Articles

Structure-Based Exploration of the Ganglioside GM1 Binding Sites of *Escherichia coli* Heat-Labile Enterotoxin and Cholera Toxin for the Discovery of Receptor Antagonists†

Wendy E. Minke,‡ Claudia Roach,§ Wim G. J. Hol,‡,§ and Christophe L. M. J. Verlinde*,‡

Department of Biological Structure, Biomolecular Structure Center, and Howard Hughes Medical Institute, University of Washington, Box 357742, Seattle, Washington 98195

Received November 6, 1998; Revised Manuscript Received March 4, 1999

ABSTRACT: Ganglioside GM1 is the natural receptor for cholera toxin (CT) and heat-labile enterotoxin (LT), which are the causative agents of cholera and traveler's diarrhea, respectively. This observation suggests that small molecules interfering with this recognition process may prevent entry of the toxins into intestinal cells, thereby averting their devastating effects. Here, the terminal sugar of ganglioside GM1, galactose, was chosen as a lead in designing such receptor antagonists. Guided by the experimentally determined binding mode of galactose, we selected a "substructure" for searching the Available Chemicals Database, which led to the purchase of 35 galactose derivatives. Initial screening of these compounds in an LT ELISA revealed that 22 of them have a higher affinity for LT than galactose itself. A structurally diverse subset of these galactose derivatives was selected for determination of IC_{50} values in the LT ELISA and IC_{50} values in a CT assay, as well as for the determination of K_d's using the intrinsic fluorescence of LT. The best receptor antagonist found in this study was *m*-nitrophenyl α-galactoside with an IC_{50} of 0.6 (2) mM in the LT ELISA and 0.72 (4) mM in the CT assay, 100-fold lower than both IC_{50} values of galactose. Careful analysis of our binding data and comparison with crystal structures led to the derivation of correlations between the structure and affinity of the galactose derivatives. These characteristics will be used in the design of a second round of LT and CT receptor antagonists.

Heat-labile enterotoxin from enterotoxigenic *Escherichia coli* (LT)[1] and the closely related cholera toxin (CT) are the major virulence factors of traveler's diarrhea and cholera, respectively. Both toxins, belonging to the AB_5 toxin family (*1*), exploit the intrinsic trafficking mechanisms of the host cell to gain access to the cytosol (*2*), where they exert their detrimental activity (*3*). After being released into the jejunum by the bacteria, the toxins recognize the oligosaccharide portion of GM1 ganglioside molecules on the surface of the epithelial cells through their B-subunits (*4*, *5*). Subsequently, the entire AB_5 toxin is internalized in noncoated vesicles and transported to the Golgi (*6*). There, most of the A-subunit, usually termed A1, is separated from the rest of the toxin through reduction of a disulfide bridge (*7*). A1 then follows a retrograde route into the endoplasmatic reticulum (*6*) and translocates to the cytosol, likely via the ER-associated protein degradation pathway (*8*). There it carries out its toxic activity, the irreversible ADP ribosylation of $G_{s\alpha}$ leading to persistent activation of adenylate cyclase (*9*, *10*). In turn, the high levels of cyclic AMP inhibit sodium reabsorption and activate chloride excretion (*11*), resulting in a massive efflux of fluids from the intestinal cells.

These fascinating cellular events are unfortunately the source of tremendous human misery. It was estimated that there are annually more than 1 billion cases of diarrhea due to enterotoxigenic *E. coli*, resulting in 1 million deaths (*12*). Cholera is endemic in 80 countries and affects several hundreds of thousands of people with a death toll of 120 000 annually (*13*). Our goal is to design and discover agents that block each of the steps in the "bioactivation" of the toxins on the basis of their three-dimensional structure. Here we discuss intervening with the first step, the recognition process between the GM1 receptor and heat-labile enterotoxin from *E. coli* (LT) or cholera toxin (CT).

LT and CT bind specifically to the oligosaccharide portion of ganglioside GM1, whose ceramide moiety is firmly embedded in the outer layer of the plasma membrane. The oligosaccharide (OS) is a branched pentasaccharide with the

† This work was supported in part by National Institutes of Health Grant R29GM54618.

* Corresponding author. Fax: (206) 685-7002. E-mail: verlinde@gouda.bmsc.washington.edu.

‡ Biomolecular Structure Center.

§ Howard Hughes Medical Institute.

[1] Abbreviations: ACD, Available Chemicals Directory; BSA, bovine serum albumin; CT, cholera toxin; esd, estimated standard deviation; GD1b, Galβ1−3GalNAcβ1−4(Neu5Acα2−8Neu5Acα2−3)Galβ1−4Glcβ1-ceramide; GM1, Galβ1−3GalNAcβ1−4(Neu5Acα2−3)Galβ1−4Glcβ1-ceramide; Gal, galactose; GalNAc, 2-acetamido-2-deoxy-galactoside; Glc, glucose; LT, heat-labile enterotoxin from *E. coli*; Neu5Ac, *N*-acetylneuraminic acid; OPD, *o*-phenylenediamine; OS, oligosaccharide; PBS, phosphate-buffered saline.

10.1021/bi982649a CCC: $18.00
Published on Web 04/17/1999

sequence Galβ1−3GalNAcβ1−4(Neu5Acα2−3)Galβ1−4Glc, with all sugars in the D configuration; the glucose moiety is covalently linked to ceramide. Structural insight into the binding process comes from crystallographic studies of (I) the toxins in the absence of ligand [LT (*14*, *15*) and CT (*16*)], (II) LT in the presence of fragments of the GM1-OS [galactose (*17*), lactose (*18*), and Galβ1−3GalNAc (*19*)], and (III) CT in the presence of GM1-OS (*20*). The following facts were established. (I) The structures of the GM1 binding B pentamers of CT and LT are very similar; their backbones can be superimposed within 0.5 Å, reflecting the fact that their sequences are 80% identical. (II) The sugars bind between loops originating from the orthogonally packed β-sheets of the B-subunit, a fold now commonly termed the OB-fold (*21*). (III) One of the loops, formed by residues 54−60, becomes more ordered upon sugar binding. (IV) The GM1-OS binds to CT mainly via its terminal sugars galactose and sialic acid, the remaining sugars acting essentially as linkers. (V) The binding mode of the terminal galactose is virtually identical in all complexes, whether bound to CT or LT, and even when the galactose is attached to glucose via a 1−4 linkage as in lactose instead of the 1−3 linkage of GM1. (VI) The porcine LT structure is an excellent model for human LT because, of the four residues out of 103 that differ between them, the only residue that is involved in GM1 recognition interacts via its backbone N atom (Arg13 in porcine and His13 in human LT). Hence, from the structural studies it appears that D-galactose and sialic acid might be leads for the design of "B-subunit antagonists" for both toxins.

