Proc. Nat. Acad. Sci. USA
Vol. 71, No. 10, pp. 4178–4182, October 1974

The Three Dimensional Structure of the Lysozyme from Bacteriophage T4

(protein structure/x-ray diffraction/mutants)

B. W. MATTHEWS AND S. J. REMINGTON

Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oreg. 97403

Communicated by Virgil Boekelheide, August 1, 1974

ABSTRACT The three dimensional structure of the lysozyme from bacteriophage T4 has been determined from a 2.5 Å resolution electron density map. About 60% of the molecule is in a helical conformation and there is one region consisting of antiparallel β-structure. The polypeptide backbone folds into two distinct lobes linked in part by a long helix. In the region between the two lobes, there is a cleft which deepens into a hole or cavity, about 6–8 Å in diameter, extending from one side of the molecule to the other. This opening is closed off by side chains which extend to within 3–5 Å of each other. A number of mutant lysozymes in which residues in the vicinity of the opening are modified have markedly reduced catalytic activity, suggesting that this region of the molecule may be catalytically important. The three dimensional structure of T4 phage lysozyme is quite different from that of hen egg-white lysozyme although it is not clear at this time whether or not the mechanisms of catalysis of the respective enzymes are related.

T4 phage lysozyme is an enzyme produced in cells of *Escherichia coli* after infection with bacteriophage T4. The enzyme has similar catalytic activity to that of hen egg-white lysozyme, both being endoacetylmuramidases (1). The molecular weight of the enzyme is 18,700, and the amino-acid sequence has been determined (2). Furthermore, a number of lysozymes have been isolated from mutant strains carrying frame-shift or amber mutations in the phage genome and have been used to demonstrate *in vivo* certain features of the genetic code (3–5).

In this preliminary report, we describe the three dimensional structure of T4 phage lysozyme as determined by x-ray crystallography from a 2.5 Å resolution electron density map. Additional details will be given in a subsequent publication.

Crystallization

The protein was purified using essentially the method of Tsugita *et al.* (1) except that 1 mM mercaptoethanol was added to all buffers. Conditions for obtaining the crystals have been described previously (6). Before x-ray photography the crystals were equilibrated with a standard mother liquor consisting of 1.05 M K_2HPO_4, 1.26 M NaH_2PO_4, 0.23 M NaCl, 1.4 mM mercaptoethanol, pH 6.7. In this solution the crystals float.

The crystals have space group $P3_221$ with cell dimensions $a = b = 61.1$ Å, $c = 96.3$ Å and one molecule in each of the six asymmetric units.

Data collection

Diffraction data were recorded photographically using conventional Buerger precession cameras and integrated intensities measured with a computer controlled drum film scanner (7). Sets of 12 films sufficed to measure 92% of the data to 2.5 Å and 53% of the data between 2.4 Å and 2.5 Å. With the films used, the Bijvoet differences were obtained for about half the reflections. Systematic errors in the measurement of the Friedel pairs were reduced by a method of local scaling to be described elsewhere (ref 8, B. W. Matthews *et al.*, manuscript in preparation). Altogether, about 27,000 intensities were measured for the parent crystals and the two heavy atom derivatives, and were reduced to about 7600 unique reflections. Excluding the weak reflections, the agreement between structure amplitudes measured on different films was about 4%.

Heavy atom derivatives

Isomorphous heavy atom derivatives were sought in the usual way by soaking the crystals in a variety of reagents containing heavy atoms. In particular, efforts were made to exploit potentially available groups on the protein, and two such attempts, utilizing the cysteine and the methione residues, proved successful.

Although exposure of the crystals to mercurials caused cracking, it was determined that diffraction data could be collected from crystals soaked in solutions of mercuric ion. As a result of a series of experiments, the conditions finally adopted were to soak the crystals for 2 days in 2.5×10^{-4} M $HgCl_2$ in standard mother liquor, in the absence of mercaptoethanol. After this period, a solution 1.0×10^{-6} M in mercuric ion was substituted. Under these conditions, both cysteines appear to have been fully substituted with mercury.

In the case of methione, it was possible to obtain substitution of three of the five residues in the molecule by soaking the crystals for 5 days in 3.3×10^{-4} M K_2PtCl_4 in standard mother liquor, without mercaptoethanol.

