J|A|C|S
ARTICLES
Published on Web 07/04/2002

Characterization and Crystal Structure of a High-Affinity Pentavalent Receptor-Binding Inhibitor for Cholera Toxin and *E. coli* Heat-Labile Enterotoxin

Ethan A. Merritt,[†] Zhongsheng Zhang,[‡] Jason C. Pickens,[§] Misol Ahn,[†] Wim G. J. Hol,*[,†,||] and Erkang Fan[‡]

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

Received February 19, 2002

Abstract: Multivalent ligand design constitutes an attractive avenue to the inhibition of receptor recognition and other biological events mediated by oligomeric proteins with multiple binding sites. One example is the design of multivalent receptor blockers targeting members of the AB_5 bacterial toxin family. We report here the synthesis and characterization of a pentavalent inhibitor for cholera toxin and *Escherichia coli* heat-labile enterotoxin. This inhibitor is an advance over the symmetric pentacyclen-derived inhibitor described in our earlier work in that it presents five copies of *m*-nitrophenyl-α-D-galactoside (MNPG) rather than five copies of β-D-galactose. The approximately 100-fold higher single-site affinity of MNPG for the toxin receptor binding site relative to galactose is found to yield a proportionate increase in the affinity and IC50 measured for the respective pentavalent constructs. We show by dynamic light scattering that inhibition of receptor binding by the pentavalent inhibitor is due to 1:1 inhibitor:toxin association rather than to inhibitor-mediated aggregation. This 1:1 association is in complete agreement with a 1.46 Å resolution crystal structure of the pentavalent inhibitor:toxin complex, which shows that the favorable single-site binding interactions of MNPG are retained by the five arms of the 5256 Da pentavalent MNPG-based inhibitor and that the initial segment of the linking groups interacts with the surface of the toxin B pentamer.

Introduction

Cholera toxin is a secreted toxin with overall AB_5 architecture. Five identical receptor-binding sites on the cholera toxin B pentamer mediate binding to the epithelial cell surface of the human host through specific interaction with ganglioside GM1. Receptor binding by cholera toxin is both an attractive model system for exploring the design of multivalent antagonists and a compelling target for the design of prophylactic drugs against the initial, acute, result of infection by *Vibrio cholerae*. The same mode of binding to GM1 is exhibited by the closely related *Escherichia coli* heat-labile enterotoxin, the causative agent of traveler's diarrhea. We have undertaken the design of multivalent receptor-binding antagonists against cholera toxin and heat-labile enterotoxin, with a particular focus on exploiting the 5-fold symmetry of the binding sites on the toxin B pentamer.

Previous designs for multivalent ligands have usually employed a generic scaffold, to which are attached many copies of a chemical moiety with relatively poor individual affinity for the target site.[1−15] In such cases the resulting geometry of the ligands being presented is variable, and the affinity gain from multivalent presentation is due primarily to an increase in the effective local concentration of the monovalent ligand. Furthermore, for highly multimeric ligands it is difficult to distinguish ligand-mediated aggregation of the protein from an actual gain in effective affinity. In contrast, if the multivalent presentation of the liganding group is geometrically restrained to match the specific arrangement of binding sites on the target protein, then a substantial additional increase in affinity is to

* Corresponding author. E-mail: hol@gouda.bmsc.washington.edu.
[†] Department of Biochemistry and Biomolecular Structure Center.
[‡] Department of Biological Structure and Biomolecular Structure Center.
[§] Department of Chemistry.
[||] Howard Hughes Medical Institute.

(1) Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755−2794.
(2) Gordon, E. J.; Sanders, W. J.; Kiessling, L. L. *Nature* **1998**, *392*, 30−31.
(3) Reuter, J. D.; Myc, A.; Hayes, M. M.; Gan, Z.; Roy, R.; Qin, D.; Yin, R.; Piehler, L. T.; Esfand, R.; Tomalia, D. A.; Baker, J. R. J. *Bioconjug. Chem.* **1999**, *10*, 271−278.
(4) Thoma, G.; Duthaler, R. O.; Magnani, J. L.; Patton, J. T. *J. Am. Chem. Soc.* **2001**, *123*, 10113−10114.
(5) Kramer, R. H.; Karpen, J. W. *Nature* **1998**, *395*, 710−713.
(6) Profit, A. A.; Lee, T. R.; Lawrence, D. S. *J. Am. Chem. Soc.* **1999**, *121*, 280−283.
(7) Glick, G. D.; Toogood, P. L.; Wiley, D. C.; Skehel, J. J.; Knowles, J. R. *J. Biol. Chem.* **1991**, *266*, 23660−23669.
(8) Kanai, M.; Mortell, K. H.; Kiessling, L. L. *J. Am. Chem. Soc.* **1997**, *119*, 9931−9932.
(9) Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 4179−4190.
(10) Connolly, D. T.; Townsend, R. R.; Kawaguchi, K.; Bell, W. R.; Lee, Y. C. *J. Biol. Chem.* **1982**, *257*, 939−945.
(11) Lee, Y. C. *FASEB J.* **1992**, *6*, 3193−3200.
(12) Kingery-Wood, J. E.; Williams, K. W.; Sigal, G. B.; Whitesides, G. M. *J. Am. Chem. Soc.* **1992**, *114*, 7303−7305.
(13) Spevak, W.; Nagy, J. O.; Charych, D. H.; Schaefer, M. E.; Gilbert, J. H. *J. Am. Chem. Soc.* **1993**, *115*, 1146−1147.
(14) Thompson, J. P.; Schengrund, C.-L. *Glycoconj. J.* **1997**, *14*, 837−845.
(15) Roy, R. *Curr. Opin. Struct. Biol.* **1996**, *6*, 692−702.