Competitive binding experiments with GM1-derived gangliosides have demonstrated the importance of galactose and sialic acid for GM1 affinity. GM2, which lacks the terminal galactose, bound 16 times weaker than GM1 to CT and 119 times weaker to LT; asialo-GM1, devoid of the sialic acid moiety, showed no detectable binding to CT or LT at concentrations that were 320 and 160 times higher than the GM1 IC_{50}, respectively (*22*). Whereas the affinity of the GM1-OS for the toxin is strong (K_d = 952 nM at 37 °C, as determined by isothermal titration calorimetry) (*23*), monosaccharide fragments are poor ligands. GalNAc, Glc, and sialic acid did not affect CT binding to GM1 on liver membranes at concentrations of up to 40 mM; only galactose reduced the extent of binding, by 20% (*24*). Competitive binding experiments with B-type erythrocytes gave similar results (*25*). In view of the inactivity of sialic acid, we decided to focus initially on galactose as a lead for drug design.

EXPERIMENTAL PROCEDURES

ACD3D Database Computer Screening. Two-dimensional substructure searches were conducted in the Available Chemicals Directory, ACD3D 95.1 (Molecular Design Ltd., San Leandro, CA), consisting of 164 566 compounds. The software used for searching was the program ISIS (Molecular Design Ltd.) (*26*).

Protein and Chemicals. Porcine LT holotoxin was expressed from plasmid EWD299 using the thermoinducible overexpression vector pROFIT in *E. coli* strain MC1061 (*27*). The protein purification protocol of Uesaka et al. (*28*) was followed except that the clarified cell lysate was subjected to a 30% ammonium sulfate precipitation in 20 mM Tris-HCl buffer (pH 7.5) prior to affinity chromatography on immobilized D-galactose. The purity was estimated to be at least 95% on the basis of silver-stained SDS−PAGE gels.

Porcine LT B_5 was obtained as follows. The sequences encoding the porcine LT B_5 were excised from plasmid EWD299 as an *Eco*RI−*Pst*I fragment and ligated into the cloning site of the pROFIT plasmid to generate plasmid pROFIT-LTB. The *E. coli* strain MC1061 was used as an expression host, and expression and purification protocols of the holotoxin were followed. The purity was estimated to be at least 95% on the basis of silver-stained SDS−PAGE gels.

Anti-LT B monoclonal antibody mAB 118-87 was a kind gift from T. Hirst (University of Bristol, Bristol, U.K.). Commercially obtained materials were as follows: CT B_5 horseradish peroxidase conjugate (List Biological Laboratories, Campbell, CA), IgG horseradish peroxidase conjugate (Boehringer Mannheim, Indianapolis, IN), GD1b (Matreya, Pleasant Gap, PA), GM1-OS (Toronto Research Chemicals, Toronto, ON), galactose and methyl 2-acetamido-2-deoxy-α-D-galactoside (Calbiochem, San Diego, CA), and all other galactose derivatives, including GM1 (Sigma, St. Louis, MO).

LT GD1b ELISA. Microtiter plates (C96 Maxisorp; Fisher Scientific, Pittsburgh, PA) were incubated at 37 °C for 16 h with 100 μL of 2 μg/mL ganglioside GD1b dissolved per well in phosphate-buffered saline (pH 7.2) containing 150 mM NaCl and 10 mM potassium phosphate (PBS). Unattached ganglioside was removed by washing the wells two times with PBS. Additional binding sites on the plate surface were blocked by incubating the wells with 200 μL of a 1% (w/v) bovine serum albumin (BSA)−PBS solution for 30 min at 37 °C and then washing them with PBS three times. The test samples, diluted in 0.1% BSA−PBS, were added in 100 μL volumes per well and incubated for 30 min at room temperature. Unbound toxin was removed by washing three times with PBS. Toxin bound to GD1b was then revealed by the following sequence: (I) incubation at room temperature with 100 μL of anti-LT-B murine monoclonal antibody 118-87, diluted 1:1500 in 0.1% BSA−PBS; (II) three wash steps with 0.05% Tween 20−PBS; (III) incubation at 37 °C for 1 h with 100 μL of anti-mouse IgG conjugated with horseradish peroxidase, diluted 1:3000 in 0.1% BSA−0.05% Tween 20−PBS; (IV) three wash steps with 0.05% Tween 20−PBS; (V) incubation with 100 μL of freshly made *o*-phenylenediamine (OPD) solution (10 mg of OPD, 5 mL of 0.1 M sodium citrate, 5 mL of 0.1 M citric acid, and 4 μL of 30% hydrogen peroxide) for 30 min at room temperature; and (VI) recording of the A_{450} on a Molecular Devices V-max ELISA microtiter plate reader.

Test samples consisted of 0.2 μg/mL porcine LT toxin preincubated with potential ligand for 2 h at room temperature. All experiments were carried out in quadruplicate and validated against a concentration gradient of 0, 0.1, 0.2, and 0.3 μg/mL toxin. IC_{50} values were calculated from at least five different concentrations of competitive ligand by nonlinear regression as described previously (*29*) with the statistical package S-PLUS (Mathsoft, Inc., Cambridge, MA). For each ligand, the IC_{50} was determined at least two times. In case the IC_{50} values of the first two determinations were not within 25% of one another, the IC_{50} was determined at

least three times. The reported IC_{50} of a ligand is the weighted average (weight = $1/esd^2$) of the IC_{50} values from its different determinations.

CT GD1b Enzyme-Linked Adhesion Assay. Microtiter plates were coated with ganglioside GD1b and subsequently blocked with BSA as described for the LT GD1b ELISA. Then, they were washed three times with 0.05% Tween 20−PBS. The test samples, diluted in 0.1% BSA−0.05% Tween 20−PBS, were added in 100 μL volumes per well and incubated for 30 min at room temperature. Unbound toxin was removed by washing three times with 0.05% Tween 20−PBS. Toxin bound to GD1b was then revealed by the following subsequent steps: (I) incubation with 100 μL of freshly made *o*-phenylenediamine (OPD) solution (10 mg of OPD, 5 mL of 0.1 M sodium citrate, 5 mL of 0.1 M citric acid, and 4 μL of 30% hydrogen peroxide) for 30 min at room temperature and (II) recording of the A_{450} on a Molecular Devices V-max ELISA microtiter plate reader.

Test samples consisted of 6 ng/mL CT B_5 horseradish peroxidase conjugate preincubated with potential ligand for 2 h at room temperature. All experiments were carried out in quadruplicate and validated against a concentration gradient of 0, 0.3, 0.6, and 0.9 μg/mL toxin peroxidase conjugate. IC_{50} values were calculated as described for the LT ELISA experiments.

