Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 907–920
Bioorganic & Medicinal Chemistry

3,5-Substituted phenyl galactosides as leads in designing effective cholera toxin antagonists: synthesis and crystallographic studies

Daniel D. Mitchell,[a,†] Jason C. Pickens,[a,b,†] Konstantin Korotkov,[a] Erkang Fan[a,*] and Wim G. J. Hol[a,c,*]

[a]*Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA*
[b]*Department of Chemistry, University of Washington, Seattle, Washington 98195, USA*
[c]*Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA*

Received 25 July 2003; accepted 18 December 2003

Abstract—With the aim of developing high-affinity mono and multivalent antagonists of cholera toxin (CT) and *Escherichia coli* heat-labile enterotoxin (LT) we are using the galactose portion of the natural receptor ganglioside GM1 as an anchoring fragment in structure-based inhibitor design efforts. In order to establish a better structure–activity relationship for guiding these studies, we designed and prepared a small focused library of twenty 3,5-substituted phenylgalactosides based on two previous leads. The compounds were tested for their ability to block CTB_5 binding to immobilized ganglioside receptor and compared to the two previous leads. The crystal structures of the most promising compounds bound to either CTB_5 or LTB_5 were then determined in order to understand the basis for affinity differences. The most potent new compound yielded a six-fold improvement over our benchmark lead *m*-nitrophenyl-α-D-galactopyranoside (MNPG), and a two-fold improvement in IC_{50} over a newer MNPG derivative. These results support the notion that the *m*-nitrophenyl moiety of MNPG and its derivatives is an important element to retain in future optimization efforts. Additionally, a consensus binding-pocket for the alkylmorpholine or piperazine moiety present in all of the designed antagonists was established as an important area of the GM1 binding site to target in future work.

1. Introduction and background

Cholera toxin (CT) and heat-labile enterotoxin (LT) are two closely related AB_5 heterohexameric virulence factors from *Vibrio cholerae* and enterotoxigenic *Escherichia coli* (ETEC), respectively. Cholera toxin is the causative agent of cholera, which is currently a disease endemic in many developing countries that kills several thousand individuals and infants each year.[1] Heat-labile enterotoxin is also responsible for a large number of deaths in developing countries annually. For example, ETEC-related diarrhea was estimated to be responsible for approximately 500,000 deaths in 1998.[2] CT and LT are structurally and functionally very similar toxins, sharing 80% sequence identity for both the A and the B chains.[3] Both heterohexamers are assembled in the periplasm of the respective bacterium and contain one catalytic A subunit, and 5 identical B subunits. The A subunit is responsible for the ADP-ribosylation of the regulatory membrane-linked $G_{s\alpha}$ protein in the cytosol of host epithelial cells lining the lumen of the gut.[4] This ADP-ribosylation constitutively activates $G_{s\alpha}$, causing continuous stimulation of adenylate cyclase. As a consequence, the cellular cAMP levels increase and activate several membrane ion channels. The resulting imbalance of ions causes an efflux of cellular water to the gut, and produces host dehydration and diarrhea.

In order for CT and LT to exert these toxic effects, recognition and binding to at least one copy of the receptor ganglioside GM1 [Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Gal-β1-4Glc-Ceramide], displayed on the exterior of intestinal epithelial cells, is required prior to internalization of the toxins.[5] Blocking this critical first interaction is an attractive avenue for developing therapeutics for the treatment and prevention of diarrhea in humans stemming from infection with *V. cholerae* and enterotoxigenic *E. coli*. Because the disease usually

* Corresponding authors. Tel.: +1-206-685-7044; fax: +1-206-685-7002; e-mail: erkang@u.washington.edu; wghol@u.washington.edu
† These authors contributed equally to this work.

0968-0896/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.12.019

emerges in areas with a contaminated water supply, the toxins usually do the most damage during large-scale outbreaks, where thousands of deaths can occur before aid arrives. Since there are no effective small-molecule prophylactic or therapeutic drugs available for CT and LT- related diarrhea presently that block the action of the toxin, the development of such drugs would be of great benefit world-wide.

