Journal of Supramolecular Structure 6:503–518 (1977)
Molecular Aspects of Membrane Transport 333–348

Crystallographic and Chemical Studies of the L-Arabinose-Binding Protein From E. coli

F. A. Quiocho, G. L. Gilliland, D. M. Miller, and M. E. Newcomer

Department of Biochemistry, Rice University, Houston, Texas 77001

The crystal structure of the L-arabinose-binding protein (ABP), an essential component of the high affinity L-arabinose transport system in E. coli, has been determined at 3.5- and 2.8-Å resolutions. The Fourier maps indicate that the molecule is ellipsoidal with overall dimensions of 70 × 35 × 35 Å (axial ratio ≃ 2:1) and consists of 2 distinct globular domains (designated "P" and "Q"). A tentative trace of the polypeptide backbone is presented. The 2 domains are arranged to create a deep and narrow cleft, the base of which is formed by 3 polypeptide chain segments linking the 2 domains. The arrangements of the secondary structure of the 2 domains are remarkably similar and can be related by a pseudo-twofold axis. Each domain has a pleated sheet core with 2 helices on either side of the plane of the β sheet. This secondary structural arrangement is similar to that found in other proteins, specifically the dehydrogenases and kinases. The structural similarity is particularly intriguing in light of the recent finding in this laboratory that the dye 2′,4′,5′,7′-tetraiodofluorescein, an adenine analogue which has been shown to bind to several dehydrogenases and kinases, binds to ABP with a dissociation constant of 30 μM.

Experiments performed with protein, modified with the chromophoric probe 2-chloromercuri-4-nitrophenol (MNP), suggest that the binding site is near an essential cysteine residue: modification of the thiol with the mercurial dramatically decreases the ligand-binding affinity of ABP, and conversely, the sugar protects the cysteine from reaction with MNP. The binding of L-arabinose to MNP-labeled protein perturbs the nitrophenol absorbance spectrum. The essential cysteine has been assigned to position 64 in the proposed chain tracing, which is consistent with the amino acid sequence. As an explanation for the failure of the difference Fourier analyses to locate the sugar-binding site, it is postulated that the structure has been solved with the sugar bound. Electron density to which no amino acid residue can be assigned and which could be the sugar molecule is within van der Waals distance of the sulfur atom.

Key words: L-arabinose-binding protein, three-dimensional structure, spectrochemical studies, active transport, chemotaxis

Abbreviations: ABP – L-arabinose-binding protein; MNP – 2-chloromercuri-o-nitrophenol; PCMBS – p-chloromercuribenzenesulfonic acid; TIF – 2′,4′,5′,7′-tetraiodofluorescein

Received for publication March 14, 1977; accepted May 25, 1977

Following the discovery of the sulfate-binding protein by Pardee in 1966 (1), a number of low-molecular-weight, water-soluble proteins binding a variety of small substrates (sugars, inorganic ions, and amino acids) have been isolated from the osmotic shock fluid of gram-negative bacteria (for reviews see Refs. 2–4). The localization of binding proteins to the peptidoglycan region between the cytoplasmic and outer membranes (5–6) and the correlation of transport-negative mutants with reduced amounts or altered forms of binding proteins (2–4, 7) suggest that binding proteins are essential components of some high-affinity uptake systems. It has also been demonstrated that binding proteins participate in bacterial chemotaxis (8–10). Recent evidence suggests that binding-protein transport (11) and chemotaxis (12–13) systems are linked to other, presumably membrane-bound, protein components which either facilitate the movement of ligand into the cell or signal the flagella to suppress twiddling activity.

One of these binding proteins, the L-arabinose-binding protein (ABP), has been purified from E. coli B/r (14) and crystals suitable for high resolution x-ray analysis have been obtained (15). The L-arabinose-binding protein is an essential component of high-affinity L-arabinose transport (16), and like all binding proteins, is composed of a single polypeptide chain. The amino acid sequence of the L-arabinose-binding protein has recently been determined by Hogg and Hermodson (17), and the molecular weight derived from this analysis is 33,200. The results of 5-Å and 3.5-Å resolution structural analyses, indicating an ellipsoidal and bilobate molecular structure of the binding protein with an axial ratio of 2:1, have recently been reported (18).

In this paper we present results of our crystallographic analysis of the L-arabinose-binding protein and evidence, based upon studies of the protein modified with the chromophoric probe, 2-chloromercuri-4-nitrophenol (MNP), that the ligand binding site is near an essential thiol residue.

METHODS AND MATERIALS

The L-arabinose-binding protein was purified from E. coli B/r strain UP1041 (ara A39) according to the procedure of Parsons and Hogg (14) with the exclusion of the final step (denaturation in 8 M urea to remove residual L-arabinose). (Cells are induced with 20 mM L-arabinose during the exponential phase of growth.) Crystals suitable for x-ray analysis were grown in 2-methyl-2,4-pentanediol and 3 mM potassium phosphate, pH 6.5 (15). These crystals belong to the space group $P2_12_12_1$ with one molecule per asymmetric unit. The unit cell dimensions recently obtained by least squares analysis of orientation parameters for a number of crystals are: a = 55.44 ± 0.08, b = 71.72 ± 0.15, and c = 77.64 ± 0.23 Å. Diffraction patterns extend to a spacing of approximately 2 Å.

X-ray diffraction intensities were measured with the omega step-scan procedure of Wyckoff et al. (19). A more detailed description of the procedure used for collecting and processing of diffraction data is given elsewhere (18).

ABP utilized for chemical studies was stored at −20°C as a crystalline suspension in 60% 2-methyl-2,4-pentanediol, 0.1% β-mercaptoethanol, and 3 mM potassium phosphate, pH 6.5. Prior to reaction with the thiol-specific reagent, 2-chloromercuri-4-nitrophenol, the protein was dialyzed overnight into a solution of 6 M guanidine-hydrochloride, 5 mM ehtylenediaminetetraacetic acid, 2 mM dithiothreitol, 10 mM Tris-HCl, pH 7.4 and subsequently dialyzed exhaustively against the appropriate buffer.