The sites of heavy atom substitution were first determined from difference Patterson syntheses and confirmed by difference Fourier methods and parameter refinement (9) in the usual way. In these preliminary tests, the centrosymmetric (*h0l*) zone, which in space group $P3_221$ gives the x, y, and z coordinates of each atom, was used exclusively, including data, up to a resolution of 2.5 Å. Final refinement of the heavy atom parameters was by a "lack of closure" procedure (10) including all the data, using a program written by Dr. Lynn Ten Eyck. The refined parameters are given in Table 1.

The space group was shown by two methods similar to those used for thermolysin (11, 12) to be $P3_221$ rather than $P3_121$.

Phase angles were calculated using the method of Blow and Crick (13), including the anomalous scattering data (14). The overall figure of merit was 0.57 and the average value of

TABLE 1. *Refinement statistics for T4 lysozyme*

Derivative	Reaction site	Z	x	y	z	B	$\langle f_H \rangle$	E	R
$HgCl_2$	Cys 54	104	0.551	0.660	0.025	25	201	103	0.48
	Cys 97	88	0.572	0.965	0.083	38			
K_2PtCl_4	Met 106	127	0.680	0.340	0.034	26	267	126	0.44
	Met 120	139	0.127	0.453	0.114	84			
	Met 1	109	0.711	0.953	0.069	49			

Z is the occupancy of the heavy atom site; x, y, z the fractional coordinates; B, the isotropic thermal parameter in A^2; $\langle f_H \rangle$, the root mean square heavy atom scattering; E, the root mean square lack of closure error, and R, the reliability index for centrosymmetric reflections. Z, $\langle f_H \rangle$, and E are on the same arbitrary scale.

[lack of closure/E (or E')] at the most probable phase was 0.66, indicating that the errors in the data were somewhat less than had been estimated (15).

Electron density map

After tracing the electron density map onto sheets of transparent film, a number of helices were immediately apparent and the direction of the polypeptide chain within such helices could be ascertained from the characteristic forward thrust of the carbonyl oxygens and backward tilt of the side chains. On consideration of the two mercury binding sites as possible locations for the respective cysteines, it became apparent that a length of helix running from site Hg 1 must correspond to the sequence Cys 97, Ala 98, . . . Strong density, practically the highest in the map, could be seen at a position compatible with Met 102, apparently in an interior region, while further along the helix Met 106 was located at a position such that it accounted for the major site of chloroplatinite binding. From this initial interpretation, and by making use of the amino-acid sequence, it was not difficult to trace the entire course of the polypeptide chain. There are a few regions where a detailed interpretation is difficult, notably including residues 50–56, which are at an extremity of the molecule and include Cys 54, the site of reaction of Hg 2. Also, the density of the two residues at the carboxyl terminus is very weak, indicating free motion of these residues in the crystal, and consistent with the observation that they can be removed by carboxypeptidase A with no effect on the enzymatic activity (5). Nevertheless, the overall conformation is unambiguous and is confirmed by the coincidence of three of the five methionines with the chloroplatinite binding sites (the other two methionines are internal) and the two cysteines with the two mercury binding sites (Table 1). In addition, characteristic electron density for many of the side chains can be identified, and there is no obvious incompatibility between the electron density map and the amino acid sequence as determined by chemical methods (2).

In the initial interpretation of the electron density map, markers were placed at each alpha-carbon position. Subsequently, a Kendrew model was constructed using a modified Richards optical comparator (16), in which the electron density sections were mounted parallel to the mirror (12).

Description of the structure

The molecule of T4 phage lysozyme has overall dimensions about 50 × 30 × 30 Å and the general arrangement of the polypeptide backbone is illustrated in Figs. 1 and 2.

As is clear from the figures, the structure consists of two quite distinct lobes, with the carboxy terminal part of the polypeptide chain lying exclusively in the upper domain, while the lower domain contains most of the amino terminal portion of the molecule, although the amino terminal helix interacts extensively with the upper domain and appears to help link the two lobes together.

In the lower lobe residues, 18–34 form three strands of somewhat distorted antiparallel pleated sheet, this being the only extended β-structure in the molecule. On the other hand, the upper domain is markedly helical in character and appears to be based on a combination of five helices, each of about 10 residues, arranged so that their axes form the walls of a cylinder surrounding a pronounced hydrophobic core. This hydrophobic region connects through the waist region of the molecule with another hydrophobic region in the lower half of the molecule.

The most obvious connection between the two domains is a long helix of 20 residues which extends almost from one extremity of the molecule to the other. In addition, the two lobes are also connected through the amino terminal helix as mentioned above.

