Proc. Nat. Acad. Sci. USA
Vol. 72, No. 1, pp. 38–42, January 1975

The Structure of a Yeast Hexokinase Monomer and Its Complexes with Substrates at 2.7-Å Resolution*

(protein conformation/nucleotide binding/kinase structure)

ROBERT J. FLETTERICK, DAVID J. BATES†, AND THOMAS A. STEITZ

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520

Communicated by Julian M. Sturtevant, October 17, 1974

ABSTRACT **From a 2.7-Å resolution electron density map we have built a model of the polypeptide backbone of a monomer of yeast hexokinase B (EC 2.7.1.1). This map was obtained from a third crystal form of hexokinase, called BIII, which exhibits space group $P2_12_12_1$ and which contains only one monomer per asymmetric unit. The 51,000 molecular weight monomer has an elongated shape (80 Å by 55 Å by 50 Å) and is divided into two lobes by a deep central cleft. The polypeptide chain is folded into three structural domains, one of which is predominantly α-helical and two of which each contain a β-pleated sheet flanked by α-helices.**

Both glucose and AMP bind to these crystals and produce significant alterations in the protein structure. Glucose binds in the deep cleft, as was observed previously in the BII crystal of the dimeric enzyme. AMP, however, binds to a site that is different from the major intersubunit ATP binding site observed in the crystalline dimer. The AMP is found near one of the β-pleated sheets.

From our current interpretation of this electron density map we conclude that neither of the two nucleotide binding regions has the same structure as has been observed for the nucleotide binding regions of the dehydrogenases, adenylate kinase, and phosphoglycerate kinase, although some similarities exist.

We have been using x-ray diffraction techniques to study three different crystal forms of yeast hexokinase B (EC 2.7.1.1; ATP:D-hexose 6-phosphotransferase), designated BI, BII, and BIII (1, 2, ‡), in order to understand the mechanism of phosphorylation of hexoses by ATP-Mg catalyzed by hexokinase and the regulation of this activity by substrates and other metabolites (3–5). In crystal form BI the two subunits in the asymmetric unit are related by a molecular screw relationship of a 180° rotation combined with a 3.6-Å translation. Because BI crystals do not bind substrate ligands and because heavy atom derivatives of BI have proven to be nonisomorphous at high resolution, our structural studies have focused on forms BII and BIII. Our results at 7-Å resolution with the dimeric enzyme in crystal form BII have shown (2, ‡) that (*i*) the subunits are also nonequivalently associated but with a different molecular screw relationship of a 156° rotation and a 14-Å translation; (*ii*) only one nucleotide binds strongly per dimer at an intersubunit site; (*iii*) there is an allosteric interaction between the sugar and nucleotide sites such that bound hexose is required for nucleotide binding at the subunit interface and that bound nucleotide promotes hexose binding; (*iv*) extensive tertiary structural changes accompany hexose binding; and (*v*) nucleotides bind weakly at two additional equivalent sites, one site per subunit, in the absence of hexoses. Unfortunately, crystal form BII, which has an asymmetric unit composed of one dimer of 102,000 daltons, diffracts x-rays rather poorly at resolution higher than 3.5 Å and the crystals are easily damaged by radiation.

We report here the tertiary structure of the hexokinase monomer determined from x-ray diffraction data to 2.7-Å resolution from BIII crystals. Although the intersubunit nucleotide binding site observed in the BII dimer is not functioning in this crystal form, a different region of the monomer binds AMP or ADP. We do not know in which of these two nucleotide binding sites ATP can function as the phosphoryl donor in the hexokinase reaction. Although the backbone chain trace has some ambiguities associated with it, we conclude that neither of the domains forming the two nucleotide binding sites has a structure that is homologous to that observed in the nucleotide binding domain of the dehydrogenases, phosphoglycerate kinase, or adenylate kinase. However, certain features of AMP binding to hexokinase are similar to some aspects of nucleotide binding by these enzymes.

Structure determination

Crystals of all three forms of yeast hexokinase B can occur simultaneously from the same solution of enzyme at room temperature at 2.2–2.4 M potassium phosphate (pH 7.0). BIII crystals can be more reliably obtained by the crystallization of proteolytically modified (6) enzyme in the presence of 0.4 mM *o*-toluoylglucosamine, a competitive inhibitor. The BIII crystals exhibit the space group $P2_12_12_1$ with unit cell dimension of a = 166.5 (± 0.1), b = 59.2 (± 0.05), and c = 58.5 (± 0.05) Å; the asymmetric unit contains one subunit of 51,000 daltons.