10.1021/ja0202560 CCC: $22.00

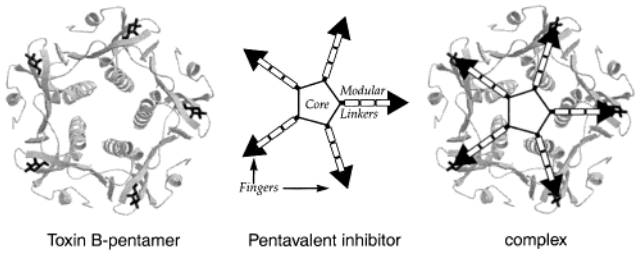

Figure 1. Conceptual design for the pentavalent ligand, showing modular design based on symmetric core, variable numbers of linker units, and monovalent "fingers" that block the toxin's receptor binding site. The toxin B pentamer target structure is shown with galactose (black) bound at each of the five receptor binding sites.

be expected. In such cases the binding of a ligand to one site partially orients and positions additional copies of the ligand to bind at other sites. A portion of the net entropic cost due to the loss of rotational and translational freedom of the ligand in solution is thus paid once for the entire set of multiple binding events, rather than being paid in full for binding at each site. The expected decrease in the entropic cost of binding for a rigid multivalent ligand should lead to a dramatic increase in affinity relative to both free, monovalent, ligands and to randomly presented multivalent ligands such as those based on dendrimers or linear polymers.

Multivalent inhibition of AB_5 toxins has previously been investigated by several groups. Presentation of four to eight copies of the full GM1 receptor oligosaccharide via poly(propylene imine) dendrimers yielded up to a 1000× gain in IC50 for cell binding by cholera toxin.[14,16] Kitov et al.[17] designed a decavalent inhibitor for Shiga toxin, with over a millionfold gain in affinity, based on an asymmetric scaffold. Shiga toxin presents three receptor binding sites per B monomer, of varying affinity, and the intent was to block up to 10 of these sites upon binding a single decavalent inhibitor. In this case, however, the toxin:inhibitor complex was found to consist of two toxin pentamers sandwiching a single inhibitor, with the five highest affinity binding sites occupied on each of the two toxin pentamers.[17] The expected gain in affinity due to multivalent presentation of a ligand to a pentavalent target has been modeled by Gargano et al.[18] and quantified experimentally for linear polymers derived from the Shiga toxin receptor trisaccharide. The observed gain of $>5 \times 10^3$ compared well with the predicted gain of $>10^4$.[18] A gain of $>10^4$ was also found by Fan et al.,[19] who explored the effect of optimizing the effective radius of a symmetric pentavalent inhibitor for cholera toxin and *E. coli* heat-labile enterotoxin using a modular design with multiple flexible linker units. This same design has also been chosen for the work presented here and is used to link five single-site antagonists to a central pentacyclen core. The linkages in the current design are not rigid, leading to a star-shaped molecule with five flexible arms, each arm terminating in a single "finger" which binds to a GM1 binding site on the toxin B pentamer (Figure 1).

In this report, we describe solution studies addressing two important issues in structure-based multivalent ligand design. The first question is whether proportional affinity gains can be realized in a multivalent ligand if it is built from better monovalent ligands. The second question is whether discrete complexes, rather than random aggregates, are formed in solution between the designed multivalent ligand and its target protein. The solution studies point to the absence of multipentamer aggregates in the presence of the pentavalent inhibitor. We also report a high-resolution crystal structure of the complex between our designed pentavalent ligand and the B pentamer of cholera toxin. The crystal structure both supports the solution study results and highlights new opportunities for the design of improved pentavalent ligands by revealing interactions between the first linkage group and the toxin's surface.

Results

Synthesis. The synthesis of the monovalent and pentavalent ligands is shown in Scheme 1. The acid-bearing MNPG analogue **1** was synthesized according to previously reported procedures[20] using galactose pentaacetate and 3-hydroxy-5-nitrobenzoic acid. One mono-Boc protected linker fragment **2** was attached to **1** through amide bond formation.[19] Subsequent acetate removal and enzymatic digestion produced **3** in the pure α-anomeric form.[20] TFA deprotection produced the soluble monovalent ligand **4**. The synthesis of the pentavalent ligand **7** was again straightforward through activation of **4** by dimethyl squarate and subsequent reaction with the core-linker module **6**, following our established synthetic route.[19] Although ideally one would like to obtain a pentavalent ligand with the optimal linker length as revealed in our previous study,[19] the poor solubility associated with longer MNPG-based finger + linker units during synthesis prevented us from obtaining the final product for longer linkers. As a consequence, only ligand **7** with two basic linker units was prepared.

Inhibition. The affinity of the ligands toward toxin B pentamer was measured in terms of ability to inhibit receptor binding (IC50) using a cholera toxin B pentamer (CTB) enzyme-linked adhesion assay.[21] As listed in Table 1, the pentavalent ligand **7** showed about 260-fold affinity enhancement over the monovalent ligand **4**. This enhancement is very close to the equivalent gain when a galactose-based finger is used in the equivalent pentavalent ligand.[19]

Light Scattering. To further study the state of pentavalent ligand−toxin B pentamer complexes in solution, we performed dynamic light scattering (DLS) experiments with varying concentration ratios of ligand **7** and the *E. coli* heat-labile enterotoxin B pentamer (LTB), at concentrations well above the measured IC50 levels. The high concentration ensures that ligand−protein complex formation is maximized. DLS measures the distribution of the hydrodynamic radius (R_h) of solution species, making it a useful tool for determining aggregation behavior in solution. The DLS measurements can distinguish whether the primary mode of interaction in solution is aggregation or the formation of a 1:1 complex. In the latter case one would observe little change in R_h as the ligand is added to the free toxin B pentamer,[19] while in the case of aggregation one

(16) Thompson, J. P.; Schengrund, C. L. *Biochem. Pharmacol.* **1998**, *56*, 591−597.
(17) Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669−672.
(18) Gargano, J. M.; Ngo, T.; Kim, J. Y.; Acheson, D. W. K.; Lees, W. J. *J. Am. Chem. Soc.* **2001**, *123*, 12909−12910.
(19) Fan, E.; Zhang, Z.; Minke, W. E.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J. *J. Am. Chem. Soc.* **2000**, *122*, 2663−2664.
(20) Pickens, J. C.; Merritt, E. A.; Ahn, M.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *Chem. Biol.* **2002**, *9*, 215−224
(21) Minke, W. E.; Roach, C.; Hol, W. G. J.; Verlinde, C. L. M. J. *Biochemistry* **1999**, *38*, 5684−5692.