Fluorescence Spectroscopy and K_d Determination. Prior to titration, the porcine LT B_5 concentration was determined with a Beckman 7300 Amino Acid Analyzer following the procedure of Hamilton (*30*) after extensive dialysis against the storage buffer [50 mM Tris-HCl (pH 7.4) containing 1 mM EDTA, 3 mM NaN_3, and 200 mM NaCl]. An ϵ_{280} of 14.0 (4) $\times 10^3$ M^{-1} cm^{-1} was established. Binding isotherms for the formation of the complex between porcine LT B_5 and ligand were determined by fluorescence spectroscopy at room temperature using a SPEX Fluorolog-2 F111A spectrofluorimeter (Yvon Jobin, Edison, NJ). The intrinsic fluorescence of Trp88 of the protein was recorded by registering the emission between 325 and 350 nm after excitation at 282 nm. Spectra were obtained at a scanning rate of 2 s per 1 nm wavelength increment. The effective path length was 10 mm.

Titrations were performed through gradual dilution of the ligand with the protein concentration fixed at 20 μM. All measurements were taken after an incubation time of 2 min. No time-dependent change in the signal was observed when the incubation time was increased.

For determination of the binding constants, the following integral was calculated for each binding isotherm, in which I_λ is the fluorescence intensity in CPS at wavelength λ:

$$S = \sum_{\lambda=325}^{\lambda=350} I_\lambda \lambda \qquad (1)$$

Integrals were determined in absence of ligand, at ligand concentrations [L], and at a ligand concentration which saturates the protein. These integrals are termed S_{free}, S_L, and S_{bound}, respectively. The integrals were fitted to eq 2, for one binding mode, and to eq 3 for two binding modes in the case of melibionic acid. Fitting was performed using the statistical package S-PLUS (Mathsoft, Inc.) on the basis of nonlinear regression analysis (*31*).

$$S_L = S_{free} - [L](S_{free} - S_{bound})/(K_d + [L]) \qquad (2)$$

$$S_L = S_{free2} - [L](S_{free2} - S_{free1})/(K_{d1} + [L]) - [L](S_{free1} - S_{bound})/(K_{d2} + [L]) \qquad (3)$$

RESULTS

ACD3D Database Computer Screening. To exploit D-galactose as a lead for receptor antagonist design, opportunities for substitution were analyzed in detail. The sugar rests on top of the Trp88 side chain and is involved in a complex hydrogen bond network between its hydroxyls and four protein side chains (Figure 1). Obviously, O1 is a candidate for substitution as it is involved in the linkage to GalNAc in the natural ligand GM1. Also, substitution of O2 should be allowed as one of its partners is a water molecule that could be displaced. Substitution of the O3 and O4 hydroxyls is likely to be unsuccessful as each of them is functioning as both a hydrogen bond donor and an acceptor with the protein; i.e., substitution would lead to unsatisfied hydrogen bond donors and acceptors. Alternative epimers at C3 and C4 are not allowed since O3 and O4 would bump into Trp88 (distances of <2.6 Å). However, there is no obvious reason alternative epimers should be precluded at C1 or C2. Substitutions at O6 were disregarded in this first round of drug design for the sake of simplicity as no less than two conserved water molecules bind to this atom. In summary, galactose without its 1- and 2-hydroxyl groups was considered the most logical substructure with which to search for commercially available galactose derivatives.

A search of the ACD3D database revealed 235 available compounds with the desired substructure. After elimination of presumably insoluble, high-molecular mass (M_r > 600 Da), and prohibitively expensive compounds, 35 were purchased for toxin affinity testing.

LT GD1b ELISA. Although the GM1 ELISA described by Svennerholm and Holmgren (*34*) is an excellent assay for screening tightly binding ligands, it is less suitable for low-affinity ligands because of solubility problems at high ligand concentrations. Therefore, we developed a new competitive ELISA based on ganglioside GD1b instead of GM1. GD1b, which has an extra sialic acid linked to the glycerol moiety of GM1, binds to the same binding site of the toxin but with an affinity that is 11 times lower (*35*). Thus, the new assay is more sensitive and allows for the detection of ligands with weaker affinities than the original GM1 ELISA.

The LT GD1b ELISA was used for initial screening of the inhibitors, with the inhibitors at a concentration of either 20 mM or at the limits of their solubility. Twenty-two of the 35 tested galactose derivatives showed an increased extent of inhibition of toxin binding compared to galactose (Table 1). Subsequently, IC_{50} values were determined for eight of these strong inhibitors in addition to IC_{50}'s for galactose and GM1-OS (Table 1). Strong inhibitors were not further pursued when they were similar to compounds for which the IC_{50} had already been determined (Table 1). Five compounds were not soluble in the LT GD1b ELISA solutions: 4-methyl umbelliferyl-α-D-galactoside, 4-methyl umbelliferyl-β-D-galactoside, naphthol AS-bi-β-D-galactoside, β-naphthyl β-D-galactoside, and resorufin β-D-galactoside. Therefore, they could not be assayed.

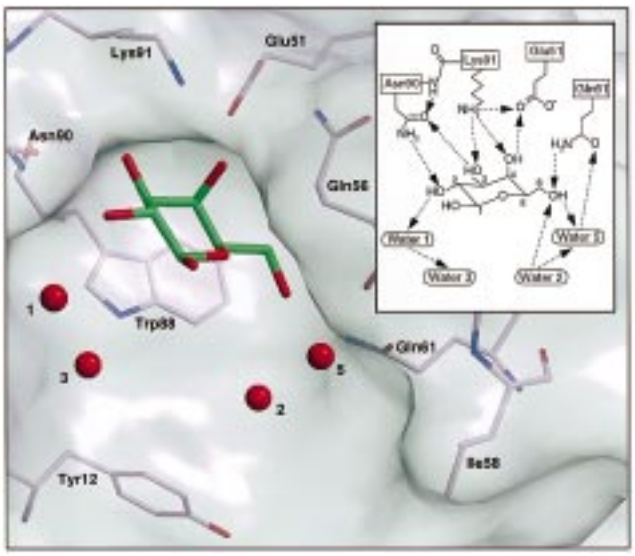

FIGURE 1: Binding mode of galactose in the GM1 binding site of pLT as observed in its crystal structure (*17*). In the inset is depicted the hydrogen bonding network between galactose and LT. An identical binding mode for galactose has been observed in the crystal structures of pLT−lactose (*18*), pLT−Galβ1−3GalNAc (*19*), and CT−GM1-OS (*20*). The molecular surface shown was calculated in GRASP (*32*). The figure was produced using Raster3D (*33*).