Assays were routinely performed by equilibrium dialysis in plexiglass microcells modeled after the design of Willis et al. (20). The purified protein exhibits native binding

activity (3×10^{-7} M and 4×10^{-7} M for the dissociation constants of L-arabinose and D-galactose, respectively) and migrates as a single band on polyacrylamide gels (14). Protein was determined from either the absorbance at 280 nm ($\epsilon = 0.94\ 1\ g^{-1}\ cm^{-1}$) (14) or by the method of Lowry et al. (21) using lyophilized ABP as a standard. All spectrophotometric measurements were carried out with the use of a Cary 118 or Varian 635 Spectrophotometer. Mercury analysis by atomic absorption spectrophotometry was performed with an Instrumentation Lab Inc. Absorption-Emission Spectrophotometer-153.

RESULTS

Structure

The calculation of a much improved 3.5-Å electron density map, followed shortly by a 2.8-Å map, has provided considerable information about the 3-dimensional structure of the L-arabinose-binding protein. The improvement of the initial 3.5-Å resolution map (18) was achieved by the addition of 2 more good heavy-atom derivatives, namely, p-chloromercuribenzenesulfonic acid (PCMBS) and CdI_2. A 3.5-Å refinement of these 2 derivatives, together with 2-chloromercuri-4-nitrophenol, K_2PtCl_4, and K_2IrCl_6 yielded a figure of merit of 0.76 for the 4,000 reflections phased. This figure of merit is a significant improvement over 0.66, the value obtained in the initial 3.5-Å phase refinement (18). Phases used to calculate electron density maps were obtained by the method of alternate cycles of least-squares refinement of heavy-atom parameters and multiple isomorphous replacement phase determination (22, 23).

The 2.8-Å resolution Fourier map recently calculated was based entirely upon the phases derived by extending the PCMBS and CdI_2 derivatives to 2.8-Å resolution. The mean figure of merit was 0.65. The heavy atom parameters (Table I) utilized for the 2.8-Å phasing were derived from a refinement of the final heavy-atom parameters used in the phasing of the 3.5-Å map. A summary of the phasing statistics resulting from the 2.8-Å resolution is presented in Table II. A more detailed description of the 2.8-Å resolution phase refinement will be presented elsewhere.

The complete 2.8-Å electron density map for the L-arabinose-binding protein is viewed up the z-axis in Fig. 1A–E. It will be evident from the description of this map which follows that the molecule is an ellipsoid consisting of 2 distinct globular domains, designated for convenience as the "P" and "Q" domains. The molecule lies inclined to the x-y plane with the P domain visible in the lower left at higher z values and the Q domain in the upper right. The electron density of the protein molecule is quite distinguishable from the low electron density of the mother liquor.

In Fig. 1A the central region of electron density is a portion of the C-terminal helix, and the upper right-hand region is part of the P domain. The regions of electron density around the border are from adjacent ABP molecules.

The protein boundary is clearly delineated in Fig. 1B with electron density from a neighboring molecule appearing at the lower left. The Q domain is fully outlined (upper right) and the P domain is just beginning to emerge (lower left). This section depicts 2 of the 3 chain segments which connect the 2 domains.

Fig. 1C reveals the 2:1 axial ratio of the molecule. The division of the P and Q domains is not clear in this view of the map due to the perspective. There is only one true connection between the 2 domains in this region. Also, there are several helical regions visible in both domains.

TABLE I. Heavy-Atom Parameters Based on the 2.8-Å Resolution Refinement

Site no.	x	y	z	B(Å)^{2a}	Fractional occupancy
		CdI_2 derivative			
1	0.4658	0.9749	0.5988	46.1	1.11
2	−0.0769	1.1326	0.5498	49.6	1.11
3	0.0086	1.1921	0.5457	98.4	1.21
4	0.0618	1.1957	0.7398	76.5	0.77
5	0.3715	1.1388	0.7334	50.8	0.32
6	0.3764	0.2951	0.8578	73.3	0.34
		Hg, PCMBS derivative			
1	0.1078	1.1313	0.7708	31.1	0.77
2	0.3075	0.7365	0.7407	7.8	0.82
3	0.0061	1.1611	0.5286	14.3	0.16
		Hg, MNP derivative			
1	0.1035	1.1275	0.7704	57.3	0.86
2	0.3051	0.7373	0.7409	10.0[b]	0.27
		$(NH_4)_2PtCl_4$ derivative			
1	0.4521	0.8189	0.6951	28.5	1.02
2	0.3978	0.3113	0.6653	177.1	1.06
3	0.0294	1.1707	0.5287	30.1	0.65
4	0.0634	1.1081	0.7636	10.0[b]	0.11
5	0.5313	0.8256	0.6677	10.0[b]	0.15
6	0.4993	0.6735	0.7719	10.0[b]	0.19
7	0.1912	1.1272	0.6873	17.0	0.18
8	0.3834	0.8418	0.6789	10.0[b]	0.14
9	0.4917	0.7346	0.8842	74.2	0.31
		K_2IrCl_6 derivative			
1	0.5297	0.7766	0.9345	245.1	0.79
2	0.5326	0.7219	0.8879	153.3	0.42
3	0.4812	0.8744	0.6654	12.8	0.13
4	−0.0537	1.1487	0.5027	216.8	0.28
5	0.0994	1.1287	0.7558	188.1	0.18

[a]Isotropic temperature factor.
[b]Parameters which were not refined.

Several interesting features are seen in Fig. 1D. The molecular boundary is outlined by the dashed line, and the division between the P and Q domains, which is quite pronounced, is indicated by the arrow. Four strands of the P-domain β-sheet are clearly visible. A long segment of helix appears in the Q domain (upper right) and a screw-related helix in another molecule is seen in the upper left. Neighboring molecules protrude into this region of the map at all 4 corners.

TABLE II. Summary of the 2.8-Å Phasing Statistics (Overall Figure of Merit = 65%)

Heavy atom (resolution)	RMS-E[a]	RMS-F_H[b]	RMS-F_H/RMS-E	R factor[c]
Hg, PCMBS (2.8 Å)	53	129	2.43	0.078
Hg, MNP (3.5 Å)	48	81	1.69	0.062
CdI_2 (2.8 Å)	54	90	1.67	0.080
$(NH_4)_2PtCl_4$ (3.5 Å)	104	144	1.38	0.129
K_2IrCl_6 (5.0 Å)	37	55	1.49	0.051

[a]Root mean square lack-of-closure error $E = (\Sigma_h \epsilon_{hj}^2/n)^{1/2}$, ϵ_{hj} = lack of closure for reflection h of derivative j and n = number of reflections.
[b]Root mean square heavy atom contribution $F_H = (\Sigma_h f_{hj}^2/n)^{1/2}$, f_{hj} = heavy atom scattering factor.
[c]Kraut R factor = $(\Sigma||F_D| - |F_N + f_D||)/\Sigma|F_D$, F_D = structure factor of the derivative, F_N for the protein and f_D for the heavy atoms.