We exploited the fact that two of the subunits' four sulfhydryl groups are reactive in order to prepare three different mercury derivatives. Dimercury acetate (a gift of M. G. Rossmann) reacts with only one sulfhydryl group, while mersalyl and methyl mercury react with both reactive thiols; fortunately, the positions of the mercury of mersalyl and methyl mercury bound to the second thiol are different. In addition to the mercury derivatives, three platinum compounds and *o*-iodobenzoylglucosamine, an iodinated analogue of *o*-toluoylglucosamine, were used in the phasing procedure.

The structure of the BIII crystal form was obtained inde-

* The previous paper in this series is W. F. Anderson and T. A. Steitz (1974), *J. Mol. Biol.*, in press.

† Deceased May 10, 1974.

‡ W. F. Anderson and T. A. Steitz (1974) *J. Mol. Biol.*, in press.

TABLE 1. *Refinement parameters at 2.7-Å resolution*

Derivative	No. of sites	$\langle\|\Delta F\|\rangle/\bar{F}$		20–7 Å	7–4 Å	4–2.7 Å	Average
MeHg*	3	0.12	R_c†	0.36	0.60	0.82	0.66
			rms E†	6.1	7.6	6.8	6.9
			rms f†	20.8	15.7	8.5	11.1
Mersalyl	2	0.10	R_c		0.66	0.91	0.80
			rms E		7.7	6.7	6.9
			rms f		13.7	7.3	8.9
DMA*	3	0.11	R_c	0.47	0.66	0.81	0.70
			rms E	7.8	7.7	6.3	6.7
			rms f	21.0	15.6	9.1	11.5
o-IBG*	1	0.10	R_c	0.64	0.81	0.86	0.79
			rms E	6.4	6.9	7.2	7.0
			rms f	10.4	9.9	8.5	9.2
$Pt(NO_2)_6^-$	5	0.23	R_c	0.57	0.92		0.70
			rms E	18.9	24.0		22.7
			rms f	37.8	26.4		29.2
$Pt(CN)_4^-$	6	0.11	R_c	0.65	0.78		0.70
			rms E	9.2	9.9		9.7
			rms f	14.8	11.0		12.0
$PtEnCl_2^-$*	3	0.13	R_c	0.51	0.82		0.68
			rms E	9.9	11.3		10.9
			rms f	21.6	14.7		16.4
No. of reflections				850	2950	8700	12500
Average figure of merit				0.90	0.80	0.58	0.65

* Abbreviations for heavy atom derivatives: MeHg, methylmercuryacetate; DMA, dimercuryacetate; *o*-IBG, *o*-iodobenzoylglucosamine; $PtEnCl_2^-$, platinum ethylene diamine dichloride.

† rms E, root-mean-square lack of closure error; rms f, root-mean-square heavy atom structure factor.

$$R_c = \frac{\Sigma||F_{PH} - F_P| - |f_H||}{\Sigma||F_{PH}| - |F_P||}, \text{ for centric reflections.}$$

pendently of the other two crystal forms. The heavy atom coordinates were found by standard difference Patterson and difference Fourier techniques and were refined by use of alternate cycles of phasing and least-squares refinement (7). The platinum compounds were not used in the phasing beyond 4.5 Å due to large heavy atom temperature factors and the obvious lack of isomorphism exhibited at higher resolution. The centroid phase angles were calculated for 12,500 of the approximately 15,000 reflections to 2.7-A resolution; the average figure of merit was 0.65. Table 1 lists some of the refinement parameters as a function of resolution.

Since the amino-acid sequence is not yet available, a polyalanine backbone model was constructed from Labquip models (1 cm/Å scale) with a Richards optical comparator (8). Independently, a mini-map drawn to a scale of 0.3 cm/Å was interpreted by placing markers at proposed α-carbon positions. Fig. 1 shows a 10-A thick slice through this map with the α-carbon positions marked. A model of the α-carbon backbone was built from the coordinates measured from the mini-map and had the same chain conformation as the polyalanine model. A schematic drawing of the course of the polypeptide backbone is shown in Fig. 2.