Scheme 1. Pentavalent Ligand Synthesis

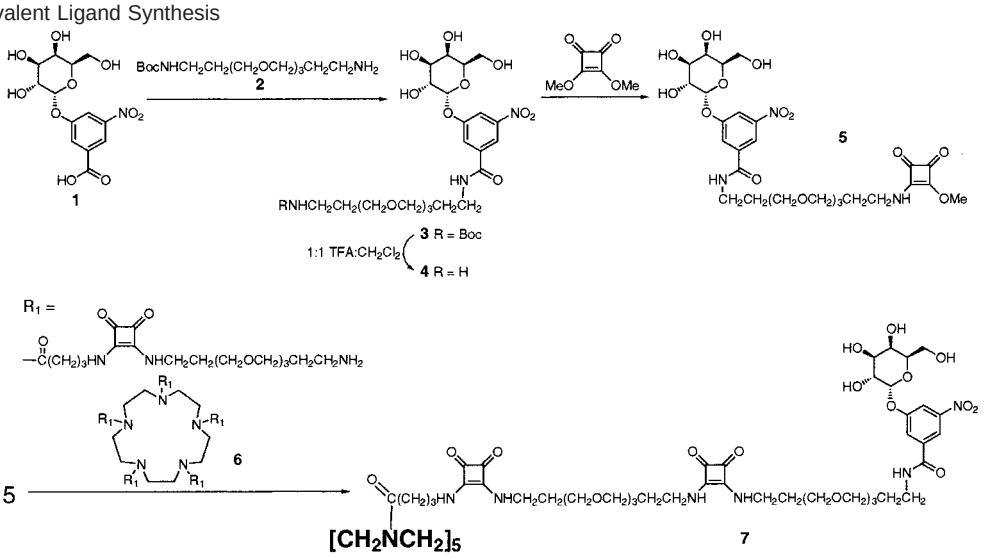

Table 1. Inhibition of Receptor Binding

compound	IC50[a] (esd) (μM)	gain in inhibition compared to single finger
galactose-based finger[19]	5000 (200)	
galactose-based pentavalent ligand[19]	16 (8)	312
MNPG-based finger **4**	195 (96)	
MNPG-based pentavalent ligand **7**	0.9 (0.2)	263

[a] IC50 values given are a weighted average, with the corresponding estimated standard deviation, derived from either three or four replicate ELISA experiments.

Table 2. Dynamic Light Scattering

solution components	R_h (δ) (nm)	MW (kDa)
LTB (11.06 μM)	3.191 (0.251)−3.442 (0.454)	54
7:LTB (11.06:22.12 μM)	3.186 (0.254)−3.338 (0.393)	54
7:LTB (11.06:11.06 μM)	3.135 (0.276)−3.226 (0.227)	51
7:LTB (22.12:11.06 μM)	3.186 (0.254)−3.311 (0.114)	54
CPRG:LTB[19]	12.15 (7.27)	1160

would observe scattering from much larger particles. The DLS experiments were carefully performed by separately prefiltering ligand and protein stock solutions prior to mixing. Then appropriate volumes of ligand and protein solutions were mixed to reach the desired concentration and ratio of ligand to protein in solution. These samples were then used directly for measurement without further treatment or filtration. This eliminated the potential for removal or breakup of aggregates if they were ever formed in substantial quantities. The results from DLS are listed in Table 2. As it clearly shows, the R_h measured for ligand−protein mixtures is very close to that of a single toxin B pentamer, regardless of the ratio of ligand to protein in solution. Therefore, one can safely conclude that, indeed, formation of a 1:1 complex between the pentavalent ligand **7** and the toxin B pentamer is the major mode of association in solution.

Table 2 shows the hydrodynamic radius (R_h) as measured by DLS for varying concentration ratios of the *E. coli* heat-labile enterotoxin B pentamer (LTB) and the pentavalent ligand **7**. The mean R_h along with the associated distribution width δ (polydispersity) is shown as the total range of values obtained from four replicate experiments. The corresponding calculated molecular weight is an approximation based on the assumption of scattering by spherical particles, $MW = (1.68R_h)^{2.3398}$. The final row in Table 2 is a positive control consisting of a mixture of LTB with a monovalent ligand (CPRG) known to cause aggregation (Wendy Minke, personal communication). Unlike the preparation of the **7**:LTB mixtures, the CPRG:LTB preparation was filtered after mixing to remove visible particulate precipitation; thus the measured R_h is an underestimate of the extent of aggregation.

Table 3. Crystallographic Data

data collection	
resolution (highest shell)	50−1.46 Å (1.51−1.46Å)
unique data measured	84 902 (7 083)
completeness	97% (82%)
R_{merge} on intensities	0.081 (0.475)
model	
R	0.149
R_{free}	0.191
protein atoms	4098, $\langle B_{iso}\rangle = 14$ Å^2, $\langle A\rangle = 0.50$
ligand atoms (finger)	120, $\langle B_{iso}\rangle = 21$ Å^2, $\langle A\rangle = 0.59$
ligand atoms (linker)	105, $\langle B_{iso}\rangle = 34$ Å^2
water molecules	567, $\langle B_{iso}\rangle = 32$ Å^2, $\langle A\rangle = 0.51$
data used in refinement	
reflections (working set)	80 506
reflections (R_{free} set)	4 244
cutoffs	27−1.46 Å
stereochemistry	
rms nonideality bond lengths	0.019 Å
rms nonideality 1−3 lengths	0.039 Å
overall coordinate ESU (Cruickshank DPI)	0.08 Å
overall coordinate ESU (maximum likelihood)	0.04 Å

Crystal Structure. Crystals of the toxin:inhibitor **7** complex are isomorphous to those of the toxin:receptor complex, and similarly diffract to near-atomic resolution.[22] Refinement of the present structure against 1.46 Å data yielded an excellent crystallographic model (Table 3), and allowed modeling of a large portion of the inhibitor. The asymmetric unit of the crystal contains one toxin pentamer and one pentavalent inhibitor. The five crystallographically independent receptor binding sites are each to be occupied by one finger of the pentavalent inhibitor. The terminal phenyl-α-D-galactoside moiety adopts in each case

(22) Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. J. *J. Mol. Biol.* **1998**, *282*, 1043−1059.