Altogether, about 60% of the molecule is comprised of helices, including residues 3–11, 39–49, 60–79, 82–90, 95–106, 115–123, 129–134, 137–141, and 143–155. Adding residues 18–34, which are in antiparallel β-structure, approximately 70% of the protein is involved in more-or-less regular secondary structure. Spectroscopic studies of the enzyme in solution also indicate a high helix content (M. A. Elwell and J. A. Schellman, personal communication).

Extending across the waist region, roughly from left front to right rear in Figs. 1 and 2, there is a depression in the surface of the molecule which, partway along its length, deepens into what can best be described as a hole, or opening, about 6–8 Å in diameter, extending from one side of the molecule through to the other. In Figs. 1 and 2, the direction of view is approximately normal to the "hole" which is located near the center of the figures. In the three dimensional structure, the polypeptide backbone in the vicinity of Thr 21–Glu 22 is about 8 Å from the backbone in the region of Glu 141–Thr 142 and the respective side chains of these residues extend to within 3–5 Å of each other so that the opening is more occluded than appears in Figs. 1 and 2. The guanido group of Arg 145 also extends across the mouth of the opening to within about 4 Å of the carboxyl of Glu 21, although the two groups do not obviously form a salt link.

This opening is perhaps the most unusual feature of the structure, and is without precedent among the proteins whose three dimensional structures have been determined to date. It remains to be determined whether or not this region of the molecule accommodates either the alternating *N*-acetyl-

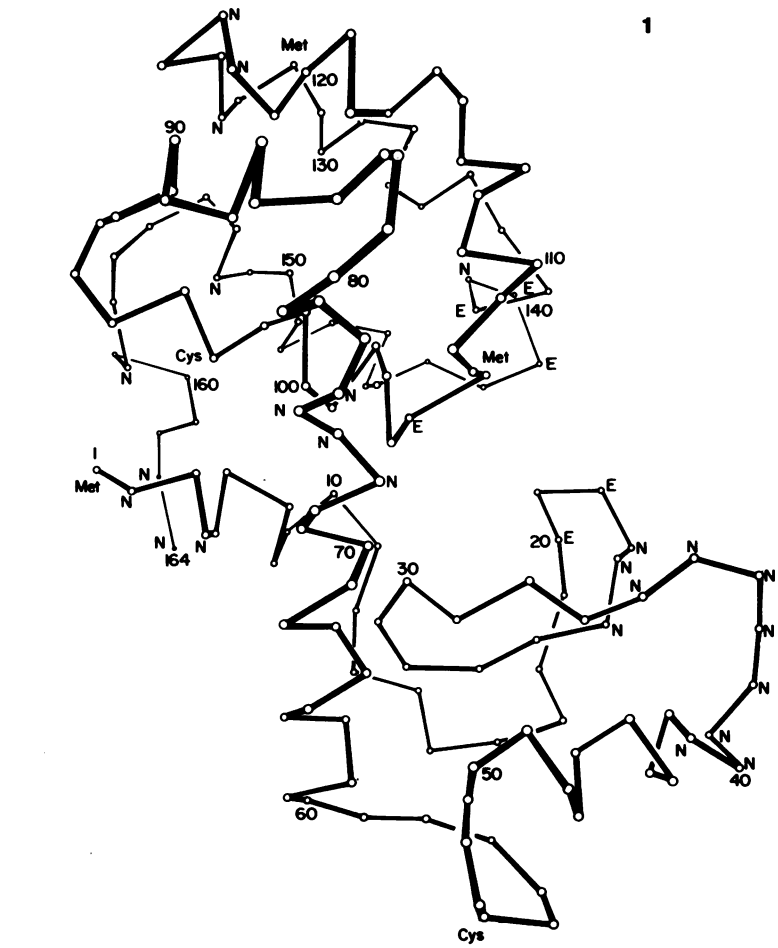

FIG. 1 (top). Perspective drawing illustrating the polypeptide backbone of bacteriophage T4 lysozyme. The approximate position of each alpha carbon atom is indicated by an open circle. Also shown are the methionine and cysteine residues involved in heavy-atom binding. Residues labeled E and N were respectively designated on the basis of genetic studies as "Essential" and "Non-essential" for full catalytic activity, as described in the *text*.

FIG. 2 (bottom). Stereo diagram illustrating the fold of the polypeptide chain in T4 phage lysozyme. The direction of view is close to that in Fig. 1 and the amino terminus is indicated by a small circle.

muramyl-*N*-acetylglucosaminyl portion of the cell wall or the interstrand polypeptide cross-link emanating from each *N*-acetylmuramyl unit (17), but if either of these should prove to be the case, then it seems certain that in order to allow the substrate to enter, the enzyme would have to undergo a fairly substantial conformational change.