Seven galactose derivatives for which the IC_{50} values were determined were better ligands than galactose (Table 1). As expected, the toxin accommodated several different substitution patterns: a β1−O linkage (as in GM1), a β1−S linkage, and an α1−O linkage. The best ligand, *m*-nitrophenyl α-D-galactoside, was 100 times more potent than galactose. This is significantly better than galactose, but there is still a long way to go since galactose binds 6 million times weaker than GM1-OS to LT.

CT GD1b Enzyme-Linked Adhesion Assay. The simpler variant assay of the ELISA based on a brief report by Idota et al. (*36*) appeared to work well. All ligands for which the IC_{50} values were determined in the LT ELISA were tested against CT (Table 1). Their affinities for CT were very similar to the LT results (Table 1 and Figure 2); R^2 is 0.90 for 10 data points.

Fluorescence Spectroscopy and K_d Determination. Binding of GM1 to CT leads to a 12 nm blue shift in the emission spectrum as reported by Fishman et al. (*37*). A similar shift was detected in the emission spectrum of LT B_5 upon binding of galactose, phenyl ethylgalactoside, melibionic acid, or thiodigalactoside. In addition, a reduction in the emission intensity was observed (Figure 3). LT B_5 was therefore titrated with these four galactose-containing compounds to determine the binding constants. Titration curves were constructed (Figure 4), from which the following K_d values for LT B_5 were obtained: 13 (1) mM for galactose, 2.6 (1) mM for phenyl ethylgalactoside, and 0.36 (7) mM for thiodigalactoside. The titration curve for melibionic acid suggested the presence of two binding modes with a K_{d1} of 5 (1) mM and a K_{d2} of 0.13 (4) M (see also the Discussion).

DISCUSSION

In this study, we succeeded in identifying several galactose derivatives which are inhibitors of LT and CT binding in GD1b assays. Thirty compounds were selected from the ACD to be tested for inhibition of LT binding to GD1b. Eighteen of these compounds were better inhibitors than galactose, which gives our approach a success rate of 60%. The best inhibitor, *m*-nitrophenyl α-galactoside, had an affinity for LT and CT that was 100 times higher than that of galactose.

It was gratifying that the results obtained in the different assays correlate well with one another. There is a linear relationship between the IC_{50} values obtained with the LT ELISA and the IC_{50} values obtained with the CT assay (Table 1 and Figure 2). Also, there appears to be a linear relationship between the IC_{50} values obtained with the LT ELISA and the binding constants. Therefore, the screening of potential ligands with the LT ELISA is a good method for obtaining compounds that inhibit binding of both LT and CT to GD1b. In addition, the results obtained in the LT ELISA are a good indication for how strong the compounds bind to LT B_5.

Our results obtained with GD1b-coated microtiter plates (LT ELISA and CT assay) also agree well with the results obtained using whole cells. Sugii (*25*) reported inhibition of ^{125}I-labeled CT to neuraminidase-treated human type B erythrocytes; the inhibition by lactose was slightly weaker than the inhibition by galactose, while inhibition by methyl α-galactoside was stronger. Our data agree with the observations of Sugii (*25*) (Table 1). Cuatrecasas (*24*) studied the inhibition of ^{125}I-labeled CT to liver cell membranes. This author reported that the inhibition by methyl α-galactoside, methyl β-galactoside, methyl β-thiogalactoside, or isopropyl β-thiogalactoside was slightly better than the inhibition by galactose. The inhibition by thiodigalactoside or *p*-aminophenyl β-galactoside was even stronger. In addition, *o*-nitrophenyl galactoside did not exhibit any inhibition at all. These data also correspond well with our results. The only discrepancy concerns lactose. Cuatracasas (*24*) did not notice any inhibition, while we found that the inhibition by lactose was similar to the inhibition by galactose. However, our result is in accordance with the observations of Sugii (*25*). The crystal structure of the pLT−lactose complex confirms these findings; the galactose moiety is bound as in all other complex structures, and the glucose moiety does not interact with the protein (*18*).

The intrinsic fluorescence of LT B_5 and CT B_5 arises from Trp88, the only tryptophan residue in the B-subunits. Crystal structures of complexes with either CT or LT show that Trp88 is important for ligand binding (*17*−*20*). In all cases, galactose-containing ligands have their galactose moieties positioned against the side chain of Trp88. This stacking creates a more hydrophobic environment for the indole ring, which results in a blue shift in the fluorescence emission spectrum when GM1 binds to CT (*37*). Binding of galactose, lactose, or fucose to CT B_5 also yields a blue shift in the emission spectrum, in addition to a decrease in intensity (*38*). We observed similar shifts and intensity changes in the emission spectrum of LT B_5 upon binding of galactose, phenyl ethylgalactoside, melibionic acid, and thiodigalactoside. Unfortunately, no fluorescence emission of LT could be observed in the presence of *m*-nitrophenyl α-galactoside, since nitrobenzene absorbs at 322 nm. Titration studies with galactose resulted in K_d values on the same order of magnitude; the K_d of CT for galactose is 40 mM (*38*), and

Table 1: Inhibition of Toxin Binding to GD1b-Coated Microtiter Plates by D-Galactose and Its Derivatives[a]