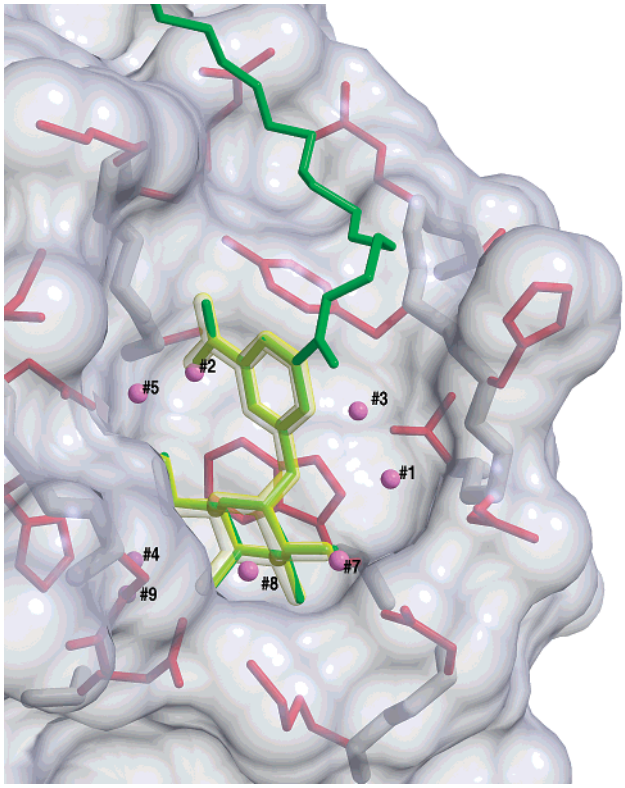

Figure 2. Binding mode of monovalent receptor inhibitor MNPG (yellow), showing α-anomer of galactose and canonical tightly associated water molecules including the water at site 2 that is displaced by the nitrophenyl group of the ligand.[23,24] Superposition (green) of the crystallographically observed conformation of one finger of the pentavalent inhibitor **7**.

the favorable bound conformation that was our design target, analogous to that seen for MNPG. The water molecule tightly associated with other CTB:sugar complexes through hydrogen bonding to the amide hydrogen of residue Gly 33 and to the hydroxyl O6 of galactose is displaced by one oxygen of the nitrophenyl (Figure 2).[23,24] This was by no means guaranteed to be the case, as earlier attempts to vary or extend the core phenyl-α-D-galactoside of MNPG resulted in other, less favorable, binding modes.[25,26] Difference electron density for the first of the two linker units attached to each finger is present at each site and can be modeled up to the first squaric acid moiety (Figure 3). Density for this portion of the inhibitor at the five crystallographically independent sites is consistent with a single favored binding mode. The root-mean-square deviation of refined positions for the inhibitor atoms, based on least-squares superposition of C^{α} atoms in the five corresponding toxin monomers, was 0.25 Å. The second linker group and the pentacyclen core of the pentavalent ligand are not visible in the crystal structure, but the dimensions of a simple model are appropriate to fit both the CTB pentamer and the surrounding crystal lattice (Figure 4).

The crystal structure is indicative of 1:1 binding of the pentavalent ligand to the 5-fold symmetric CTB pentamer. The crystal packing is such that the receptor-binding surface of each toxin pentamer is paired with the receptor-binding surface of a crystallographically related pentamer. The gap between such pentamer pairs in the crystal is large enough to accommodate two pentacyclen cores, as required by the observed fully occupied receptor-binding sites. The central pores of these crystallographically related pentamer pairs are not perfectly aligned, but they are close enough that we cannot rule out a binding mode in which one or more monovalent arms from each toxin:ligand complex swap reciprocally with arms from the paired complex. Such swapping would be stochastic, in that it would vary from one unit cell to the next in the crystal, and could contribute to the lack of clear electron density for the first set of linking groups attached to the pentacyclen core. Nevertheless, the simplest interpretation of the well-defined portions of the toxin:ligand complex in the crystal is that there is 1:1 association of pentavalent inhibitor to toxin pentamer. This is borne out by the absence of higher order assemblies of toxin:ligand multimers seen in solution by dynamic light scattering.

Discussion

The binding interaction between cholera toxin and the branched oligosaccharide of ganglioside GM1 has been studied crystallographically at high resolution,[22,27] revealing that the receptor binding site on the toxin is relatively broad and shallow. This makes it a difficult target for inhibitor design (see Figure 4 in Pickens et al.).[20] The pentasaccharide of the natural receptor, ganglioside GM1, is a fairly rigid molecule whose interaction with the toxin buries 400 Å^2. The terminal galactose residue of GM1 is the most deeply buried portion of the receptor after binding, and all of the more than 25 receptor fragments and galactose-based inhibitors investigated crystallographically are observed to bind with the galactose moiety in the same position and conformation. Thus there is good reason to expect that, as long as the terminal galactose is retained, chemical variation of the linker and core moieties of a multivalent inhibitor will not perturb the fundamental binding mode of the fingers.