Figure 1E shows an almost total disappearance of the Q domain with only 2 parallel helical regions of the P domain remaining. Electron density from the 4 neighboring molecules is still visible.

A tentative trace of the course of the polypeptide chain was first derived from an interpretation of the 3.5-Å map. With a few minor modifications, this trace proved to be essentially identical to the one based upon the interpretation of the 2.8-Å map. Attempts were made to fit the complete amino acid sequence (17) to the improved 3.5-Å and later to the 2.8-Å maps. Approximately two-thirds of the amino acid sequence have been successfully fitted to the provisional trace. The identification of the N-terminal peptide (up to residue 95) and the C-terminal peptide (starting from 189) was verified by fitting the known sequence. A stereo drawing of the provisional trace is shown in Fig. 2. Depicted in Fig. 3 is a schematic drawing of the molecule. The molecular dimensions are 70 × 35 × 35 Å. The fact that the L-arabinose-binding protein is ellipsoidal, with an axial ratio of 2:1, could account for the initial overestimation of the molecular weight of 38,000 for the protein by sedimentation and gel filtration techniques (11).

In our proposed trace, approximately the first 35% of the polypeptide chain forms the major part of the P domain and the next 43% forms the entire Q domain. The remainder of the chain (~ 22%) protrudes as a loop or "handle," then extends back into the P domain, and finally forms the C-terminal helix shared by both domains.

The arrangements of the secondary structure ("super secondary structure") of the 2 domains are remarkably similar and both domains can be related by a pseudo-twofold axis (Fig. 3). Each domain has a central pleated sheet core consisting entirely of parallel strands (with the exception of the sixth antiparallel strand in the Q domain), and on either side of the plane of the β sheet lie 2 helices antiparallel to the β sheet. This arrangement of secondary structure is similar to that found in other proteins, specifically the dehydrogenases and kinases, and has been termed the "nucleotide-binding fold" (24). A detailed quantitative comparison of the 2 domains of ABP and of similar domains found in other proteins is presently underway.

This structural similarity is particularly significant in light of a recent finding in this laboratory that the dye 2′,4′,5′,7′-tetraiodofluorescein (TIF) binds to ABP with a dissociation constant of approximately 30 μM. TIF has been used as a spectral and crystallo-

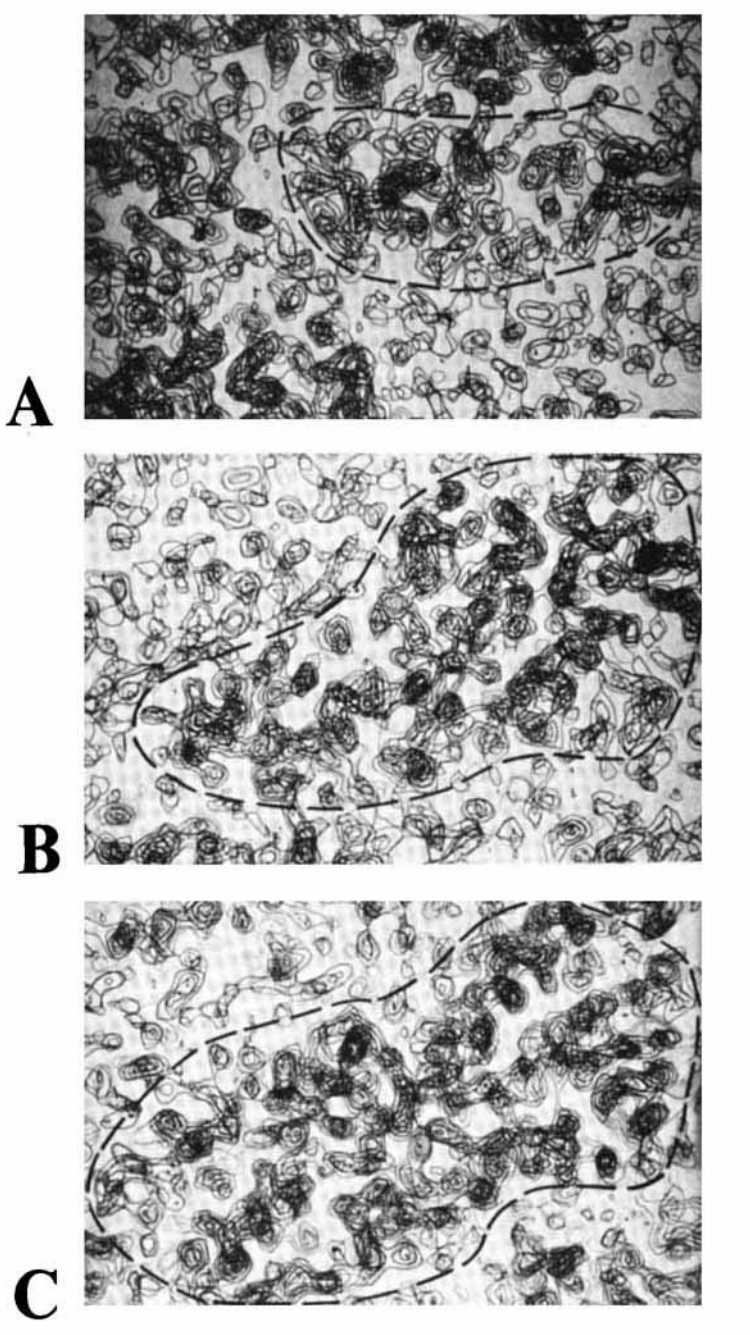
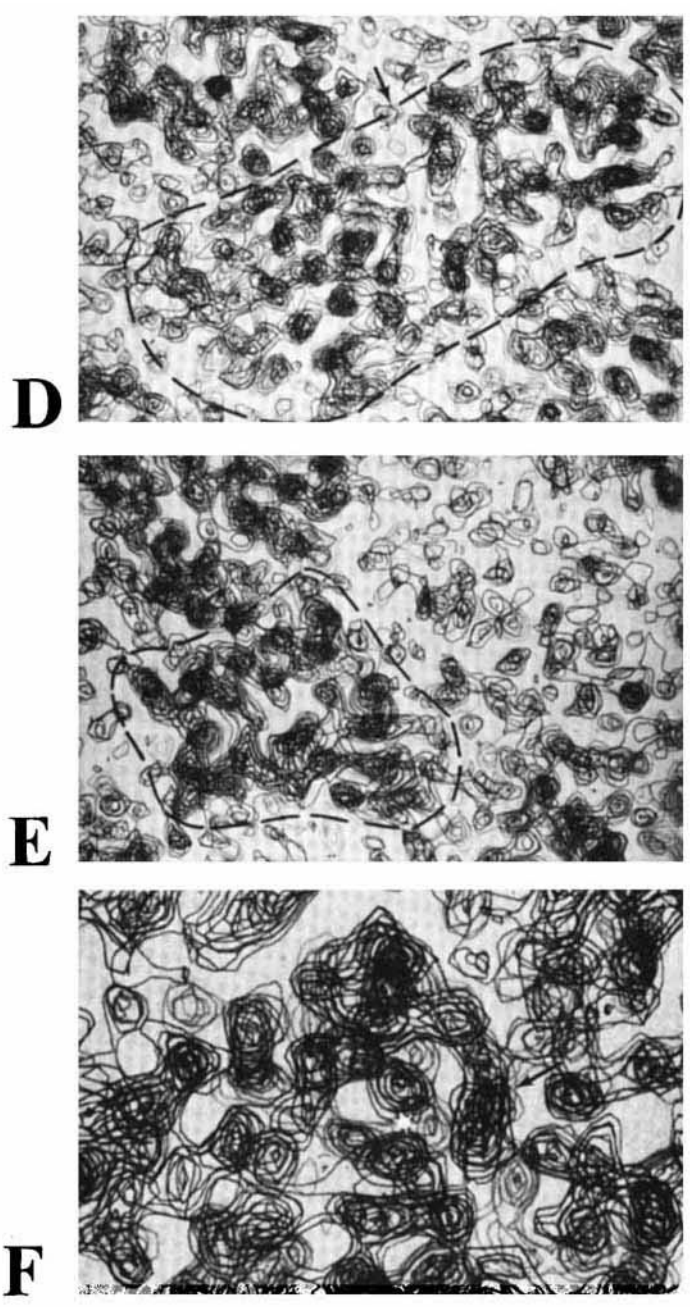