Although there is no direct evidence to demonstrate that substrates might bind in the vicinity of the cleft and/or the opening described above, there is indirect evidence from genetic studies, described in the following section, which strongly suggests that this region of the molecule is catalytically important.

Also, there is some chemical evidence to support this assumption. Firstly, reaction of the two sulphydryls with *p*-chloromercuribenzoate (2) does not reduce the enzymatic activity to any appreciable extent, and in the three dimensional structure both cysteines are far from the region of the cleft (Fig. 1). Secondly, the chloroplatinite ion, used to obtain an isomorphous heavy atom derivative, reacts with three methionines, including Met 106 which is located toward one end of the cleft region, and the activity of the $PtCl_4^{2-}$-substituted enzyme is 30% of that of the unsubstituted enzyme (unpublished experiments performed in collaboration with F. W. Dahlquist and A. Y. Maynard), suggesting that the $PtCl_4^{2-}$ group either obstructs the substrate binding site or perturbs the conformation of the enzyme required for optimum catalytic activity.

Mutant lysozymes

The genetics of T4 phage lysozyme have been studied extensively over a number of years by Streisinger, Tsugita and colleagues, and a variety of mutant enzymes have been isolated and characterized (e.g., see refs. 3–5). From the analysis of a series of frame-shaft mutants, Tsugita (5) suggested that the residues Asp 20, Glu 22, Glu 105, Trp 138, Asn 140, and Glu 141 are all essential for full catalytic activity, on the grounds that changes in any of these amino acids drastically reduce the catalytic effectiveness of the enzyme. It is striking to find that each of these residues, indicated by the letter E in Fig. 1, is located in the vicinity of the "hole". In contrast, residues which Tsugita classed as non-essential, denoted N in Fig. 1, are scattered throughout the molecule.

It is also of interest to consider, in the light of the tertiary structure, those mutants in which additional amino acids have been inserted in the polypeptide chain. For example, it is hardly surprising that the replacement of Thr-Glu at positions 21,22 by Lys-Thr-Glu reduces the enzymatic activity almost to zero. One the other hand, it might be expected that the replacement of Ala 74, in the long helix connecting the two lobes of the molecule, by Asp-Val, would have drastic consequences; yet the activity observed for this mutant is almost the same as that for the wild-type enzyme. Obviously, a more detailed study of selected mutants of this type is in order.

Comparison with hen egg-white lysozyme

The lysozymes from T4 bacteriophage and from hen egg-white have similar catalytic activity but non-homologous amino acid sequences (2). It is, therefore, of interest to compare the tertiary structure of the phage enzyme with that of hen egg-white lysozyme determined by Phillips and collaborators (18–20).

Apart from the fact that the respective molecules are both folded into two lobes, there seems to be no similarity whatsoever between their respective three dimensional structures.

It has been suggested by Dunnill (21) that the amino-acid sequences of the phage and hen egg-white lysozymes are correlated and that the two enzymes have an evolutionary relationship, but the present structure determination does not support this hypothesis. For example, using the sequence alignment proposed by Dunnill (21), the respective helices of the two enzymes do not obviously correlate and, in addition, Asp 47 and Glu 64 in T4 phage lysozyme, postulated by Dunnill to be catalytically important, are located away from the presumed substrate binding site and are 20 Å apart.

Comparisons of the structure of α-chymotrypsin with that of subtilisin (22), and carboxypeptidase A with that of thermolysin (12), have shown that enzymes with similar catalytic activity may have quite different tertiary structures, yet have active sites with some elements in common. We have therefore examined the T4 phage lysozyme structure for possible analogues of Glu 35 and Asp 52, the two catalytically important residues in hen egg-white lysozyme (23). Two possible candidates might be Glu 11 and Asp 20, which are located on opposite sides of the opening described above, and have their carboxyl groups about 8 Å apart, comparable with 7 Å between the two carboxyls in the hen egg-white enzyme (20); however, there is an obvious difference in that neither acid group is located in a predominantly nonpolar region, as is Glu 35 in hen egg-white lysozyme. In the phage enzyme, Glu 11 is located in the inner wall of the opening through the molecule and forms a salt link with Arg 145, an interaction for which there is no obvious counterpart in hen egg-white lysozyme. Asp 20 is located in a "hairpin bend" which juts across the mouth of the molecular opening (Fig. 1) and is included in the residues designated as "essential" on the basis of the genetic studies (5).