The starting point for our inhibitor design was therefore the identification and characterization of small galactose derivatives that bind to the toxin as single-site inhibitors.[21] These are poor competitive inhibitors for the natural receptor, but interpretation of their relative affinities as correlates of the detailed molecular interactions observed crystallographically allowed us to identify key structural features. Among these is an extensive hydrogen bonding network involving side chains from the toxin, particularly from residues in the 50−60 loop, and a well-defined set of tightly associated water molecules at the toxin surface.[23] In particular we found improved affinity from α-anomeric sugar substituents and from displacement of one or more of the tightly associated water molecules (Figure 2). Both of these features are exhibited by the compound *m*-nitrophenyl-α-D-galactoside (MNPG),[24,26] and the present crystal structure of the **7**:CTB complex shows that these favorable binding interactions are retained by a pentavalent receptor antagonist presenting five MNPG moieties linked to a central 5-fold symmetric core.

Structure-guided inhibitor design based on derivatizing the anomeric oxygen of the galactose has yielded in the best instance to date an affinity of $K_d \cong 10\ \mu$M.[20] This is $>10^3$ better than

(23) Merritt, E. A.; Sixma, T. K.; Kalk, K. H.; van Zanten, B. A. M.; Hol, W. G. J. *Mol. Microbiol.* **1994**, *13*, 745−753.
(24) Merritt, E. A.; Sarfaty, S.; Feil, I. K.; Hol, W. G. J. *Structure* **1997**, *5*, 1485−1499.
(25) Minke, W. E.; Pickens, J.; Merritt, E. A.; Fan, E.; Verlinde, C. L. M. J.; Hol, W. G. J. *Acta Crystallogr., Sect. D* **2000**, *D56*, 795−804.
(26) Fan, E.; Merritt, E. A.; Zhang, Z.; Pickens, J.; Roach, C.; Ahn, M.; Hol, W. G. J. *Acta Crystallogr., Sect. D* **2001**, *D57*, 201−212.
(27) Merritt, E. A.; Sarfaty, S.; van den Akker, F.; L'hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.* **1994**, *3*, 166−175.

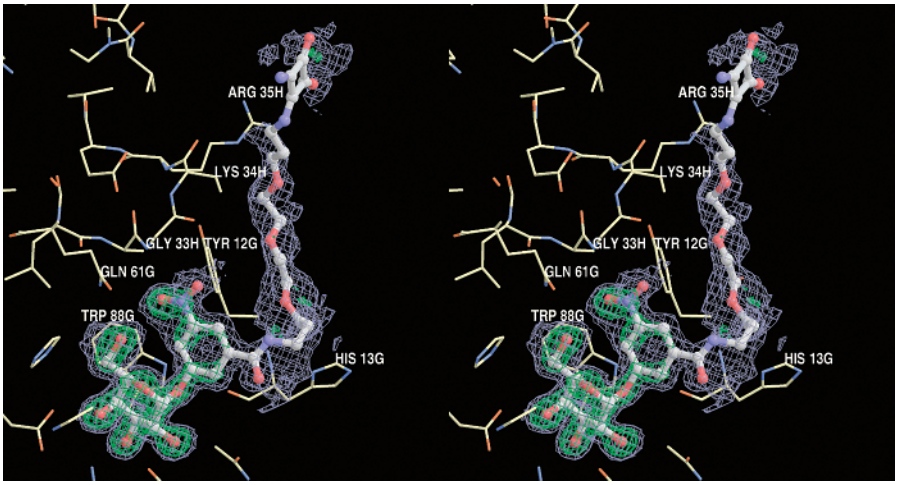

Figure 3. Stereopair of the finger and first linker unit of **7** fit into electron density at one of the five identical binding sites in the crystal structure of the CTB:**7** complex. Experimental electron density is contoured at 1.5σ (gray) and 3.0σ (green) in a σ_A-weighted (mFo − Fc) difference map calculated before addition of any ligand atoms to the crystallographic model.

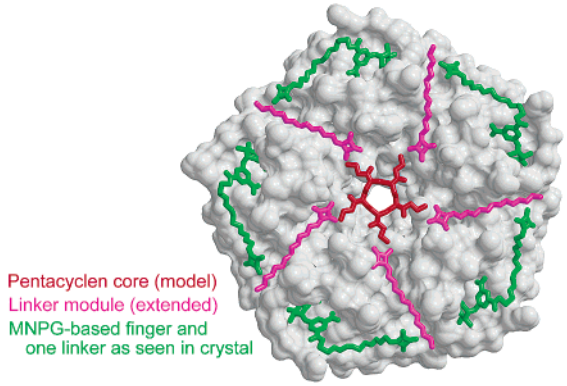

Figure 4. Molecular surface of the CTB pentamer shown with a composite model of the modular components of the pentavalent inhibitor **7**. The finger + linker substructure observed in the crystal structure is shown in green; an energy-minimized model for the pentacyclen core is shown in red; a second linker unit is shown in purple in a fully extended conformation to indicate the maximum distance spanned by a single linker unit. Note that the location of the squaric acid moiety belonging to the first linker unit, although shown here in green, is poorly indicated by experimental electron density as compared to the rest of the finger + linker substructure (Figure 3).

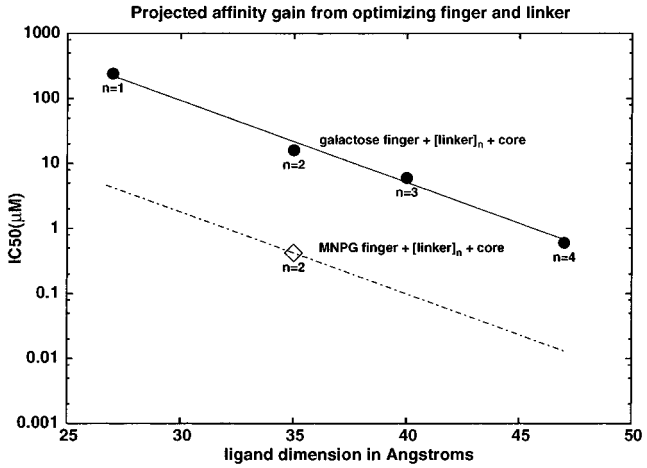

Figure 5. Extrapolation of the affinity gain that may be achieved by further optimization of the pentavalent MNPG-based inhibitor. The improvement in IC50 gained by increasing the number of linker modules in a galactose-based inhibitor is shown in the upper line of the figure.[19] The best affinity was found for a construct containing four linker modules, corresponding to an expected mean ligand dimension matching that of the target toxin B pentamer. The MNPG-based pentavalent inhibitor reported here contains two linker modules, corresponding to a mean ligand dimension in solution that is smaller than the predicted optimum. If a similar dependence of IC50 on total linker length holds for the MNPG-based inhibitor, then further improvement should be possible, as indicated by the second line shown running parallel to the line fit to previously observed data.[19]