Description of the hexokinase monomer

The hexokinase monomer has an elongated shape of approximate dimensions 80 Å by 55 A by 50 Å, and is divided into two lobes by a deep cleft. We have located 451 amino-acid residues from this map. A comparison of the side chains in our electron density map with the amino-acid sequence of the first 30 residues kindly provided by Dr. E. Barnard suggests that, since they are not apparent in the map, residues 1 to 25 are either disordered in this crystal and/or have been cleaved by a yeast protease (9, 6). As shown in Fig. 2, the two lobes differ dramatically in the type of secondary structure they contain; one lobe consists largely of α-helix while the other contains substantial β structure in addition to helix. Approximately 40% of the polypeptide backbone is found in the α-helical conformation and 30% in β-sheet structure.

There are two extensive regions of β structure, each flanked by α-helices; these two β-pleated sheets can also be described as forming an 11-stranded, barrel-like structure. Barrel-like β-structures consisting of six or seven strands have been described for other enzymes, notably α-chymotrypsin (10). The 11-stranded barrel in hexokinase is much larger in diameter and is found to be filled with two α-helices in its center.

The general structural features of this kinase are dissimilar to adenylate kinase (11) and phosphoglycerate kinase (12, 13) with the notable exception of some aspects of the domain forming one of the two nucleotide binding sites (see below).

Substrate binding

Sugar. A 3.0-Å resolution difference electron density map shows that glucose binds in the cleft that separates the molecule into two lobes and in about the same position as was observed earlier with the BII crystals of the dimeric enzyme‡. Furthermore, the binding of glucose is accompanied by extensive alterations in protein structure throughout the monomer. Particularly striking structure changes involve both the backbone and side chains of several strands of β-structure which are in contact with the nucleotide binding sites. It is

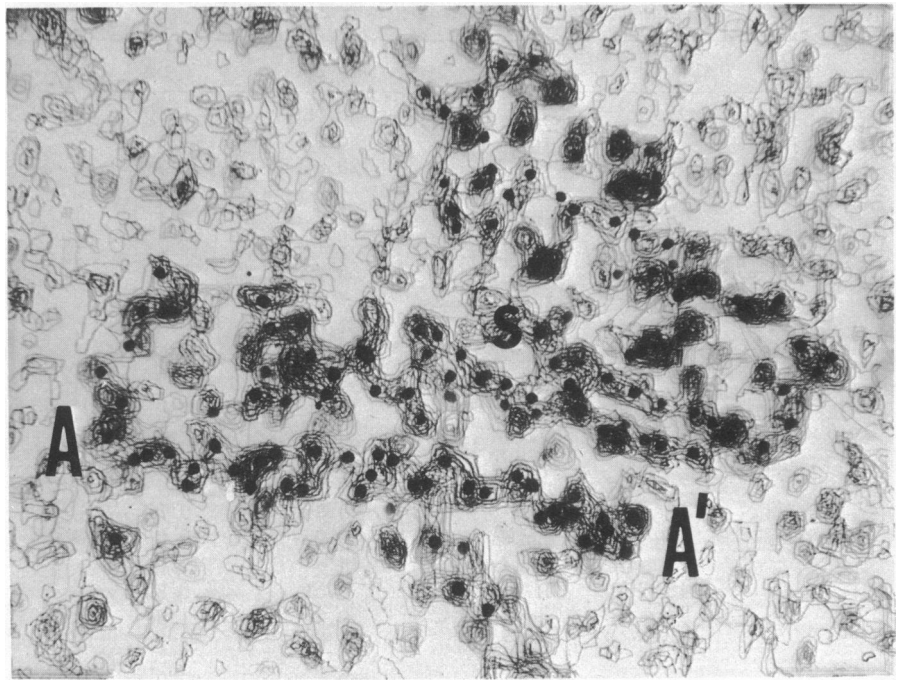

FIG. 1. A portion of the 2.7-Å resolution electron density map. A 50-Å long helix is indicated by A,A'. The large cleft in the molecule can be seen, with the bound *o*-toluoylglucosamine molecule indicated by S.

difficult to make detailed interpretations of these structural changes at this stage in the analysis in the absence of the amino-acid sequence.

A difference Fourier map between an *o*-toluoylglucosamine (a competitive inhibitor) complex and the glucose complex clearly shows the phenyl ring of the inhibitor and no difference electron density at the position of the glucose. Thus, at 3.0-Å resolution, the orientation of glucose and that of the sugar moiety of *o*-toluoylglucosamine appear to be very similar or perhaps the same.