Fig. 1. The complete 2.8-Å electron density map of L-arabinose binding protein viewed up the z axis. Each of the views, A–E, consist of 10 superposed sections and portions of the molecule are enclosed by a dashed line. The bounds of the map are: x = −0.20 to 0.70 (~ 50 Å), y = 0.25–1.25 (~ 72 Å), and z = 0.389–0.918 (~ 41 Å) with the x axis vertical and the y axis horizontal. The upper left corner is at



x = −0.20, y = 0.25. A) z = 0.389–0.486. B) z = 0.497–0.594. C) z = 0.605–0.702. D) z = 0.713–0.810. The arrow marks the separation of the 2 domains. E) z = 0.821–0.918. F) An enlargement of the proposed sugar binding site region consisting of 12 superposed sections with z = 0.626–0.745 [y = 0.47–0.93 (~ 33 Å)]. The arrow indicates the position of an "extraneous" electron density and the asterisk is above the single "essential" cysteine, residue no. 64.

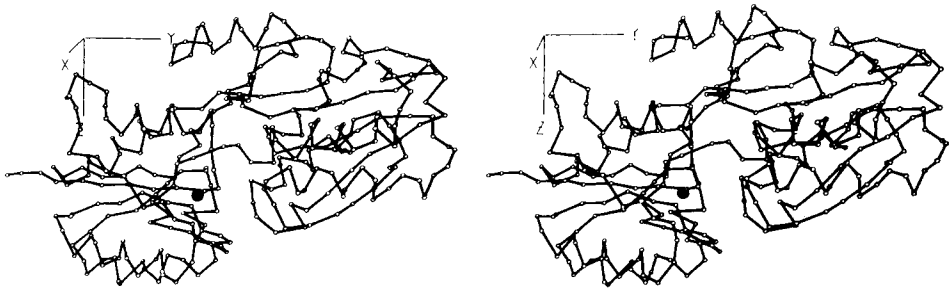


Fig. 2. Stereo drawing of the proposed polypeptide chain trace for L-arabinose-binding protein. The black dot indicates the essential cysteine.

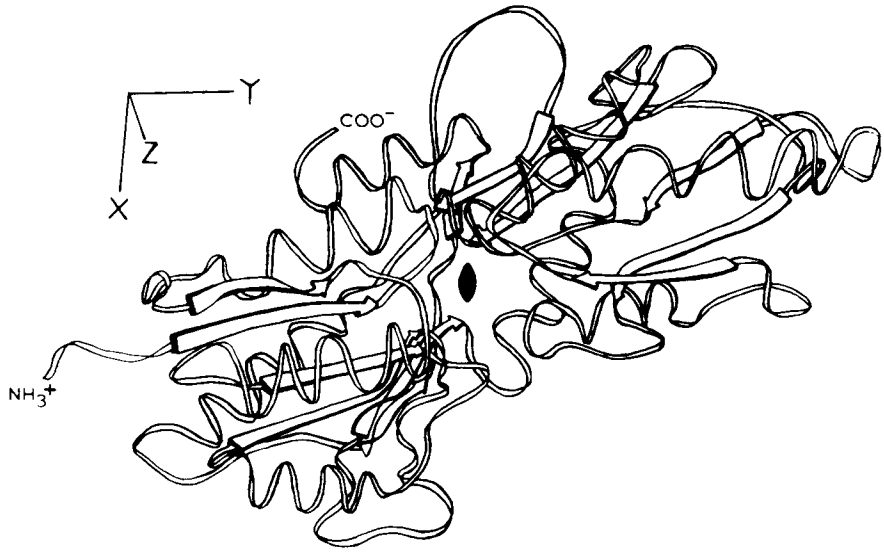


Fig. 3. A schematic representation of L-arabinose-binding protein. The β sheets are indicated by arrows pointing from N to C. A pseudo-twofold axis is located directly in the center of the molecule, perpendicular to the plane of the page.

graphic probe of the nucleotide binding site in a variety of enzymes, e.g., lactate dehydrogenase (25), aspartate transcarbamylase (26), and creatine kinase (27). The difference spectra obtained upon binding of TIF to ABP (Fig. 4) is characteristic of spectra observed for TIF binding to the aforementioned enzymes and suggests the existence of a "super secondary structure" akin to the "nucleotide fold." A similar TIF difference spectrum was also observed for the D-galactose-binding protein. The significance of 2 structurally