the affinity of galactose itself, but still many orders of magnitude worse than that of the natural receptor. Thus there is at present only a remote prospect of finding a simple galactose derivative with sufficient single-site affinity to be medically relevant. These compounds are nonetheless of interest as candidates for multivalent presentation. The present structure bears this out, showing that our five-armed core + linker + finger construct retains the desired favorable binding mode of the single-site inhibitor MNPG (Figure 2). Furthermore, we have now shown that improved affinity in a single-site ligand confers a parallel improvement in the affinity of a pentavalent construct presenting five copies of this ligand. Pentavalent presentation of β-galactose in a symmetric construct with two linker units resulted in a 310-fold increase in affinity over a single β-galactose based finger−linker unit.[19] Pentavalent presentation of MNPG in a symmetric construct with two linker units results in a 260-fold increase compared to a single MNPG−finger−linker unit (**4**) (Table 1).

In an earlier paper we showed that the affinity of a pentavalent receptor antagonist for the *E. coli* heat-labile enterotoxin B pentamer (LTB) was strongly dependent on the effective length in solution of the five arms.[19] In that pentavalent design, based on a much weaker single-site ligand, the optimal number of linker units was shown to be four, corresponding to an effective solution radius of the pentavalent inhibitor that matches the distance between nonadjacent receptor binding sites on the toxin pentamer (Figure 5). By comparison the pentavalent construct characterized here has only two linker units per arm, and is thus too short for optimal complementarity to the target set of binding sites on the protein. The present two-linker construct with fingers based on MNPG already exhibits 100-fold greater inhibition of toxin:receptor binding than does the analogous two-linker pentavalent ligand based on galactose fingers (Table 1). We therefore predict that optimization of the arm length, using modified linker units with greater solubility, would produce a further 100-fold gain in affinity, parallel to the gain previously

seen for the galactose-based construct, bringing the net IC50 into the subnanomolar range (Figure 5).

Characterization of the specific mode of interaction between a multivalent ligand and its target protein is of great theoretical interest, yet can pose a great challenge. For example, when a multivalent ligand exhibits increased activity compared to a monovalent ligand, it can be difficult to prove whether this gain is due to an increase in intrinsic affinity or due to other effects such as aggregation, precipitation, or the combination of several factors. This is especially true when the gain is characterized through inhibition assays rather than by direct measurement of affinities at equilibrium. Therefore, it is not surprising that there is still active debate about the true gain in intrinsic affinity from multivalency.[28,29] Any heterogeneity of the multivalent ligand itself adds to the difficulty of interpreting assay results. Thus conventional multivalent approaches based on polymers,[4,8,9,30] dendrimers,[3,14−16] or other nonhomogeneous backbones[1,12,13] suffer from a lack of precise control over the number and the geometry of monovalent ligands attached to each backbone molecule. In contrast, a modular multivalent approach such as we report here enables the synthesis and isolation of a single species multivalent ligand. We are therefore able to use a variety of biophysical tools to study the solution behavior of the ligand−protein complex. We chose to use DLS to study the interaction between **7** and toxin B pentamer in solution because DLS is very sensitive to aggregate formation. The DLS experiments provide crucial evidence on the nature of complexes formed in solution between the pentavalent ligand **7** and toxin B pentamer. As listed in Table 2, there is no sign of significant formation of large aggregates or precipitation in this pentavalent ligand−toxin B pentamer system under a variety of conditions. Over a range of ligand:toxin concentration ratios from 1:2 to 2:1, the observed effective hydrodynamic radius of the complex does not significantly vary (Table 2). This is strong evidence that 1:1 complex formation between ligand **7** and the toxin B pentamer is the major event in solution at micromolar concentration or lower. Therefore, one can safely conclude that the gain in inhibition potency measured in our competitive inhibition assay is due not to aggregation, but rather to a gain in the intrinsic affinity of the pentavalent ligand for the toxin B pentamer.

An ideal receptor antagonist would exhibit perfect complementarity to the full set of binding sites on the target; in the case of cholera toxin and heat-labile enterotoxin, it would contain five rigidly linked GM1 antagonists, each of which would bind to one site on the toxin and block toxin:receptor interactions at that site. Our current work falls short of this ideal, as the core and linker units that join the binding site ligands are flexible. Nevertheless, a substantial portion of one linker unit of each arm is well-ordered in the current **7**:CTB complex (Figures 3 and 4), making van der Waals contact with the side chains of residues Glu 11, Tyr 12, His 13, Lys 34, and Arg 35. This may partially explain the increase in IC50 observed for the MNPG-based finger **4** relative to MNPG itself (Table 1), and suggests that a rigidified variant of the linker can be designed that would optimize these interactions with the protein. Ideally this would both yield enthalpic gains from the specific interactions and yield entropic gains due to a smaller loss of entropy arising from the decreased number of rotatable bonds in the unbound ligand. It is also possible that a modified core moiety, replacing **6** in the current scheme, can be designed to interact favorably with the toxin surface in the region of the central pore (Figure 4). The molecular weight and physical dimensions of such a fully optimized pentavalent ligand will be quite large, but the site of biological action, epithelial cell surfaces in the intestinal lumen, means that ligand size is not by itself a problem. A receptor-binding antagonist for these toxins should be orally delivered but need not, and ideally would not, be transported into the bloodstream. This is in distinction to the more typical requirement that a drug enter the bloodstream and be delivered to intracellular targets elsewhere in the body. Thus multivalent approaches to toxin inhibition such as described here may one day lead to prophylaxis of cholera and *E. coli* induced traveler's diarrhea.