The difference map between these two sugar complexes shows many fewer positive and negative density features than either of the difference maps of the complexes with respect to native enzyme. This observation suggests that most of the

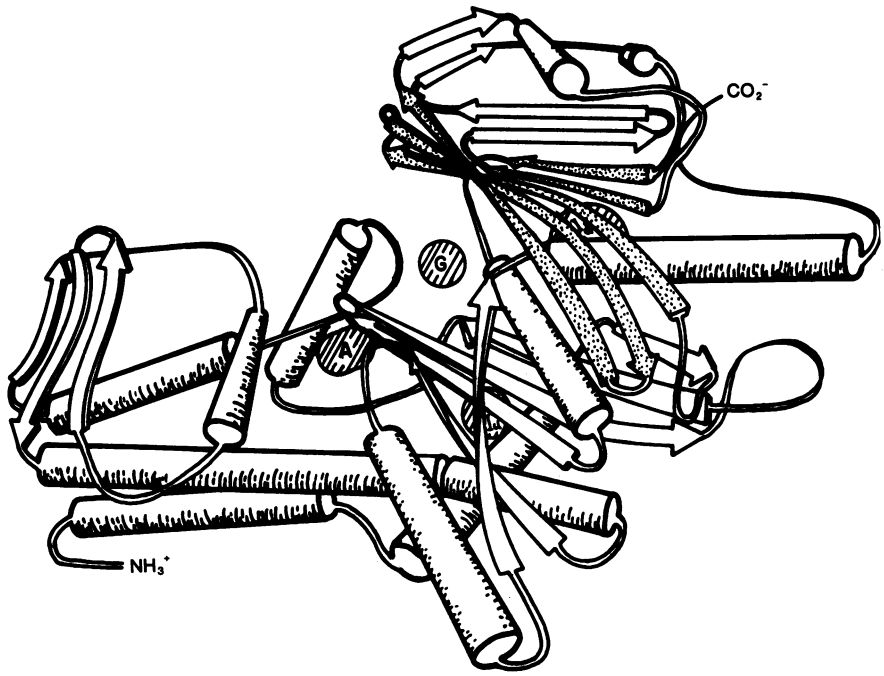

FIG. 2. A schematic drawing of the course of the polypeptide backbone of a subunit of hexokinase B. Helical regions are indicated by tubes and regions of β-structure by arrows. One of the two β-sheets that form a barrel-like structure has been stippled. The position to which glucose binds is indicated by a G. The binding site for AMP found in BIII crystals is indicated by an A and lies at the end of the unstippled β-structure. The positions of the two regions involved in formation of the intersubunit ATP site in BII crystals are indicated by I_u and I_d.

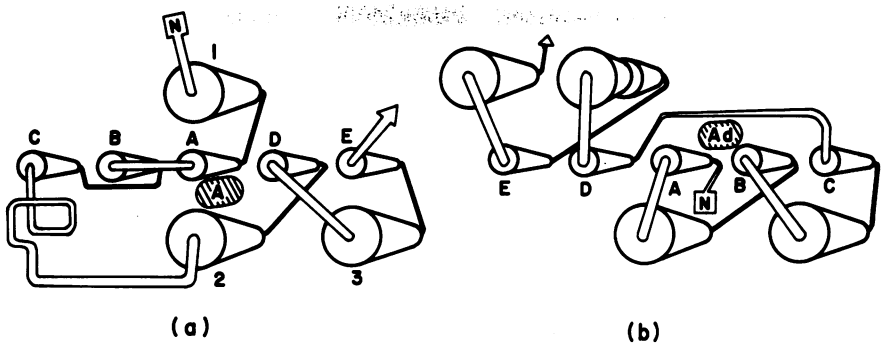

FIG. 3. Schematic comparison of the topology of the A nucleotide binding domain of hexokinase (a) and the nucleotide binding domain of lactate dehydrogenase (b). Helices are represented by large tubes and β strands by small tubes. In (a), the hexokinase binding site for AMP is indicated by A; in (b) the binding site for the adenosine of NAD is indicated by Ad. While the possibility of some errors in the tracing of the polypeptide chain of hexokinase cannot be excluded, it is not possible to connect the β structure and helices of hexokinase (a) in the same way as is observed in lactate dehydrogenase (b). For example, connecting the carboxyl end of strand A to helix 2 of hexokinase results in a topology that is more nearly like that of lactate dehydrogenase, but would require "breaking into" the middle of the 50 Å long α-helix (A–A' in Fig. 1). Such an interpretation is inconsistent with this electron density map and with the 3.5-Å map of the BII crystal form (Anderson, Fletterick, and Steitz, unpublished results). Although the adenosine binds between the second and third strands of the β-sheet in both hexokinase and lactate dehydrogenase, the adenosine binds "below" the β-sheet in hexokinase but "above" the β-sheet in lactate dehydrogenase.

