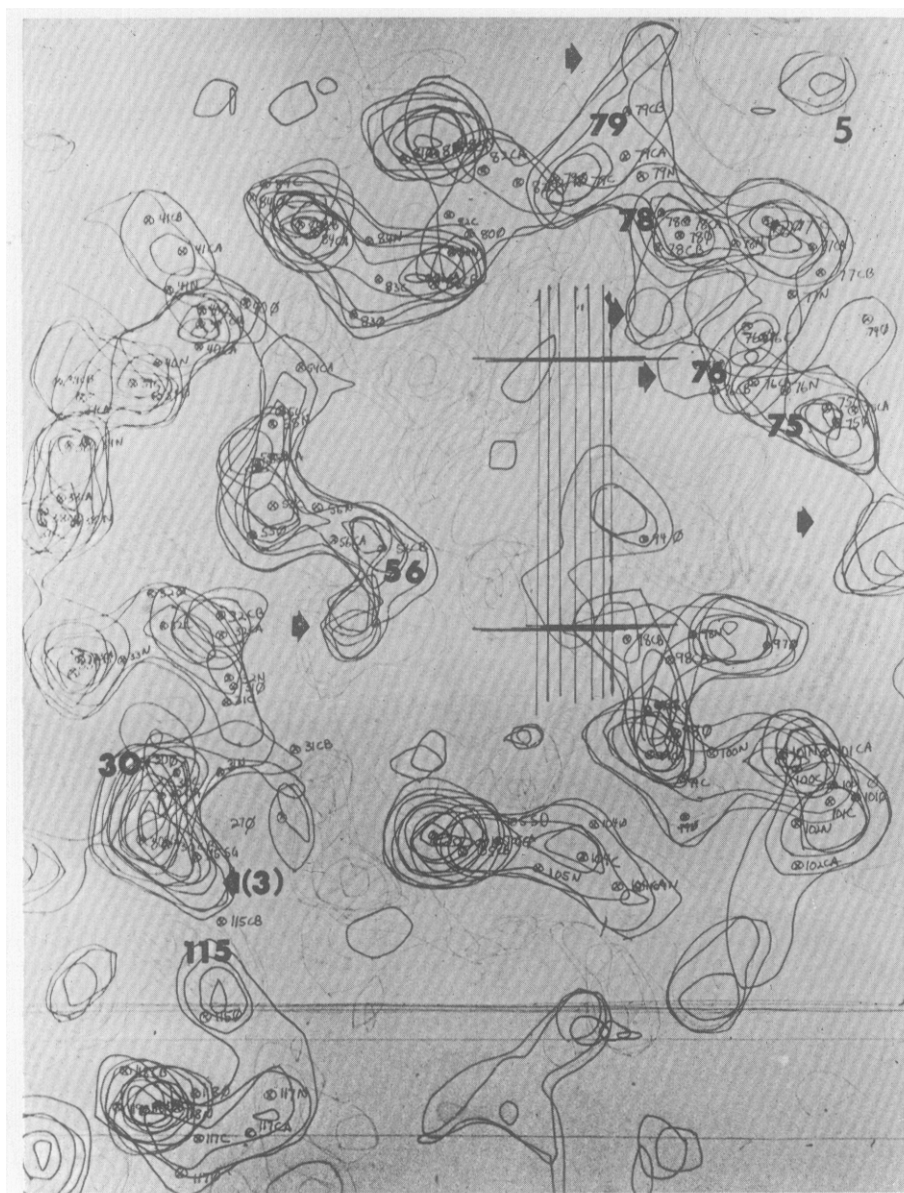


Fig. 1(b).

FIG. 1. A composite of 7 sections each of the 2.8 Å electron density map. Sections are in the YZ plane for values of  $x = 0.39$  to  $0.45$  and  $0.53$  to  $0.59$ .

and (b) shows a few sections of this difference Fourier in the region of the active cleft. The root-mean-square error in the difference map was calculated to be equal to  $0.08 \text{ e}/\text{\AA}^3$ , calculated using the expression given by Ford *et al.* (1974). The difference Fourier was therefore contoured at intervals of  $0.08 \text{ e}/\text{\AA}^3$ . The solid lines correspond to positive values of the difference density and the dotted contours indicate negative values of the difference map. In this Figure A,B,C,D,E and F denote the six subsites on the enzyme molecule thought to be involved in the binding of a hexasaccharide