It may be noted that residue 31, the only histidine in the phage lysozyme, is positioned in the depression between the two domains of the molecule, only a few angstroms from the mouth of the opening. Furthermore, the imidazole of His 31 participates in an ionic interaction with Asp 70 similar to that observed in the serine proteases (24, 25) and thermolysin (26). Whether this potential "charge relay system" plays a catalytic role in T4 lysozyme remains to be determined.

In summary, the three dimensional structures of hen egg-white lysozyme and T4 lysozyme are seen to be quite different but it is not clear at this time whether or not their respective mechanisms of catalysis may be related.

We are particularly grateful to Miss A. Yvonne Maynard for her excellent technical assistance in purifying the phage lysozyme and to Dr. Lynn Ten Eyck for his help in the refinement of the heavy atom parameters. We have benefited from discussions with F. W. Dahlquist, G. Streisinger, J. A. Schellman, and J. (Emrich) Owen, and would also like to thank H. F. Matthews and other colleagues in this Institute for their assistance. This work was supported in part by grants from the National Science Foundation (GB30823X), the National Institutes of Health (GM15423, GM20066, GM00715), and by the award to one of us (B.W.M.) of an Alfred P. Sloan Research Fellowship and a Public Health Service Career Development Award (GM 70585) from the Institute of General Medical Sciences.

1. Tsugita, A., Inouye, M., Terzaghi, E. & Streisinger, G. (1968) *J. Biol. Chem.* **243**, 391–397.
2. Tsugita, A. & Inouye, M. (1968) *J. Mol. Biol.* **37**, 201–212.

3. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.
4. Ocada, Y., Amagase, S. & Tsugita, A. (1970) *J. Mol. Biol.* **54**, 219–246.
5. Tsugita, A. (1971) in *The Enzymes*, ed. Boyer, P. D. (Academic Press, New York), Vol. 5, 3rd ed., pp. 343–411.
6. Matthews, B. W., Dahlquist, F. W. & Maynard, A. Y. (1973) *J. Mol. Biol.* **78**, 575–576.
7. Matthews, B. W., Klopfenstein, C. E. & Colman, P. M. (1972) *J. Phys. E: Sci. Instrum.* **5**, 353–359.
8. Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967) *Nature* **214**, 652–656.
9. Hart, R. G. (1961) *Acta Crystallogr.* **14**, 1194–1195.
10. Dickerson, R. E., Kendrew, J. C. & Strandberg, B. E. (1961) in *Computing Methods and the Phase Problem in X-ray Crystal Analysis* (Pergamon Press, Oxford), pp. 236–251.
11. Matthews, B. W. (1966) *Acta Crystallogr.* **20**, 230–239.
12. Colman, P. M., Jansonius, J. N. & Matthews, B. W. (1972) *J. Mol. Biol.* **70**, 701–724.
13. Blow, D. M. & Crick, F. H. C. (1959) *Acta Crystallogr.* **12**, 794–802.
14. Matthews, B. W. (1966) *Acta Crystallogr.* **20**, 82–86.
15. Matthews, B. W. (1970) in *Crystallographic Computing*, eds. Ahmed, F. R., Hall, S. R. & Huber, C. P. (Munksgaard, Copenhagen), pp. 146–159.
16. Richards, F. M. (1968) *J. Mol. Biol.* **37**, 225.
17. Weidel, W. & Pelzer, H. (1964) *Advan. Enzymol.* **26**, 193–232.
18. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1964) *Nature* **206**, 757–763.
19. Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967) *Proc. Roy. Soc. Sect. B* **167**, 365–377.
20. Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972) in *The Enzymes*, ed. Boyer, P. D. (Academic Press, New York), Vol. 7, 3rd ed., pp. 665–868.
21. Dunnill, P. (1967) *Nature* **215**, 621–622.
22. Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C. & Wilcox, P. E. (1972) *Biochemistry* **11**, 2439–2449.
23. Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967) *Proc. Roy. Soc. Sect. B* **167**, 378–388.
24. Blow, D. M., Birktoft, J. J. & Hartley, B. S. (1969) *Nature* **221**, 337–340.
25. Wright, C. S., Alden, R. A. & Kraut, J. (1969) *Nature* **221**, 235–242.
26. Matthews, B. W., Colman, P. M., Jansonius, J. N., Titani, K., Walsh, K. A. & Neurath, H. (1972) *Nature New Biol.* **238**, 41–43.