Experimental Section

Synthesis. (A) General. Common solvents (reagent grade or HPLC grade) were used as purchased from commercial sources without further purification. HPLC purification was performed on an HP 1100 quaternary pump system with a variable wavelength detector. The C18 preparative column was purchased from Vydac (21 × 250 mm, 10−15 μm). ^{1}H NMR spectra were obtained at 300 MHz on a Bruker AC-300 instrument, while mass spectra were obtained from a Bruker Esquire 3000 electrospray ion trap mass spectrometer.

(B) {3-Nitro-5-[3-(2-{2-[3-(*t*-Boc-amino)propoxy]ethoxy}ethoxy)-propylaminocarbonyl]phenyl}-α-D-galactopyranoside (3). A general synthetic procedure is as follows: To a solution containing crude **1** (300 mg, 0.58 mmol)[20] in a total of 15 mL of EtOAc, 36 mg (0.19 mmol) of cyanuric chloride was added. After the cyanuric chloride was completely dissolved, 60 μL (0.58 mmol) of *N*-methylmorpholine was added while stirring vigorously. After 2 h, the monoprotected diamine **2** (189 mg, 0.58 mmol) was added dropwise. The mixture was allowed to stir for 12 h, and then was filtered through a 0.4 μm nylon syringe filter to remove insoluble byproducts. The solvent was then removed by rotary evaporation and the crude sugar-protected product **3** brought up in 10 mL of MeOH. Approximately 10 mg of sodium was added and the solution was allowed to stir for 2 h, and then passed through 10 mL of Dowex cation-exchange resin (8% cross-linking, 200 mesh) in the acidic form. HPLC purification (solvent A, 0.1% TFA; solvent B, CH_3CN gradient 0 to 15% B over 5 min, then to 60% B over 25 min) gave 68 mg of a 1:1 anomeric mixture of **3**. Yield: 18%. The anomeric mixture was brought up in 10 mL of 0.1 M phosphate buffer (pH 7, 1 mM $MgCl_2$) containing 500 units of β-galactosidase. The enzymatic digest was continued overnight. After removing the enzyme by filtration, the solution was subjected to HPLC repurification using the same conditions as above giving 10.5 mg of **3**. Recovery: 30%. ESI-MS: 648.4 $[M + H]^+$. ^{1}H NMR: d 8.33 (s, 1H), 8.16 (s, 1H), 7.97 (s, 1H), 5.70 (s, 1H), 3.98 (m, 3H), 3.87 (t, 1H), 3.70−3.45 (m, 16H), 3.10 (t, 2H), 1.90 (m, 2H), 1.69 (m, 2H), 1.42 (s, 9H).

(C) {3-Nitro-5-[3-(2-{2-[3-aminopropoxy]ethoxy}ethoxy)propylaminocarbonyl]phenyl}-α-D-galactopyranoside (4). Compound **3** (21 mg, 32.3 μmol) was treated with 2.0 mL of 1:1 TFA:CH_2Cl_2 for 5 min at room temperature. After rotary evaporation of solvent, the residue was dissolved in 3 mL of water and neutralized with aqueous NaOH to pH 2. HPLC purification (solvent A, 0.1% aqueous TFA; B, CH_3CN using same gradient as for compound **3**) produced 15 mg of **4**. Yield: 84%. ESI-MS: 548.4 $[M + H]^+$.

(D) Pentavalent Ligand (7). Compound **3** (10 mg, 15.4 μmol) was treated with 1.0 mL of 1:1 TFA:CH_2Cl_2 for 5 min at room temperature.

(28) Dimick, S. M.; Powell, S. C.; McMahon, S. A.; Moothoo, D. N.; Naismaith, J. H.; Toone, E. J. *J. Am. Chem. Soc.* **1999**, *121*, 10286−10296.
(29) Lundquist, J. J.; Debenham, S. D.; Toone, E. J. *J. Org. Chem.* **2000**, *65*, 8245−8250.
(30) Mourez, M.; Kane, R. S.; Mogridge, J.; Metallo, S.; Deschatelets, P.; Sellman, B. R.; Whitesides, G. M.; Collier, J. R. *Nat. Biotechnol.* **2001**, *19*, 958−961.

After rotary evaporation of solvent, the residue was dissolved in ~1 mL of water and neutralized with aqueous NaOH to pH ~7. Then a solution of dimethyl squarate (5.0 mg, 35 μmol) in 1.0 mL of MeOH was added. This mixture was adjusted to pH 8 by 0.1 M $NaHCO_3$, and was stirred at room temperature overnight. After neutralization with 0.1% aqueous TFA to pH 7, the mixture was subjected to HPLC purification (solvent A, H_2O; B, CH3CN gradient 0 to 15% B over 15 min, then to 55% B over additional 30 min) to produce 7.2 μmol of the activated compound **5**. Yield: 47%. ESI-MS: 658.5 $[M + H]^+$.

This 7.2 μmol of compound **5** was mixed with 0.66 μmol of compound **6**[19] in a 4.0 mL solution of 1:1 H_2O:MeOH. The pH of the solution was adjusted to pH 9.0 by 0.1 M $NaHCO_3$ (~0.3 mL). After stirring at room temperature for 36 h, the mixture was acidified to pH 2.0 by TFA and subjected to HPLC purification (solvent A, 0.1% aqueous TFA; B, CH_3CN gradient 20 to 50% B over 30 min). A total of 0.59 μmol of final pentavalent ligand **7** in 3.0 mL of aqueous solution was obtained. The yield was determined by photometric measurement of the absorbance at 290 nm ($\epsilon_{290} = 1.568 \times 10^5\ M^{-1}\ cm^{-1}$). The final pure compound was kept in aqueous solution at pH 7 and used directly for assays and cocrystallization experiments. Yield: 89%. ESI-MS: 1052.8 $[M + 5H]^{5+}$, 1315.7 $[M + 4H]^{4+}$, and 1754.0 $[M + 3H]^{3+}$.