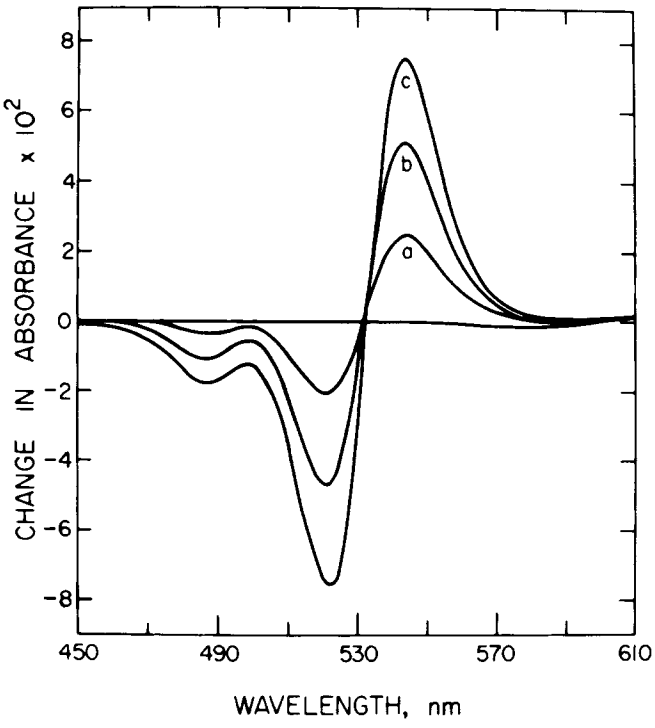


Fig. 4. Difference spectra of free vs tetraiodofluorescein bound to arabinose-binding protein. The sample contained 1.4×10^{-4} M arabinose-binding protein in 0.1 M Tris-hydrochloride, pH 8.2, and the reference cell buffer only. Curves a, b, c are spectra produced by the addition of a small volume of 1 mM tetraiodofluorescein to both cells to give a concentration of 5.0, 13, and 15 μM, respectively.

similar domains in L-arabinose-binding protein and especially the similarity of these domains with those found in unrelated enzymes remains an intriguing mystery. Their existence in this binding protein supports the notion that they can be formed readily and could have evolved independently. We are currently attempting to determine whether the L-arabinose-binding protein may bind nucleotides and various nucleotide analogs.

Essential Thiol Residue and Sugar Binding Site

Several attempts have been made using difference Fourier techniques to locate the sugar-binding site from crystals soaked or cocrystallized in solutions containing L-arabinose or D-galactose. Difference Fourier maps, however, showed no peaks above background. A likely explanation for these failures to locate the sugar-binding site is the possibility that the native protein may have been crystallized with tightly bound L-arabinose (introduced during cell growth).

The demonstration that ABP contains an "essential" cysteine residue provided us with a means to locate the sugar-binding-site region. Both chemical and crystallographic

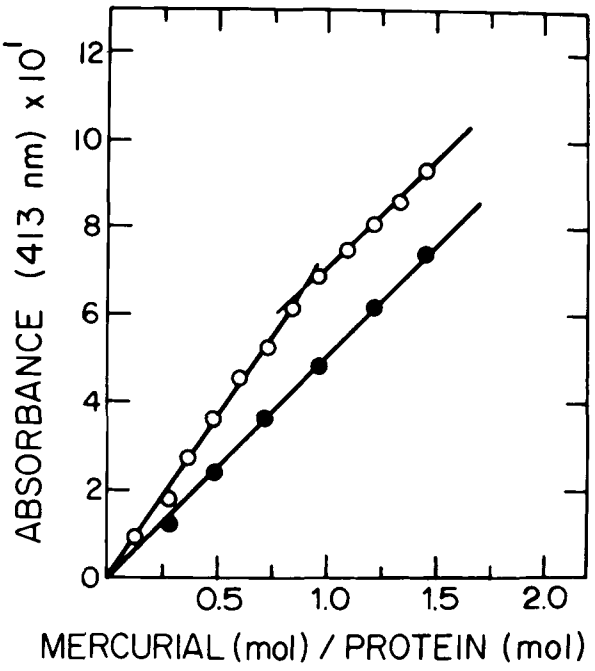


Fig. 5. Titration of arabinose-binding protein with 2-chloromercuri-4-nitrophenol. Protein (4.74×10^{-5}M) in 1.5 ml of 5 M guanidine hydrochloride, 5 mM EDTA, 0.1 M triethanolamine hydrochloride buffer, pH 8.0, was titrated with 5-μl aliquots of MNP (1.72 mM). The absorbance at 413 nm of the protein sample and blank was read after each addition. ○) Protein; ●) blank.

analyses indicate that ABP contains one cysteine residue. Titration of the protein in the denatured state with MNP, a thiol-specific chromophoric probe (28), yielded an inflection point at 0.88 ± 0.03 equivalents of added mercurial (Fig. 5) (29). These results were duplicated by a separate titration with Ellman's reagent (30) 5,5′-dithiobis(2-nitrobenzoic acid) (D. M. Miller, manuscript in preparation). Furthermore, mercury analysis of an ABP sample (5.2 mg/ml) incubated with a fivefold excess of MNP for 3 days and separated from unreacted mercurial by passage through a Sephadex G-25 column and exhaustive dialysis gave a ratio of 0.8 moles of mercury per mole of protein.

The observation that thiol-specific reagents dramatically decrease the affinity of ABP for L-arabinose (14, 29) and that L-arabinose inhibits the rate of reaction of MNP and 5,5′-dithiobis(2-nitrobenzoic acid) with the protein suggest that the thiol is near the binding site (D.M. Miller, manuscript in preparation). Figure 6 illustrates the maximal inhibition of binding achieved with the addition of one equivalent of mercurial and the protection afforded ABP by preequilibration with 10^{-5} M L-arabinose.