structural changes induced by the two sugars are the same. However, the difference map does indicate some definite changes in the protein structure, demonstrating that the enzyme structural changes induced by the substrate and the inhibitor are not completely identical. These different protein conformations induced by glucose and *o*-toluoylglucosamine may account for *o*-toluoylglucosamine being an inhibitor rather than a substrate, even though it binds to the monomer in approximately the same orientation as glucose.

Nucleotide. Due to the heterologous subunit interaction in the dimer, the intersubunit nucleotide binding site (I site) found in the BII dimer is formed by two different regions of the subunit. Thus, two regions exist in the monomer that could potentially be expected to bind nucleotide; these two regions are designated I_u and I_d in Fig. 2. Numerous attempts to bind nucleotides or analogues have failed to produce convincing crystallographic evidence of binding at either of the partial sites in the BIII crystalline state. The reasons for this failure may be that the I site requires subunit association and/or tertiary structural changes to be effective in nucleotide binding or that the crystal packing makes these regions inaccessible to nucleotide.

While we have not been able to bind AMP, ADP, or ATP to the I sites either in the presence or absence of glucose, we find that AMP or ADP binds to another site (A site). Low resolution difference electron density maps show that either 50 mM AMP or 0.05 mM diiodofluorescein can bind to BIII crystals in a small pocket that is adjacent to the sugar site (Fig. 2). Wassarman and Lentz (14) found, from a 2.8-Å resolution difference electron density map, that diiodofluorescein binds to lactate dehydrogenase in exactly the same location as the adenosine ring of NAD. This also appears to be the case with hexokinase. The iodines of the diiodofluorescein and the adenosine of AMP are found in a pocket that is located at the end of the second and third strands of a β-pleated sheet (the lower, unstippled sheet in Fig. 2). Assuming that ATP binds as AMP, model building suggests that the γ-phosphate of an ATP molecule may be as much as 10 Å from the 6-hydroxyl of glucose or as close as a few Ångstroms.

This AMP binding site is not in the same position as the major intersubunit nucleotide site (I site) observed in the BII dimer‡. Rather, the two I binding sites and the A site are on opposite sides of the glucose. A portion of the I site is located at the end of the other β-pleated sheet (the upper, stippled sheet in Fig. 2). Presumably, since this site is formed by two subunits, the complete site exists only in the dimer form of the enzyme. It is not possible to say at this point which of these two nucleotide binding sites is involved in the phosphoryl transfer. Presumably, one of these sites is at the active site while the other nucleotide site possibly functions in allosteric regulation as suggested by solution studies (3, 5).

Comparison of structure with other nucleotide binding proteins

The structures of the domains of several kinases and dehydrogenases that bind nucleotide cofactors appear to show certain similarities (15, 16, 12, 13, 11, 17, 18). In the cases of lactate, glyceraldehyde phosphate, and alcohol dehydrogenases, and phosphoglycerate kinase and adenylate kinase, the nucleotide binding domain of the protein consists of a 5- or 6-stranded parallel β-pleated sheet, with parallel α-helices arranged on either side of the sheet. Further, some similarity exists in the connection of these secondary structure features as well as the location of the adenine binding site, which is always found to lie at the carboxyl end of the sheet.

Rossmann and coworkers (15, 16) have suggested that the nucleotide binding domains of all kinases and dehydrogenases have evolved from a common ancestral gene and diverged to such an extent that the amino-acid sequences are no longer homologous by the usual standards, and that α-helices have been inserted and deleted. It has been suggested (16) that most, if not all, proteins that bind nucleotide will contain a domain of this structure.

The hexokinase domain that provides the two regions that participate in binding ATP at the intersubunit site (I site) is completely different from the lactate dehydrogenase fold, while the domain involved in binding the AMP (A site) is only superficially similar (Fig. 3). At the A site the adenosine does bind at the carboxyl end of a β-pleated sheet containing

three parallel and two anti-parallel strands. However, in our current interpretation, the connection of these strands and the arrangement of helical regions is not the same as in lactate dehydrogenase. There are two possible reasons for the difference in structure between this AMP binding site in hexokinase and the NAD binding site in lactate dehydrogenase: either we have incorrectly interpreted our electron density map, or not all nucleotide binding domains have evolved from a single ancestral gene.