IC50 Measurements. The CTB:GD1b enzyme-linked adhesion assay was performed as previously reported.[21] Test samples consisted of 6 ng/mL CTB horseradish peroxidase conjugate (Sigma-Aldrich, Milwaukee, WI) preincubated with desired ligands for 2 h at room temperature. Independent experiments for each inhibitor were carried out in triplicates and validated against a concentration gradient of 0, 1.5, 3, 6, and 12 ng/mL toxin peroxidase conjugate. IC50 values were calculated from triple data sets of at least seven different concentrations of competitive ligands by nonlinear regression,[31] and reported as a weighted average of three or four independent determinations.

Dynamic Light Scattering (DLS). Protein or ligand in phosphate buffered saline (150 mM NaCl/10 mM phosphate, pH 7.2) was individually filtered through inorganic membrane filters (Whatman, Anodisc13, 0.02 μm) into separate vials. Then various portions of protein and ligand were mixed and, without further filtration, transferred into a sample cell for measurement. DLS measurement was done on a DynaPro99 instrument (Protein Solutions Inc.) illuminated by a 25 mW, 832.8-nm-wavelength, solid-state laser at 25 °C. Data analysis was performed using the dynamics Version 5.25.44 software provided with the instrument. Inverse Laplace transform ("regularization fit") analysis was used to find the mean and standard deviation (polydispersity) of the hydrodynamic radius (R_h) distribution for the molecule/complex species in solution.

E. coli heat-labile enterotoxin B pentamer (LTB) was used for DLS experiments in the presence or absence of the pentavalent ligand **7**. The DLS measurements were carried out with the following solutions: (1) LTB (11.06 μM); (2) LTB:**7** (11.06:22.12 μM); (3) LTB:**7** (11.06:11.06 μM); (4) LTB:**7** (22.12:11.06 μM). Independent experiments were performed three to four times for each set of solution.

Protein Production. Cholera toxin B pentamer and *E. coli* heat-labile enterotoxin B pentamer were separately expressed in *E. coli*, using constructs kindly provided by Claudia Roach, and purified by galactose affinity chromatography essentially as described in Minke et al.[25]

Crystallization and Structure Determination. Crystals of CTB complexed with **7** grew from sitting drops consisting of 1 μL of protein at 5.0 mg/mL in 100 mM Tris HCl at pH 7.5, 1 μL of 0.17 mM **7** in the same buffer, and 1 μL well buffer containing 50 mM NaCl, 100 mM Tris HCl at pH 7.5, and 42% PEG 300. Crystals were isomorphous to the previously determined complex of CTB with the GM1 pentasaccharide.[27]

X-ray intensities from a single flash-frozen crystal were measured using a wavelength of 0.9793 Å at the APS Structural Biology Center beamline 19ID. The data were integrated and scaled using programs HKL2000 and TRUNCATE.[32,33] Isotropic refinement of the protein and well-ordered water molecules not in the region of the receptor-binding site was carried out in REFMAC version 4[34] using data to 1.5 Å and a riding hydrogen model. This yielded residuals $R = 0.191$ and $R_{free} = 0.221$. Clear electron density was present in σ_A weighted difference maps (mFo − Fc) for the "finger" portion of the inhibitor at all five binding sites, and fragmentary density for the start of the attached linker moieties (Figure 3). Refinement was continued with the addition of individual atomic displacement parameters U^{ij} to the model and inclusion of data to 1.46 Å, bringing the residuals to $R = 0.165$ and $R_{free} = 0.210$. At this point the inhibitor fingers (from the sugar to the amide nitrogen) were built into difference density and added to the refinement. This yielded residuals $R = 0.148$ and $R_{free} = 0.191$. Portions of the dilinker groups attached to the amide were modeled into the remaining difference density at each of the binding sites (Figure 3) and refined with isotropic thermal parameters, yielding the final model as described in Table 3.

Model fitting, placement, and real-space refinement of ligand and water molecules was carried out using XFIT.[35] The choice of restraint weights for U^{ij} parameters was guided by analysis of the overall distribution of anisotropy using PARVATI,[36] and the anisotropic treatment of water molecules was additionally restrained toward isotropy using a local modification to REFMAC.[37] The resulting mean value of atomic anisotropy, $\langle A \rangle$, is given in Table 3 for protein, ligand, and solvent atoms. No noncrystallographic symmetry restraints were used during refinement. Figures were generated using RASTER3D,[38] MSMS,[39] and XFIT.[35] Structure factors and the refined model have been deposited with the Protein Databank (accession code 1LLR).

Acknowledgment. This work was supported by the NIH (AI34501, AI44954). We thank Claudia Roach for generating the expression system for CTB, and Wendy Minke and Christophe Verlinde for stimulating discussions. Use of the Argonne National Laboratory Structural Biology Center beamline 19ID at the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Biological and Environmental Research, under Contract No. W-31-109-ENG-38.

JA0202560

(31) Bowen, W. P.; Jerman, J. C. *Trends Pharmacol. Sci.* **1995**, *16*, 413−417.
(32) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.
(33) Collaborative Computational Project Number 4. *Acta Crystallogr., Sect. D* **1994**, *D50*, 760−763.
(34) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D* **1997**, *D53*, 240−255.
(35) McRee, D. *Practical Protein Crystallography*; Academic Press: San Diego, 1993.
(36) Merritt, E. A. *Acta Crystallogr., Sect. D* **1999**, *D55*, 1109−1117.
(37) Murshudov, G.; Lebedev, A.; Vagin, A. A.; Wilson, K. S.; Dodson, E. J. *Acta Crystallogr., Sect. D* **1999**, *D55*, 247−255.
(38) Merritt, E. A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505−524.
(39) Sanner, M. F.; Spehner, J.-C.; Olson, A. J. *Biopolymers* **1996**, *38*, 305−320.