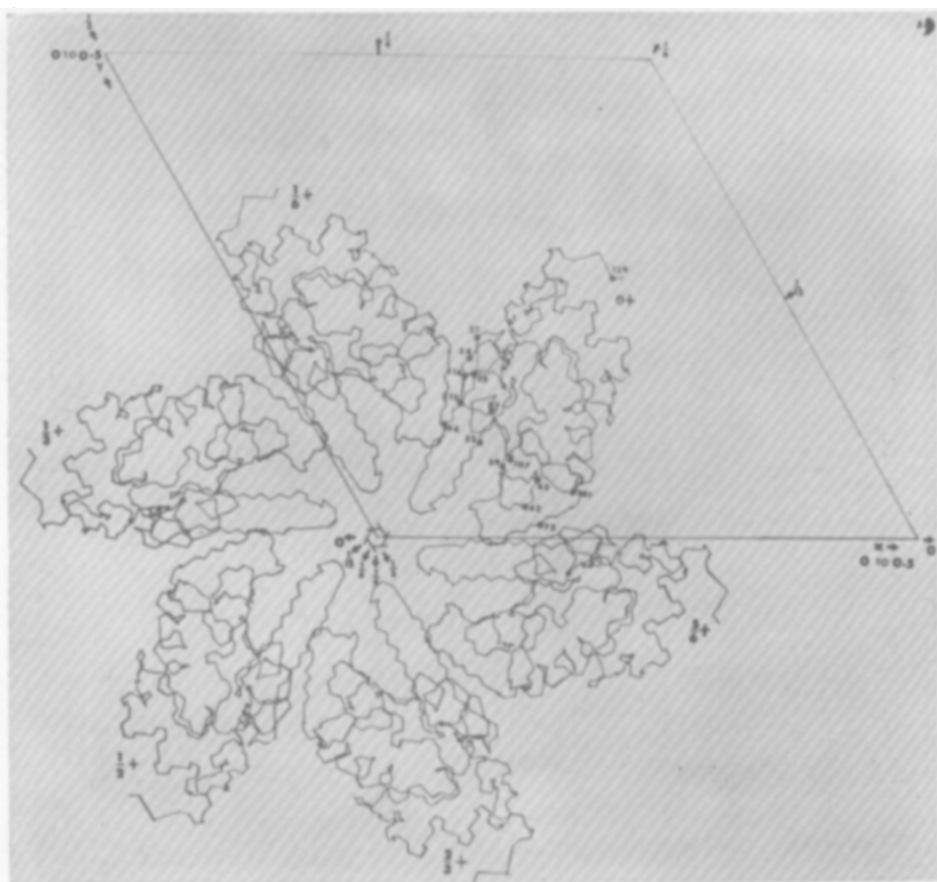


FIG. 2. A computer-generated drawing showing the packing of 6 of the 12 molecules of turkey lysozyme in the hexagonal crystal form. These 6 molecules are related by the 6-fold screw axis parallel to the  $z$  axis. Some of the carbon atoms of residues involved in binding of a substrate are marked on one of the 6 molecules. The positions of some of the dyad axes are also marked in the Figure.

substrate. It may be seen that there is no significant positive density at the A and B subsites, a large peak at the C subsite, a small peak at the D subsite and two clusters of large peaks at sites marked E and F'. There is no significant density at the F subsite. Instead, there appears a second peak connected to the saccharide bound at subsite E and marked F' in Figure 3(b). Some of the atoms of the enzyme molecule in this region are also marked on the difference map. A model of the disaccharide was built in the region of the difference density marked E and F'. Figure 4 shows a drawing of some of the interactions between residues in the enzyme molecule and the saccharide units at E and F'.

#### 4. Discussion

The decrease in the  $R$  factor observed when the co-ordinates of turkey lysozyme measured from the model are used in a structure factor calculation, instead of the transformed hen lysozyme co-ordinates is a clear indication of the correct assignment of the phase angles. In addition, the new features observed in the native electron