At 2.7-Å resolution and in the absence of amino-acid sequence information, it is certainly possible to make some incorrect connections of chains, since there are a few branch points in the electron density due to strong noncovalent interactions. We have tried to reinterpret our electron density map in a manner consistent with the lactate dehydrogenase fold; however, such an interpretation does not appear to be consistent with the electron density in this map. If our interpretation for the structure of the domains that form the A and I nucleotide binding sites is correct, then it is unlikely that these nucleotide binding domains have evolved from Rossmann's proposed ancestral gene.

As an alternative to divergent evolution from a single or small number of genes for nucleotide binding domains, it remains possible that the observed similarities among the kinases and dehydrogenases at their nucleotide binding sites are due to a combination of convergent evolution and a limitation on the number of ways of forming certain supersecondary structures. That is, there may be some functional reason that accounts for adenosine binding at the carboxyl end of parallel β-sheets. Further, there may be only a limited number of ways of folding a parallel β-pleated sheet. All proteins solved to date that contain a β-pleated sheet consisting of only parallel strands (about eight structures) bear some structural similarities to NAD binding domain of lactate dehydrogenase. Perhaps these similarities in structure provide insight into the protein folding problem.

We thank Prof. E. A. Barnard and his colleagues for the gifts of hexokinase B. We also thank David Agard for aiding in the development of the contouring program and Wayne F. Anderson for useful discussions. This research was supported by Public Health Service Grant GM18268 and a Jane Coffin Childs Memorial Fund for Medical Research postdoctoral fellowship to D.J.B.

1. Steitz, T. A., Fletterick, R. J. & Hwang, K. J. (1973) *J. Mol. Biol.* **78,** 551–561.
2. Anderson, W. F., Fletterick, R. J. & Steitz, T. A. (1974) *J. Mol. Biol.* **86,** 261–269.
3. Kosow, D. P. & Rose, I. A. (1971) *J. Biol. Chem.* **246,** 2618–2625.
4. Ainslie, G. R., Jr., Shill, J. P. & Neet, K. E. (1972) *J. Biol. Chem.* **247,** 7088–7096.
5. Colowick, S. P. (1973) in *The Enzymes* (Academic Press, New York), Vol. IX, 3rd ed., pp. 1–48.
6. Schmidt, J. J. & Colowick, S. P. (1973) *Arch. Biochem. Biophys.* **158,** 471–477.
7. Adams, M. J., Haas, D. J., Jeffery, B. A., McPherson, A., Jr., Mermall, H. L., Rossmann, M. G., Schevitz, R. W. & Wonacott, A. J. (1969) *J. Mol. Biol.* **41,** 159–188.
8. Richards, F. M. (1968) *J. Mol. Biol.* **37,** 225–230.
9. Lazarus, N. R., Derechin, M. & Barnard, E. A. (1968) *Biochemistry* **7,** 2390–2400.
10. Birktoft, J. J. & Blow, D. M. (1972) *J. Mol. Biol.* **68,** 187–240.
11. Schulz, G. E., Elzinga, M., Marx, F. & Schirmer, R. H. (1974) *Nature* **250,** 120–123.
12. Blake, C. C. F. & Evans, P. R. (1974) *J. Mol. Biol.* **84,** 585–598.
13. Bryant, T. N., Watson, H. C. & Wendell, P. F. (1974) *Nature* **247,** 14–17.
14. Wassarman, P. M. & Lentz, P. J., Jr. (1971) *J. Mol. Biol.* **60,** 509–522.
15. Rossmann, M. G., Moras, D. & Olsen, K. W. (1974) *Nature* **250,** 194–199.
16. Liljas, A. & Rossmann, M. G. (1974) *Annu. Rev. Biochem.* **43,** 475–508.
17. Bränden, C. I., Eklund, H., Nordström, B., Boiwe, T., Söderlund, G., Zeppezauer, E., Ohlsson, I. & Åkeson, Å. (1973) *Proc. Nat. Acad. Sci. USA* **70,** 2439–2442.
18. Schulz, G. E. & Schirmer, R. H. (1974) *Nature* **250,** 142–144.