Furthermore, the addition of L-arabinose to mercurial-labeled ABP perturbs the characteristic nitrophenol absorbance spectrum shown in Fig. 7 suggesting a ligand-

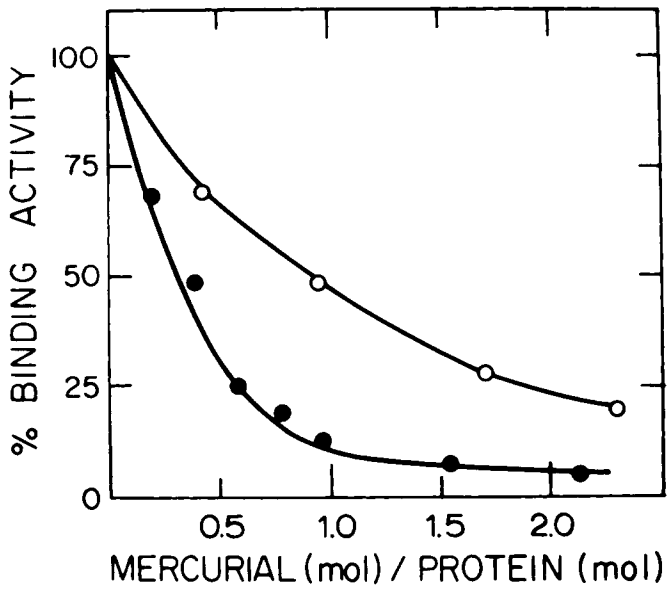


Fig. 6. Binding activity of MNP-modified arabinose-binding protein and protection by ligand. For the protection experiment, to arabinose-binding protein equilibrated (8 h) in the binding cells with L-[1-^{14}C] arabinose were added 5-μl aliquots of MNP. Following incubation overnight at 4°C, 100-μl aliquots were counted for radioactivity. In a second experiment, "sugar-free" protein was modified with MNP by adding 10-μl aliquots of the mercurial at appropriate concentrations to 0.44 ml of arabinose-binding protein and allowing the reaction to proceed overnight at 4°C. For the assay, double chambers separated by a dialysis membrane were filled with 200 μl of MNP-modified protein and 200 μl of L-[1-^{14}C] arabinose, equilibrated and counted as above. The percent of binding activity is expressed as the ratio of the apparent dissociation constant of duplicate samples to that of a control. The buffer was 10 mM Tris-HCl, pH 7.4. Arabinose-binding protein was 0.5 mg/ml and L-[1-^{14}C] arabinose was 10^{-5} M in both experiments. •) Arabinose-binding protein modified with MNP; ○) arabinose-binding protein reacted with MNP in presence of L-arabinose.

induced conformational change and/or direct interaction between the sugar and the phenolic mercurial.

A nondialyzable mercury-binding site was located by difference Fourier synthesis at position 64 in a tentative trace of the polypeptide chain. Moreover, the native electron density at site 64 is consistent with the assignment of a sulfur atom to this position. Although amino acid analysis had earlier indicated the presence of 2 half-cystine residues (14), the sequence determination showed one cysteine at position 64 (17), consistent with our assignment.

The likelihood that the structure of ABP may have been solved with bound L-arabinose prompted further examination of the 3.5- and 2.8-Å electron density maps, particularly in the region near the single essential cysteine residue. This examination shows in both maps the presence of an "extraneous" density near the thiol group which presently cannot be attributed to the protein molecule, but is of sufficient size and shape to be a

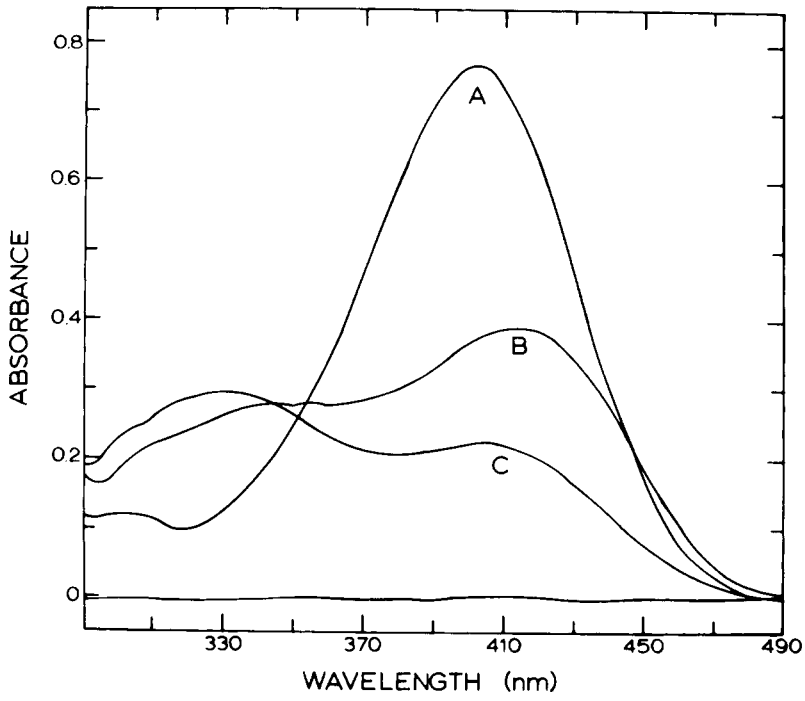


Fig. 7. Comparison of absolute spectra of 2-chloromercuri-4-nitrophenol in buffer (0.1 M Tris-HCl, pH 7.85) and bound to arabinose-binding protein in the absence and presence of L-arabinose. Arabinose-binding protein and MNP were 5×10^{-5} M. A) Absolute spectrum of MNP. B) Absolute spectrum of MNP bound to arabinose-binding protein. C) Absolute spectrum of MNP bound to arabinose-binding protein in the presence of 10^{-2} M L-arabinose.

sugar molecule. The presumed sugar molecule would be bound just to the "right" and within van der Waals distance of the cysteine residue (see Figs. 1F and 2). A native structure with sugar is consistent with the results of the soaking and crystallization experiments.

ABP is fully active in 60% 2-methyl-2, 4-pentanediol, suggesting that the failure of the difference Fourier maps to reveal bound sugar is not a result of the molecule being rendered inactive by its mother liquor. In addition, it can be seen from the packing diagram (Fig. 8) that the protein is readily accessible in the crystal lattice; various organomercurials, no smaller than the sugar substrate, are able to pentrate the crystal lattice and bind the essential thiol. Attempts to obtain "sugar-free" ABP crystals are currently underway.

DISCUSSION

The 3-dimensional structure of the L-arabinose-binding protein is the first structure to be determined in the family of transport proteins. The remarkable features of the structure may prove to be common to all periplasmic binding proteins and may be related to function.