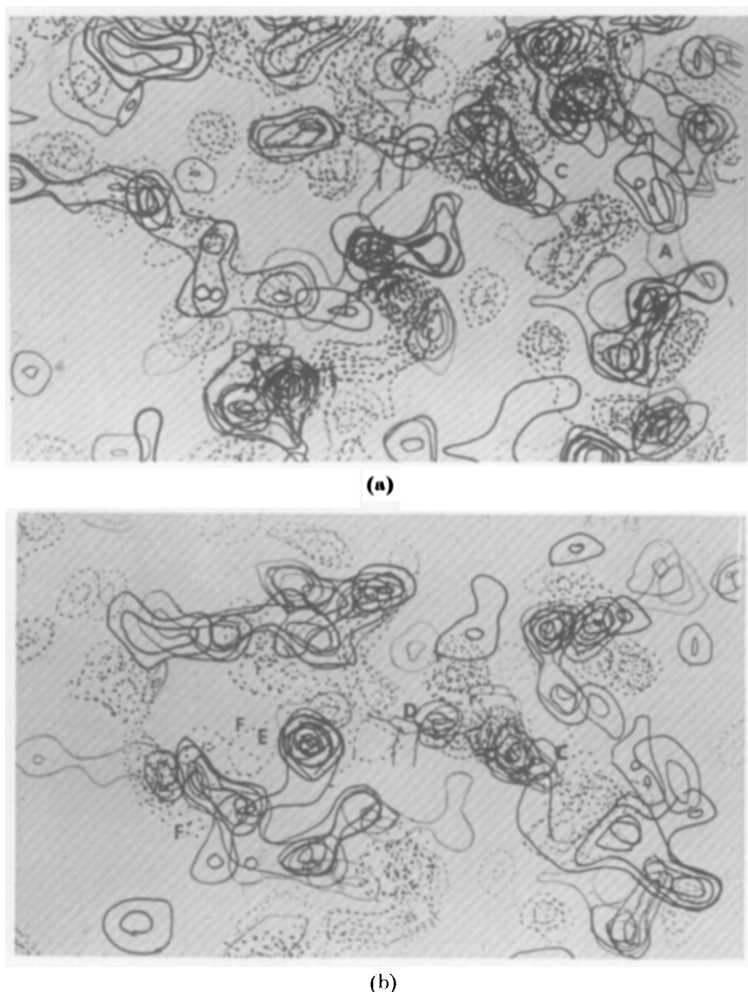


FIG. 3. A composite of 5 sections each of the 2.8 Å difference electron density map between crystals of turkey lysozyme soaked in a disaccharide and the native crystals.

density map corresponding to residues not included in the original phase calculation is an adequate proof of the correctness of the phase assignment. Proofs such as these are crucial in the present studies, because the phase angles, and hence the structure itself, have been derived using a known structure of an homologous enzyme. Another satisfactory proof used in this respect included the interpretation of a difference Fourier of a platinum derivative of turkey lysozyme using the derived phase angles and comparing the positions of the bound platinum derived from this map with the interatomic vectors seen in the difference Patterson (Bott, 1977). The difference Patterson is totally unaffected by the phase assignment, and therefore is capable of providing an independent determination of the position of the heavy metal substitution.

A comparison of the co-ordinates measured from the model of turkey lysozyme with those of hen lysozyme gives a root-mean-square deviation of about 1.0 Å for only the main-chain atoms, 3.0 Å for only the side-chain atoms and 1.65 Å for all atoms in the structure. There are two regions of the molecule that deviate more than

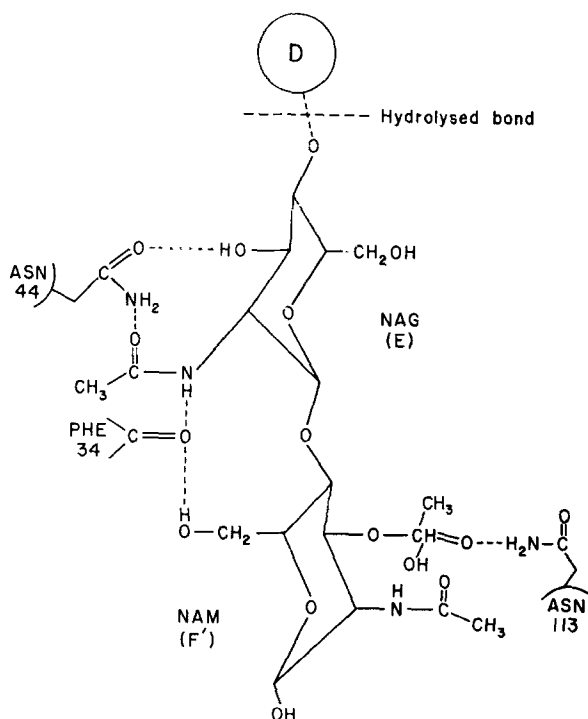


FIG. 4. A schematic drawing of parts of the enzyme molecule interacting with the disaccharide built in subsites E and F'.

NAG, *N*-acetyl glucosamine; NAM, *N*-acetyl muramic acid.

the root-mean-square value given above, these are the residues 97 to 110 and 119 to 125, respectively. These are two regions where three of the seven substitutions occur between hen and turkey lysozymes; 99 (Val to Ala), 101 (Asp to Gly) and 121 (Glu to His). A part of the residues in the region 97 to 110 form the active site of the enzyme molecule and the extended loop of residues 100 to 104 project out into solution more in hen lysozyme than in turkey lysozyme. This confirms the observation by Maron *et al.* (1972) that the immunological cross-activity of the hen lysozyme "active-site-directed" antibodies against turkey lysozyme is weaker than its interaction with the anti-hen lysozyme antibodies.