inhibitor	Gal substituent	[inhibitor] (mM)/ % inhibition	IC_{50}(pLT) (mM)	IC_{50}(CT) (mM)
(a) β1-Gal substitution				
α-naphthyl β-galactoside	*O*-1-naphthyl	5/0		
4-methyl umbelliferyl-β-lactoside	*O*-4Glc1−4methylumbelliferyl	3/0		
5-bromo-3-indolyl-β-galactoside	*O*-3-indolyl-5-bromo	1/0		
o-nitrophenyl β-thiogalactoside	*S*-*o*-nitrophenyl	<10/15		
2-(β-galactosidoxy)naphthol AS-LC	*O*-2-naphthalene-3-carboxylic acid (2,5-dimethoxy-4-chloroanilide)	<10/15		
galactose	−	20/22	58 (8)	45 (4)
lactose	*O*-4Glc	20/22	6 (2) × 10^1	46.0 (4)
lactobionic acid	*O*-4-gluconic acid	14/28		
p-aminophenyl β-lactoside	*O*-4Glc1-*p*-aminophenyl	20/30		
lactulose	*O*-4-fructosyl	20/54	15 (3)	27 (6)
4-*O*-β-galactosylmannoside	*O*-4-mannoside	20/60		
p-aminophenyl β-galactoside	*O*-*p*-aminophenyl	20/70		
methyl β-galactoside	*O*-CH_3	20/74		
p-aminobenzyl β-thiogalactoside	*S*-CH_3-*p*-aminophenyl	20/75		
methyl β-thiogalactoside	*S*-CH_3	20/80		
phenyl β-galactoside	*O*-phenyl	20/85		
phenethyl β-galactoside	*O*-$(CH_2)_2$-phenyl	20/85	9 (2)	14 (3)
isopropyl β-thiogalactoside	*S*-$CH(CH_2)_2$	20/86		
hydroxyethyl β-thiogalactoside	*S*-$(CH_2)_2$-OH	20/90		
thiodigalactoside	*S*-1-galactosyl	20/95	4 (1)	4 (1)
N-(ϵ-aminocaproyl)-β-galactosylamine	NH-$CO(CH_2)_5NH_2$	20/99	2.1 (5)	3.2 (2)
GM1 oligosaccharide	*O*-3GalNAcβ1−4(Neu5Acα2−3)Galβ1−4Glc		1 (1) × 10^{-5}	14 (4) × 10^{-6}
(b) α1-Gal substitution				
α-naphthyl α-galactoside	*O*-1-naphthyl	5/0		
o-nitrophenyl α-galactoside	*O*-*o*-nitrophenyl	20/25		
p-nitrophenyl α-galactoside	*O*-*p*-nitrophenyl	20/60		
methyl α-galactoside	*O*-CH_3	20/77		
melibionic acid	*O*-6-gluconic acid	20/85	11 (2)	5.2 (1)
p-aminophenyl α-galactoside	*O*-*p*-aminophenyl	20/95	4.8 (5)	12 (2)
m-nitrophenyl α-galactoside	*O*-*m*-nitrophenyl	20/100	0.6 (2)	0.72 (4)
(c) β1-GalNAc substitution				
p-nitrophenyl 2-acetamido-2-deoxy-β-galactoside	*O*-*p*-nitrophenyl	<10/15		
(d) α1-GalNAc substitution				
methyl 2-acetamido-2-deoxy-α-galactoside	*O*-CH_3	20/15		
benzyl 2-acetamido-2-deoxy-α-galactoside	*O*-CH_2-phenyl	<10/25		

[a] The values in this table differ slightly from the ones reported by Merritt et al. (*39*). They reported preliminary data.

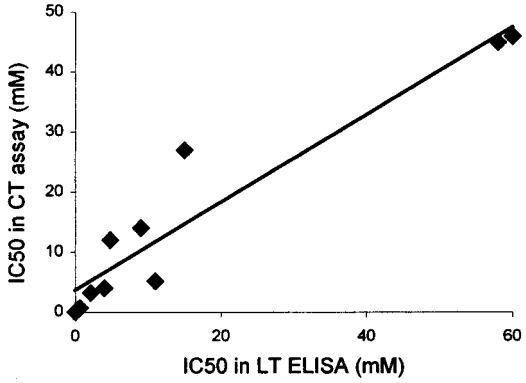

FIGURE 2: Correlation between the results obtained in the LT ELISA and the CT assay. $R^2 = 0.90$.

the K_d of LT for galactose is 13 (1) mM (this study). All available solution studies point to the fact that the galactose derivatives studied bind with the galactose moiety essentially at the same position near Trp88.

To understand the improved binding of the galactose derivatives found in this study, our colleagues cocrystallized four of our compounds with pLT (*39*). The resultant crystal structures provide a platform for discussing interactions that explain why these compounds exhibit improved binding compared to that of galactose by itself. For all four ligands, the galactose moiety was observed to bind in a manner virtually identical to that previously seen in other complexes

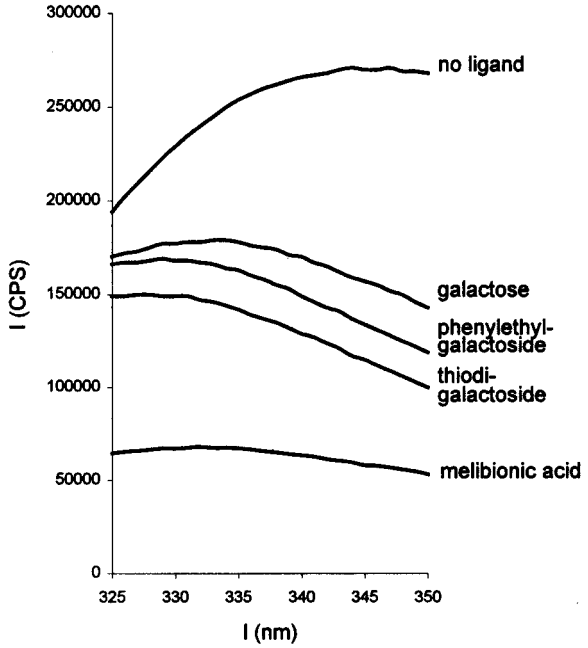

FIGURE 3: Emission spectra of LT B_5 in the absence of ligand and in the presence of galactose, phenylethylgalactose, thiodigalactose, or melibionic acid (LT B_5 concentration of 20 μM and ligand concentration of 100 mM).

(Figure 1), in agreement with the studies in solution. The effects of the galactose substituents are discussed below in

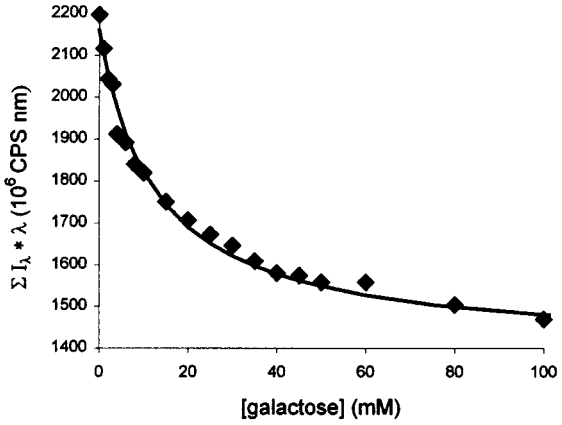

FIGURE 4: Effect of galactose on the fluorescence spectrum of LT B_5. The diamonds represent experimental data, while the line represents the fitted function. The fluorescence spectra of LT B_5 were scanned in the presence of decreasing amounts of galactose as described in Experimental Procedures (LT B_5 concentration of 20 μM). The K_d of galactose is 13 (1) mM.

a systematic way in an effort to obtain insight into the reasons for the range in affinity that is observed.

The crystal structure of the pLT−lactulose complex reveals two hydrogen bonds between the fructofuranose moiety and crystallographically conserved waters (Figure 5A). Also, this galactose substituent provides an additional surface of interaction with the protein of 110 Å^2. The interaction surface is the buried solvent accessible surface of the protein and ligand as calculated with the program GETAREA (*41*) using a probe radius of 1.4 Å, and the values are accurate within 5%. Together with the two hydrogen bonds, this additional interaction surface contributes to the 4-fold increase in the affinity of lactulose compared to that of galactose.