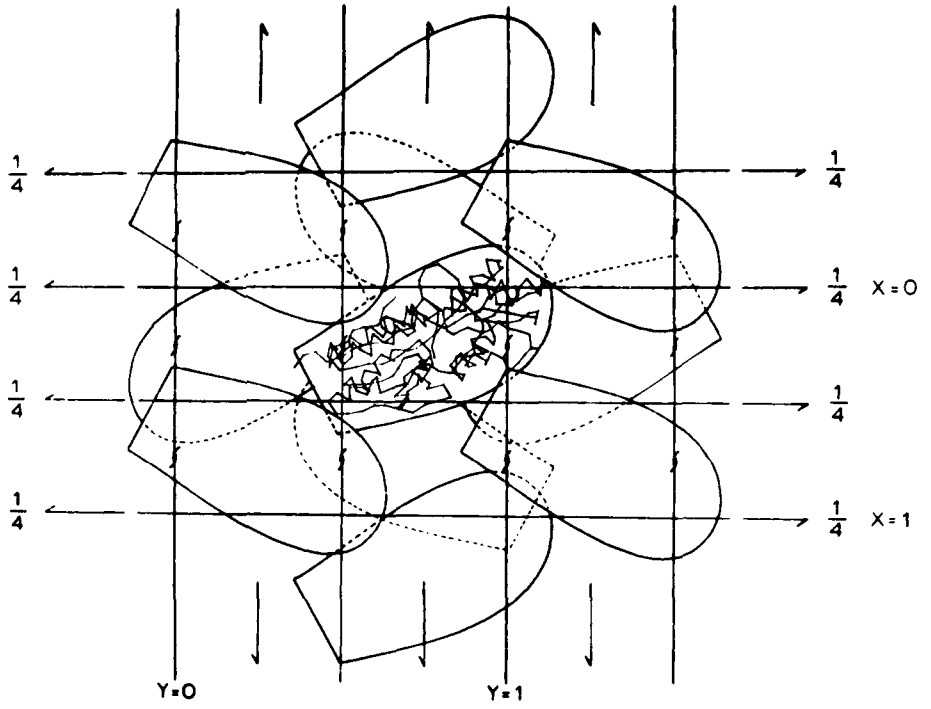


Fig. 8. Packing diagram for L-arabinose-binding protein. The equivalent positions for this space group are x, y, z; ½ −x, −z,½ + z; ½ + x, ½ − y, −z; and −x, ½ + y, ½ −z. The symmetry notation is taken from the International Tables for X-ray Crystallography (32).

All binding proteins share common functions – active transport and bacterial chemotaxis – and are dislodged from bacterial cells by mild osmotic shock as monomeric polypeptides. That binding proteins may be elongated is indicated by the axial ratio of 2:1 for ABP and the preliminary finding that the sulfate-binding protein (MW = 34,000) from S. typhimurium has an axial ratio of 4:1 (33). Furthermore, among subgroups of periplasmic binding proteins (i.e., sugar-binding proteins, amino acid-binding proteins, etc.) a number of observations indicates structural similarity as well. For instance, despite the fact that L-arabinose- and D-galactose-binding protein biosyntheses are controlled by distinctly different genes, *ara*C and *mglR* respectively (34, 35), the cross-reactivity of antibodies for each shows that the molecules share some regions of similar tertiary structure (36). Furthermore, we have shown that the dye 2′,4′,5′,7′-tetraiodofluorescein, a chromophoric probe for the nucleotide-binding site in a variety of proteins, binds to both L-arabinose-binding protein and D-galactose-binding protein. Circular dichroic measurements for the 2 sugar-binding proteins indicated similar secondary structure content (14, 31). However, the relative amount of helix predicted from circular dichroic measurements (10%) for ABP is incorrect on the basis of the structural results which indicate 35% helix. Perhaps this common molecular conformation has fullfilled a functional and/or structural requirement for binding proteins. It is not unreasonable to suggest that a 3-

dimensional structure is common to a given family of proteins; groups of proteins (i.e., cytochromes, dehydrogenases, and kinases) often have common structural features.

The existence of extensive secondary structure and the packing of this structure ("super secondary structure") is an aspect of the protein that could be relevant to the general properties of binding proteins. Perhaps this "super-structure" could account for the unusual stability of L-arabinose-binding protein and, by extension, the resistance of binding proteins in general to denaturation.

Our structural analysis has further revealed that the 2 distinct globular domains in the L-arabinose-binding protein are very similar. The orientation of the 2 domains is such that both super secondary structures can be related by an approximate local twofold axis (Fig. 3). A pseudo-dyad axis relating 2 domains has similarly been shown in rhodanese (37). A detailed quantitative comparison, using methods developed by Rossmann and coworkers (24), of the 2 domain structures in the binding protein, and with similar domains found in other, unrelated protein structures, is under further investigation. It is of interest to note that the arrangement of the secondary structure in both domains closely resembles the main structural features of the so-called "nucleotide-binding fold" domain found in some protein structures, notably the dehydrogenases and kinases (24), in the sense that all have a central pleated sheet structure consisting mainly of parallel strands and 2 helices on either side of the plane of the sheet, antiparallel to each strand. The significance of the presence of 2 structurally similar domains in L-arabinose-binding protein and especially the similarity of these domains with those found in unrelated enzymes remains a tantalizing mystery. Their occurrence in this binding protein supports the notion that they can be formed readily and could have evolved independently.

Insofar as it is known, the L-arabinose-binding site lies between 2 domains, but is predominantly associated with the P domain. The second domain (Q domain) may be involved to a large extent in binding to other components of the system. The elegant genetic studies of Ames and co-workers (11) on the histidine transport system have shown that the J protein (histidine-binding protein) has 2 distinct binding sites necessary for its function. One site is required for binding histidine, and the other is involved in a direct interaction with another protein component (the P protein) of the system which is presumed to be membrane bound. Nothing is known about the nature and location (relative to the histidine-binding site) of the other site.

The 2 domains in the binding protein could also conceivably facilitate an induced conformational change necessary for transport and bacterial chemotaxis. Since only 3 closely parallel polypeptide chain segments connect the 2 domains (the hinge region), any conformational change(s) could take place in the molecule by a mere shifting of one domain relative to the other. This conformational change could then affect the interaction with other components, the affinity for substrate, or both. The functional relationship between domains is especially important since the sugar-binding site is presumed to be between the domains, near the opening of the cleft.