Not all of the observed differences between the two lysozymes are due to genuine structural differences, some of them are perhaps due to errors in the phase angles at the present stage of the structure analysis. Further refinement of the structure will obviously clear this problem.

The most interesting result of this structure analysis is the binding of the disaccharide to the enzyme molecule. Stereochemical constraints imposed by the lactyl side chain on the muramic acid would prevent the binding of an *N*-acetyl muramic acid residue at subsite C (Blake *et al.*, 1967*b*; Phillips, 1966, 1967). There are at least three possible pairs of binding sites for this disaccharide. They are A-B, C-D and E-F.

The binding of the disaccharides *N*-acetyl glucosamine—*N*-acetyl glucosamine and *N*-acetyl glucosamine—*N*-acetyl muramic acid in the tetragonal hen lysozyme crystals (Blake *et al.*, 1967*a*) and the former in the triclinic hen lysozyme crystals (Kurachi *et al.*, 1976) have been reported to be in the region of subsites A, B and C. An examination of Figure 3(a) and (b) indicates several positive and negative features

TABLE 2

*Hydrogen bonding arrangement between a disaccharide in the second half of the cleft and the lysozyme molecule in the predicted model in hen lysozyme and observed in turkey lysozyme*

| Atom of substrate                        | Atoms on the enzyme molecule       |                                |
|------------------------------------------|------------------------------------|--------------------------------|
|                                          | Predicted model in<br>hen lysozyme | Observed in<br>turkey lysozyme |
| (1) O (3) of saccharide at E             | —CO of Gln57                       | —CO of Asn44                   |
| (2) NH of the acetamido group at E       | —CO of Glu35                       | —CO of Phe34                   |
| (3) CO of the acetamido group at E       | —NH <sub>2</sub> of Asn44          | —NH <sub>2</sub> of Asn44      |
| (4) O (6) of saccharide at F             | —CO of Phe34                       | —CO of Phe34                   |
|                                          | —NH <sub>2</sub> of Asn37          | —                              |
| (5) Ring oxygen of saccharide at F       | —NH <sub>2</sub> of Arg114         | —                              |
| (6) O (1) of saccharide at F             | —NH <sub>2</sub> of Arg114         | —                              |
| (7) CO of Lactyl side chain of NAM at F' | —                                  | —NH <sub>2</sub> of Asn113     |

both at the atomic positions of the enzyme molecule itself and away from them. The former corresponds to conformational changes produced in the substituted crystals and the latter most probably is due to the added molecule itself. There is practically no density at sites corresponding to the A and B subsites. This indicates that the binding, if any, of the glucosamine–muramic acid disaccharide at these sites is very weak indeed. This is to be expected in view of the major substitution of a Gly for Asp (101), since the latter participates in two hydrogen bonds with the saccharides at sites A and B.

There is however a strong peak observed for the saccharide binding at subsite C, apparently making all the four hydrogen bonds, as in the case of hen lysozyme (Phillips, 1967). The observed peak for the saccharide binding at subsite D is not as strong as the peak for subsite C, an equally strong negative peak appears close to the positive peak expected for the saccharide binding at D. A similar weak electron density is reported to have been observed for the sugar in site D from the binding studies of tetra-*N*-acetyl glucosamine to the tetragonal hen lysozyme crystals (Grace, Johnson, Patterson & Phillips, unpublished data, as reported by Ford *et al.*, 1974). We have interpreted the weak density in our difference map to be indicative of the weak binding at this subsite due to the unfavorable interactions at subsite D. It may also be due to a replacement of a chlorine ion or the solvation shell near Asp52 by the saccharide at the D subsite. Binding at subsite D has been observed in the tetragonal hen lysozyme crystals for the lactone moiety of a lactone derivative of tetra-*N*-acetyl glucosamine (Ford *et al.*, 1974) and for a tri-*N*-acetyl-glucosamine derivative to lysozyme inactivated by iodine (Beddell *et al.*, 1975). The orientations in site D of the lactone group in the former case and the glucopyranose ring in the latter case have been described in the above investigations. The density observed in the present difference map is not accurate enough to describe the conformation of the sugar in subsite D.