The melibionic acid crystal structure at 2.8 Å did not show clear density for the gluconic acid moiety (*39*). However, recently a new 1.45 Å structure of the same complex was determined which showed good density for the gluconic acid moiety in three GM1 binding sites (E. A. Merritt, personal communication). In two of the three sites, the carboxylate of the gluconic acid moiety forms a hydrogen bond with the backbone amide of Arg13, with a distance of 2.8 Å between the carboxylate oxygen and the backbone nitrogen, and a C═O···N angle of 128° (Figure 5B). In addition, two hydrogen bonds are formed with two waters. In the third site, the gluconic acid moiety only interacts with two waters (Figure 5B) and no direct interaction with the protein occurs. The two binding modes for the gluconic acid moiety differ up to 4.2 Å. These two binding modes observed in the 1.45 Å structure agree with our fluorescence titration data, which also suggested two binding modes for melibionic acid. It is likely that the hydrogen bond to the backbone amide of Arg13 is the main contributor to the 6-fold increased extent of binding of melibionic acid compared to that of galactose. Interestingly, it mimics the same hydrogen bond that is formed by the sialic acid carboxylate of GM1 and the backbone amide of residue 13.

The crystal structure of the complex of pLT B_5 with thiodigalactoside reveals two reasons why this compound binds 15 times better to the B-pentamer than galactose (Figure 5C). First, the second galactose moiety contributes an additional surface of interaction with the protein of 190 Å^2 compared to that of galactose. Second, O6 forms two hydrogen bonds to crystallographically conserved waters.

Finally, the crystal structure of the complex between pLT B_5 and *m*-nitrophenyl α-galactose provides insight into why this ligand has 100-fold improved affinity compared to galactose (Figure 5D). First, the *m*-nitrophenyl moiety contributes an additional surface of interaction with the protein of 180 Å^2 compared to that of galactose. Second, the nitro group of *m*-nitrophenyl α-galactose forms a hydrogen bond to the backbone amide of Gly33, with a distance of 2.8 Å between the nitro group oxygen and the backbone nitrogen, and an N−O···N angle of 113°. This hydrogen bond is likely to be relatively strong, since the nitro O···H amide distance is 0.35 Å shorter than the average hydrogen bond distance of 2.30 (1) Å for nitro O···H bonds (*42*). Perhaps most importantly, for this hydrogen bond to form, a crystallographically conserved water molecule must be displaced. The release of this water molecule into bulk solvent leads to an increase in entropy. Enthalpically, the interactions between the nitro group and pLT are similar to the interactions between the conserved water and the protein; in both cases, a strong hydrogen bond is formed to Gly33 and an additional, weaker hydrogen bond to another conserved water molecule. Therefore, the improved binding of *m*-nitrophenyl α-galactoside to LT is likely mainly entropy driven.

One of the main goals of this study was to distinguish features of galactose substituents that affect ligand binding in the GM1 binding site of LT and CT. In addition, we wanted to explain the effects of the functional groups on a structural level (Figure 6). One of the first considerations was whether to use α- or β-galactose derivatives (see label 1 in Figure 6). Although this is highly dependent on the nature of the galactose substituent, our results indicate that α-substituents are generally favored over β-substituents. This is reflected in the observations that binding of *p*-aminophenyl α-galactose is favored over binding of *p*-aminophenyl β-galactose, and binding of methyl α-galactoside over binding of methyl β-galactoside (Table 1). In addition, our best inhibitor has an α-substituent, *m*-nitrophenyl α-galactoside. In Figures 1 and 5, it can be seen that O1 of galactose at the α-anomeric position points toward the protein surface, whereas O1 at the β-anomeric position points toward the solvent. Therefore, a substituent at the α-position will tend to bury more surface area than a substituent at the β-position.

Second, for the β-substituents, it seems that a bridging sulfur is favored over a bridging oxygen (see label 2 in Figure 6), which is reflected in the observation that methyl β-thiogalactoside inhibits better than methyl β-galactoside. In addition, thiodigalactoside is one of our best inhibitors.

Third, an *N*-acetyl group on C2 of galactose (see label 3 in Figure 6) reduces the affinity of LT for galactose considerably as is shown by the very weak inhibition by methyl 2-acetamido-2-deoxy-α-galactoside, *p*-nitrophenyl 2-acetamido-2-deoxy-β-galactoside, and benzyl 2-acetamido-2-deoxy-α-galactoside (Table 1). Especially noticeable is the 5-fold decrease in the affinity upon addition of an *N*-acetyl group at the C2 position of methyl α-galactoside. The *N*-acetyl group replaces O2 of galactose, which in the crystal structures forms a 2.8 Å hydrogen bond to Nδ2 of Asn90 (Figure 1, inset). Therefore, an *N*-acetyl group at the C2 position of galactose most likely forces the saccharide to accommodate a different, less favorable position than the