For efficient transport, the sugar-binding site would be expected to be juxtaposed with other components of the system. This would imply protein:membrane surface interactions encompassing regions on the surface of the 2 domains of the protein and adjacent to the sugar-binding site. We have preliminary observations (currently under further study) which suggest that these specific regions of the binding protein surface may have some local charge properties necessary for hydrophilic interactions with the cytoplasmic membrane surface. This is based on the finding that the binding sites of all the heavy-atom derivatives used so far are unusually concentrated on these regions (or on one side) of the

protein molecule (Table I). This side of the molecule contains the presumed sugar-binding site. This remarkable distribution of the heavy-atom binding sites is not a result of the molecular packing of the crystal; Fig. 8 clearly indicates that all sides of the protein are accessible to the solvent. It is noteworthy that all these heavy-atom compounds are anionic, suggesting, in part, local concentrations of positively charged residues on the surface of the protein. These positively charged residues may suggest the possibility of protein interactions with either the phosphate moiety of the phospholipids adjacent to the other membrane-bound protein components or the negatively charged residues on the exposed part of the other protein component, or both. Such a hydrophilic interaction would be consistent with the observation that all binding proteins are loosely bound and easily removed by mild osmotic shock treatment.

Currently, the assignment of the sugar-binding site region of the L-arabinose-binding protein depends to some extent upon the identification of the single "essential" cysteine residue. However, the structure of the binding protein may have been solved with bound L-arabinose. To the "right" and within van der Waals distance of the cysteine residue is an "extraneous," nonprotein density peak which we attribute to a bound sugar molecule (Fig. 1F). The presumed sugar-binding site is in the cleft formed by the packing of the 2 domains, abutting the P domain. In addition to the cysteine at least 4 other residues, one of which appears to be a tryptophan, are in the vicinity of the sugar molecule. In most periplasmic binding proteins (e.g., L-arabinose- and D-galactose-binding proteins (14, 31); see also Ref. 2 and 4), the presence of substrates causes changes in the fluorescence of tryptophan residue(s). This fluorescence change may now be attributed more likely to changes in the microenvironment of the tryptophan as a result of direct interaction with the substrate rather than exclusively to conformational change, as has been frequently suggested.

ACKNOWLEDGMENTS

The contributions of former colleagues G. N. Phillips, Jr., V. K. Mahajan, and A. K. Q. Siu are gratefully acknowledged. This investigation was supported by a grant from the Robert A. Welch Foundation (C-581) and by NIH Grant GM-21371.

REFERENCES

1. Pardee AB, Prestidge LS, Whipple MB, Dreyfus J: J Biol Chem 241:3962, 1966.
2. Oxender DL: Annu Rev Biochem 41:777, 1972
3. Rosen BP, Heppel LA: In Lieve L (ed): "Bacterial Membranes and Walls." New York: Marcell Dekker, 1973, p 209.
4. Boos W: Annu Rev Biochem 43:123, 1974.
5. Pardee AB, Watanabe K: J Bacteriol 96:1049, 1968.
6. Nakane PK, Nichoalds GE, Oxender DL: Science 161:182, 1968.
7. Ames GF-L, Lever TL: Proc Natl Acad Sci USA 66:1096, 1970.
8. Adler J: Annu Rev Biochem 44:341, 1975.
9. Koshland DE Jr: In Jaernicke L (ed): "Biochemistry of Sensory Functions." Berlin: Springer-Verlag, 1974, p 133.
10. Adler J: J Gen Microbiol 74:77, 1973.
11. Ames GF-L, Spudich EN: Proc Natl Acad Sci USA 73:1877, 1976.
12. Ordal GW, Adler J: J Bacteriol 117:517, 1974.
13. Strange PG, Koshland DE Jr: Proc Natl Acad Sci USA 73:762, 1976.
14. Parsons RG, Hogg RW: J Biol Chem 249:3602, 1974.

15. Quiocho FA, Phillips GN Jr, Parsons RG, Hogg RW: J Mol Biol 86:491, 1974.
16. Brown CE, Hogg RW: J Bacteriol 111:606, 1972.
17. Hogg RW, Hermodson MA: J Biol Chem 252:5135, 1977.
18. Phillips GN Jr, Mahajan VK, Siu AKQ, Quiocho FA: Proc Natl Acad Sci USA 73:2186, 1976.
19. Wyckoff HW, Doscher M, Tsernoglou D, Allewell NM, Kelly DM, Richards FM: J Mol Biol 27:563, 1967.
20. Willis RC, Morris RG, Cirakoglu C, Schellenberg GD, Gerber NH, Furlong CE: Arch Biochem Biophys 161:64, 1974.
21. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ: J Biol Chem 193:265, 1951.
22. Dickerson RE, Kendrew JC, Strandberg BE: Acta Crystallogr 19:1188, 1961.
23. Lipscomb WN, Coppola JC, Hartsuck JA, Ludwig ML, Muirhead H, Searle J, Steitz TA: J Mol Biol 19:423, 1966.
24. Rossmann MG, Liljas A, Branden C-I, Banazak LJ: In Boyer PD (ed): "The Enzymes," 3rd Ed. New York: Academic Press, 1975, vol II, p 62.
25. Wasserman PM, Lentz PJ: J Mol Biol 60:509, 1971.
26. Jacobsberg LB, Kantrowitz ER, Lipscomb WN: J Biol Chem 250:9238, 1975.
27. Somerville LL, Quiocho FA: Biochim Biophys Acta 481:493, 1977.
28. McMurray CH, Trentham DR: Biochem J 115:913, 1969.
29. Miller DM: Fed Proc Fed Am Soc Exp Biol 35:1586, 1976.
30. Ellman GL: Arch Biochem Biophys 82:70, 1959.
31. Boos W, Gordon AS, Hall RE, Price HD: J Biol Chem 247:917, 1972.
32. Henry NFM, Lonsdale K (eds): "International Tables for X-ray Crystallography." London: Kynock Press, 1969.
33. Langridge R, Shinagawa H, Pardee AB: Science 84:585, 1970.
34. Hogg RW, Englesberg AS: J Bacteriol 100:423, 1969.
35. Lengeler J, Herman KO, Unsold HJ, Boos W: Eur Biochem 19:457, 1971.
36. Parsons RW, Hogg RW: J Biol Chem 294:3608, 1974.
37. Bergsman J, Hol WGJ, Jansonius JN, Kalk KH, Ploegman JH, Smit JDG: J Mol Biol 98:637.