There are two strong peaks in the cleft beyond the subsite D. These have been interpreted to be due to the binding of the disaccharide to the lower half of the active cleft. The first of these binding sites is close to the E subsite predicted for the hexa-saccharide in hen lysozyme. As in the predicted model this saccharide is in a position

to make three hydrogen bonds with the enzyme molecule. However, some of the residues involved in these hydrogen bonds are different from those of the predicted model as seen from Figure 4 and Table 2. Instead of Gln57 in the predicted model, the peak for the sixth saccharide is closer to Phe34.

There is however no peak corresponding to the saccharide at the subsite termed F in the model. Instead, there is a strong peak at a different site termed F' in our difference map. This has been interpreted to be due to the *N*-acetyl muramic acid linked to the *N*-acetyl glucosamine as subsite E. Unlike the sugar at subsite F, the sugar at subsite F' can make only two hydrogen bonds with the enzyme molecule. One of these is between the carboxyl of the lactyl side group of the *N*-acetyl muramic acid and the —NH<sub>2</sub> group of Asn113. The position of the subsite F' is about 5 Å away from the predicted subsite at F.

A second observation worth noting is while the distance between covalently bonded saccharides is approximately 5 Å, the distance between the positive density corresponding to the sugars bound at subsites D and E is about 6.5 Å. This is longer than the distance between covalently linked saccharides. These peaks are densities of saccharides that are not covalently attached. However, if this binding represents the enzyme-substrate interaction in solution, this result suggests that the enzyme might exert some strain on the leaving group by pulling it away from the site of hydrolysis.

The packing of the turkey lysozyme molecules in the hexagonal cell (Fig. 2) seems to indicate that in this form the active cleft is not blocked by any neighboring molecule. This is confirmed by the binding of the disaccharide at sites below the subsite D. Experiments are already in progress to attempt to bind a longer saccharide to this crystal form. To reduce the enzymatic catalysis of a longer saccharide, diffraction experiments will be performed at low temperatures using antifreeze buffers (Alber *et al.*, 1976).

The set of phases that have been used in calculating the present difference Fourier is not accurate enough to enable precise measurement of the co-ordinates of the two saccharides at E and F'. We propose to refine the co-ordinates of the turkey lysozyme model in order to obtain a better set of phase angles to determine more accurately the mode of binding of saccharides to the unexplored regions of the active cleft of lysozyme.

We are grateful to Dr Melvin Schindler of the Weizmann Institute of Science, Rehovot, Israel for samples of *N*-acetyl glucosamine—*N*-acetyl muramic acid, and to Dr S. T. Rao of the University of Wisconsin at Madison and to Dr N. W. Isaacs for some of the computer programs used in this work. This work was supported by United States Public Health Service grant AI-10762 from the National Institutes of Health.

#### REFERENCES

- Alber, T., Petsko, G. A. & Tsernoglou, D. (1976). *Nature (London)*, **263**, 297–300.  
Beddell, C. R., Blake, C. C. F. & Oatley, S. J. (1975). *J. Mol. Biol.* **97**, 643–654.  
Blake, C. C. F. & Swan, I. D. (1971). *Nature New Biol.* **232**, 12–15.  
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, R. (1965). *Nature (London)*, **206**, 757–761.  
Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, R. (1967a). *Proc. Roy. Soc. ser. B*, **167**, 365–377.  
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, R. (1967b). *Proc. Roy. Soc. ser. B*, **167**, 378–388.

- Bott, R. (1977). Ph.D. Thesis, State University of New York at Stony Brook.
- Bott, R. & Sarma, R. (1976). *J. Mol. Biol.* **106**, 1037-1046.
- Dodson, E. J., Isaacs, N. W. & Rollett, J. S. (1976). *Acta Crystallogr. sect. A*, **32**, 311-315.
- Fenna, R. E. & Matthews, B. W. (1975). *Nature (London)*, **258**, 573.
- Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tjian, R. (1974). *J. Mol. Biol.* **88**, 349-371.
- Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972). In *The Enzymes* (Boyer, P. D., ed.), vol. 7, pp. 665-868, Academic Press, New York.
- Kurachi, K., Sieker, L. C. & Jensen, L. H. (1976). *J. Mol. Biol.* **101**, 11-24.
- Maron, E., Eshdat, Y. & Sharon, N. (1972). *Biochim. Biophys. Acta*, **278**, 243-249.
- Phillips, D. C. (1966). *Sci. Amer.* **215**, 78-90.
- Phillips, D. C. (1967). *Proc. Nat. Acad. Sci., U.S.A.* **57**, 484-495.
- Rossmann, M. G. & Blow, D. M. (1962). *Acta Crystallogr.* **15**, 24-31.
- Rupley, J. A. (1967). *Proc. Roy. Soc. ser. B*, **167**, 416-428.