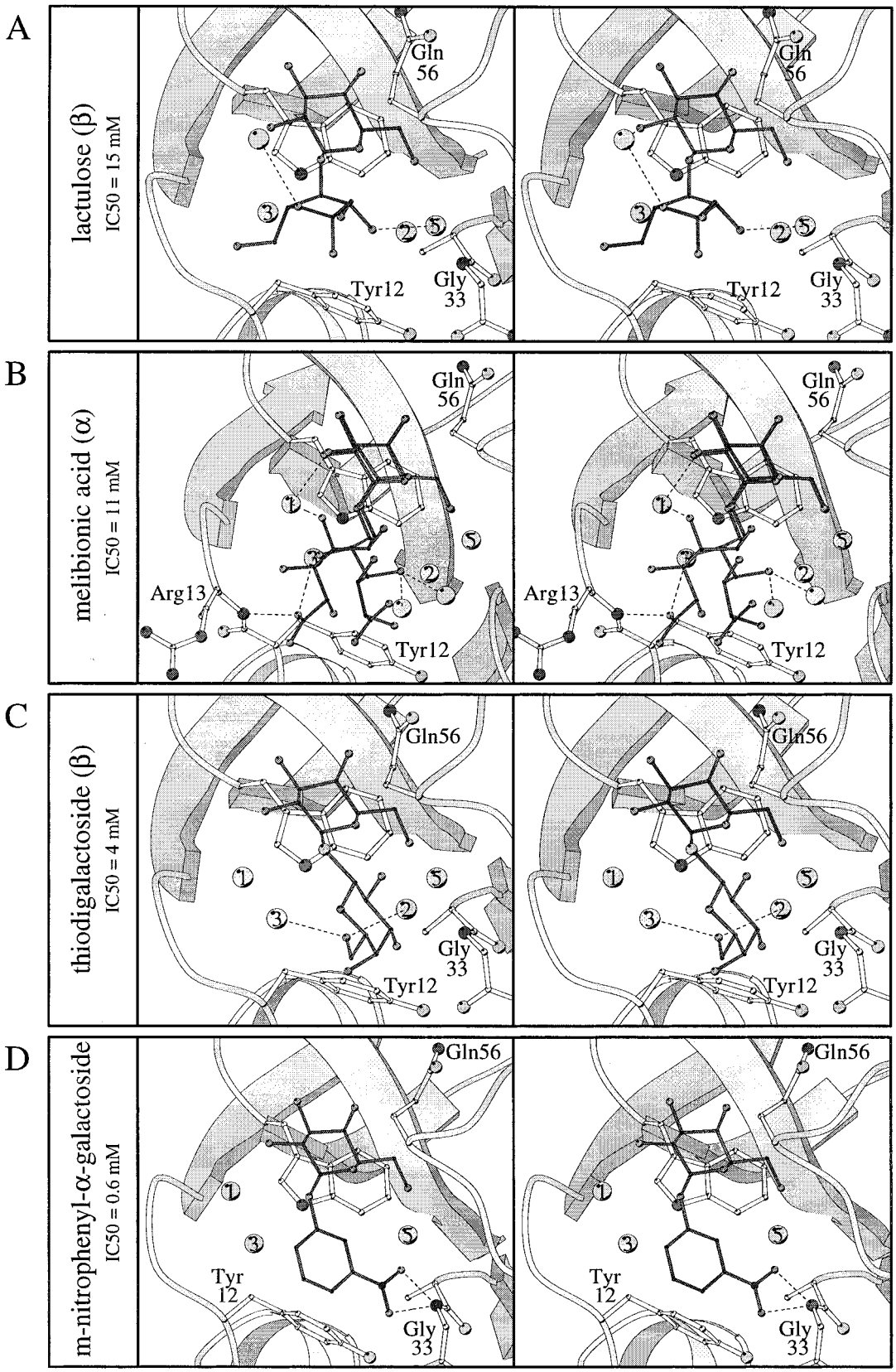

FIGURE 5: Binding mode of lactulose (A), two binding modes of melibionic acid (B), binding mode of thiodigalactoside (C), and binding mode of *m*-nitrophenyl α-galactoside (D) as observed in their crystal structure complexed to pLT B_5 (*39*). The nonlabeled water molecules are noncanonical ones; i.e., these waters are not observed in every structure. This figure was produced with MOLSCRIPT (*40*).

one observed in all crystal structures of complexes with galactose moieties without an *N*-acetyl group in the 2-position (Figures 1 and 5).

Fourth, it is very clear that a nitro group at the ortho position of phenyl galactoside is not favorable (see label 4 in Figure 6). Neither *o*-nitrophenyl α-galactoside nor *o*-

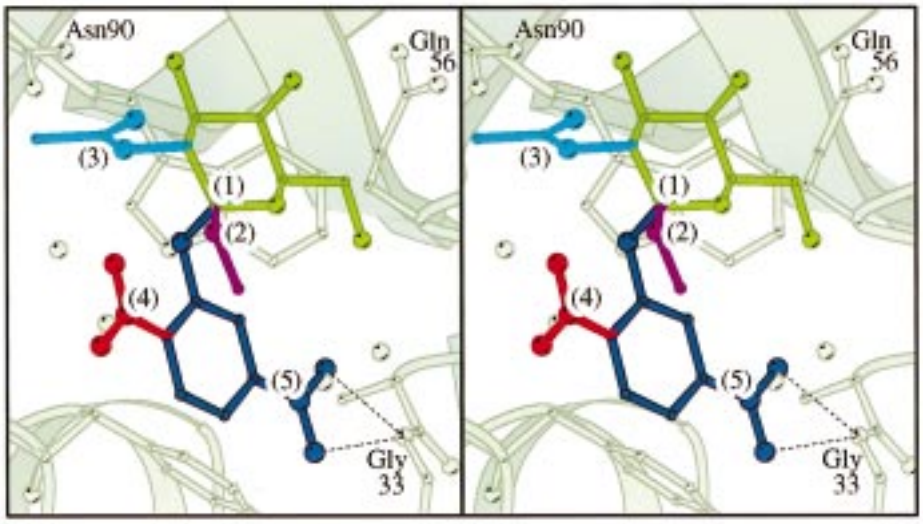

FIGURE 6: Functional groups of galactose derivatives that result in distinctive effects on the binding potential for the binding of the ligands to LT and CT. The crystallographically determined pLT−galactose complex is depicted in green; the functional groups are depicted in bright colors. (1) α-Substituents are favored over β-substituents. (2) For β-substituents, a bridging sulfur is favored over a bridging oxygen. (3) The *N*-acetyl at C2 of galactose is unfavorable. (4) NO_2 at the ortho position is unfavorable. (5) NO_2 at the meta position is favorable. This figure was produced with MOLSCRIPT (*40*).

nitrophenyl β-thiogalactoside exhibits improved binding compared to galactose, while their meta- or para-substituted or unsubstituted analogues do (Table 1). The low affinity of *o*-nitrophenyl at the α-position of galactose may be explained by its conformational preference as observed in the small molecule crystal structure of *o*,*p*-dinitrophenyl 2,3,4,5-tetraacetyl-α-glucoside (*43*). In this structure, the O5−C1−O1−C1′ dihedral angle is 63° and the C1−O1−C1′−C2′ dihedral angle is −162° (the nitro group is attached to C2′). This conformation would position the nitro group too close to Trp88; the distance between the nitro group oxygen and Nε1 of Trp88 is only 1.1 Å, after superpositioning the glucose moiety of dinitrophenyl tetraacetylglucoside onto galactose in its pLT complex structure. For investigation of the low affinity of *o*-nitrophenyl at the β-position of galactose, we analyzed the structure of *o*-nitrophenyl β-glucoside (*44*). However, this structure does not clarify the low affinity of *o*-nitrophenyl β-thiogalactoside, since after placing this structure into the sugar binding site no steric hindrance is observed.

Finally, the most striking result was the high affinity of *m*-nitrophenyl α-galactoside for the toxins; the IC_{50} for *m*-nitrophenyl α-galactoside was about 100 times improved compared to the IC_{50} for galactose in both the LT ELISA and CT assay. It is exactly the meta position of the nitro group that makes this compound such a good inhibitor; neither the ortho nor the para position performs as well (Table 1). As mentioned above, a nitro group at exactly this position is favorable, because a strong hydrogen bond between the ligand and the protein is formed which also leads to the displacement of a water molecule (see label 5 in Figure 6).

Even though *m*-nitrophenyl α-galactoside inhibits 100 times better than galactose, it still inhibits 50 000 times less well than GM1-OS (Table 1). Therefore, the results obtained in this study will be implemented in the process of designing improved LT B_5 receptor antagonists.

ACKNOWLEDGMENT

We thank Brian Jones for thoughtful discussions. Also, we are grateful to Ethan Merritt for access to unpublished crystallographic data, insightful comments, and assistance with the preparation of figures. We are indebted to the Murdock Charitable Trust for a major equipment grant to the Biomolecular Structure Center.

REFERENCES

1. Merritt, E. A., and Hol, W. G. J. (1995) *Curr. Opin. Struct. Biol. 5*, 165−171.
2. Lencer, W. I., Constable, C., Moe, S., Jobling, M. G., Webb, H. M., Ruston, S., Madara, J. L., Hirst, T. R., and Holmes, R. K. (1995) *J. Biol. Chem. 131*, 951−962.
3. Hirst, T. R. (1995) in *Bacterial toxins & virulence factors in disease* (Moss, J., Vaughan, M., Igleweski, B., and Tu, Eds.) pp 123−184, Marcel Dekker, New York.
4. King, C. A., and Van Heyningen, W. E. (1973) *J. Infect. Dis. 127*, 639−647.
5. Finkelstein, R. A., Boesman, M., Neoh, S. H., LaRue, M. K., and Delaney, R. (1974) *J. Immunol. 113*, 145−150.
6. Bastiaens, P. I. H., Majoul, I. V., Verveer, P. J., Soling, H.-D., and Jovin, T. M. (1996) *EMBO J. 15*, 4246−4253.
7. Kassis, S., Hagmann, J., Fishman, P. H., Chang, P. P., and Moss, J. (1982) *J. Biol. Chem. 257*, 12148−12152.
8. Hazes, B., and Read, R. J. (1997) *Biochemistry 36*, 11051−11054.
9. Moss, J., and Vaughn, M. (1977) *J. Biol. Chem. 252*, 2455−2457.
10. Cassel, D., and Pfeuffer, T. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2669−2673.
11. Field, M. (1993) *Arch. Surg. 128*, 273−278.
12. Holmgren, J., and Svennerholm, A. M. (1992) *Gastroenterology Clinics of North America 21*, 283−302.
13. *The World Health Report 1996* (1996) World Health Organization, Geneva, Switzerland.
14. Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., Van Zanten, B. A. M., Witholt, B., and Hol, W. G. J. (1991) *Nature 351*, 371−377.
15. Sixma, T. K., Kalk, K. H., Van Zanten, B. A., Dauter, Z., Kingma, J., Witholt, B., and Hol, W. G. (1993) *J. Mol. Biol. 230*, 890−918.

16. Zhang, R.-G., Scott, D. L., Westbrook, M. L., Nance, S., Spangler, B. D., Shipley, G. G., and Westbrook, E. M. (1995) *J. Mol. Biol. 251*, 563−573.
17. Merritt, E. A., Sixma, T. K., Kalk, K. H., Van Zanten, B. A. M., and Hol, W. G. J. (1994) *Mol. Microbiol. 13*, 745−753.
18. Sixma, T. K., Pronk, S. E., Kalk, K. H., Van Zanten, B. A. M., Berghuis, A. M., and Hol, W. G. J. (1992) *Nature 355*, 561−564.
19. Van den Akker, F., Steensma, E., and Hol, W. G. J. (1996) *Protein Sci. 5*, 1184−1188.
20. Merritt, E. A., Sarfaty, S., Van den Akker, F., L'Hoir, C., Martial, J., and Hol, W. G. J. (1994) *Protein Sci. 3*, 166−175.
21. Murzin, A. G. (1993) *EMBO J. 12*, 861−867.
22. Schengrund, C. L., and Ringler, N. J. (1989) *J. Biol. Chem. 264*, 13233−13237.
23. Schön, A., and Freire, E. (1989) *Biochemistry 28*, 5019−5024.
24. Cuatrecasas, P. (1973) *Biochemistry 12*, 3547−3558.
25. Sugii, S. (1990) *Can. J. Microbiol. 36*, 452−454.
26. Moock, T. E., Henry, D. R., Ozkabak, A. G., and Alamgir, M. (1994) *J. Chem. Inf. Comput. Sci. 34*, 184−189.
27. Feil, I. K., Reddy, R., De Haan, L., Merritt, E. A., Van den Akker, F., Storm, D. R., and Hol, W. G. J. (1996) *Mol. Microbiol. 20*, 823−832.
28. Uesaka, Y., Otsuka, Y., Lin, Z., Kurazono, H., and Takeda, Y. (1994) *Microb. Pathog. 16*, 71−76.
29. Bowen, W. P., and Jerman, J. C. (1995) *Trends Pharmacol. Sci. 16*, 413−417.
30. Hamilton, P. B. (1963) *Anal. Chem. 35*, 2055−2064.
31. Cleland, W. W. (1967) *Adv. Enzymol. 29*, 1−32.
32. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) *Biophys. J. 64*, 166−170.
33. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
34. Svennerholm, A. M., and Holmgren, J. (1978) *Curr. Microbiol. 1*, 19−23.
35. Sugii, S., and Tsuji, T. (1989) *Can. J. Microbiol. 35*, 670−673.
36. Idota, T., Kawakami, H., Murakami, J., and Sugawara, M. (1995) *Biosci., Biotechnol., Biochem. 59*, 417−419.
37. Fishman, P. H., Moss, J., and Osborne, J. C., Jr. (1978) *Biochemistry 17*, 711−716.
38. Mertz, J. A., McCann, J. A., and Picking, W. D. (1996) *Biochem. Biophys. Res. Commun. 226*, 140−144.
39. Merritt, E. A., Sarfaty, S., Feil, I. K., and Hol, W. G. J. (1998) *Structure 5*, 1485−1499.
40. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Fraczkiewicz, R., and Braun, W. (1998) *J. Comput. Chem. 19*, 319−333.
42. Allen, F. H., Baalham, C. A., Lommerse, J. P. M., Raithby, P. R., and Sparr, E. (1997) *Acta Crystallogr. B53*, 1017−1024.
43. Koeners, H. J., De Kok, A. J., Romers, C., and Van Boom, J. H. (1980) *Recl. Trav. Chim. Pays-Bas 99*, 355.
44. Jones, P. G., Kirby, A. J., Glenn, R., and Hadicke, E. (1982) *Z. Kristallogr. 161*, 79.

BI982649A