We are engaged in an effort to design both monovalent and multivalent CT and LT receptor-binding antagonists using information obtained from the crystallographic analysis of a number of CT and LT small molecule co-crystal structures. From the structures of $LTAB_5$ with lactose and CTB_5 in complex with the oligosaccharide head group of GM1 (GM1-OS) it was discovered that the terminal galactose could potentially serve as an 'anchor' point in future antagonist design.[6,7] This anchor-based approach is an especially important concept in the case of designing and optimizing small-molecule antagonists of CT and LT because its shallow receptor-binding site lacks well-defined pockets that can be exploited using traditional structure-based drug design. Galactose itself binds to the toxins with a relatively low affinity, having an approximate IC_{50} of 100 mM in our current assay system. A subsequent screen of 35 commercially available O1 substituted galactose derivatives revealed *m*-nitrophenyl-α-D-galactopyranoside (MNPG) to be the most effective antagonist of the series binding with an approximate 100-fold increase in affinity over galactose alone.[8] Crystal structures of CTB_5-MNPG and LTB_5-MNPG revealed that the MNPG molecule was bound in nearly the same binding mode in over 15 crystallographically independent binding sites examined in the LT and CT pentamers.[9,14] The binding mode featured an oxygen from the nitro group displacing canonical water #2 from the binding site in order to make a hydrogen bond with the backbone amide-hydrogen of Gly 33 from an adjacent subunit. The entropically favorable loss of this water molecule has been hypothesized to account in large part for the 100-fold affinity gain by MNPG over β-D-galactose.[9]

In another study, we described a series of designed antagonists incorporating the MNPG core focusing on two lead compounds (**1** and **2** of Fig. 1) that displayed multiple binding modes in co-crystal structures with CTB_5 and LTB_5.[10] These compounds feature a morpholine ring linked to the MNPG core by a short aliphatic linker and resulted in improved aqueous solubility and higher affinity. Based on the fact that these compounds were found to have a 10 to 14-fold affinity gain over MNPG toward LTB_5, a major conclusion of the study was that such an anchor-based approach to ligand optimization was a valid way of improving the potency of our monovalent antagonists. However, the modest overall affinity of these compounds compared to GM1-OS (low micromolar vs. low nanomolar K_d) still leaves much scope for improvement. Additionally, our ongoing efforts at incorporating an optimized set of monovalent antagonists into a variety of multivalent scaffolds require that they contain a primary amine for conjugation.[11–13] Thus, the development of potent, 'linkable' monovalent antagonists is an important facet of our broader efforts aimed at the creation of highly avid multivalent ligands.

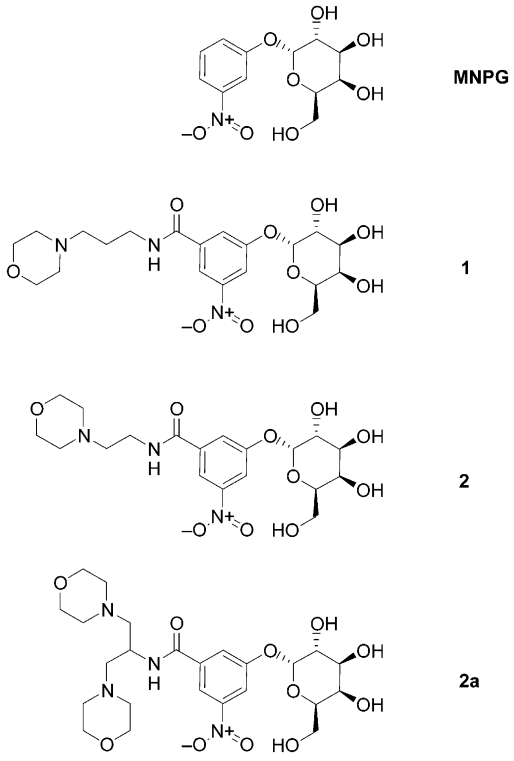


Figure 1. First generation leads derived from MNPG.

We report here the design, synthesis, and characterization of an improved series of linkable, monovalent antagonists of CT and LT receptor binding. Competitive binding assays and X-ray crystallographic studies of the most promising compounds reveal the basis for observed trends in binding affinity. The resulting improved structure–activity relationships (SAR) will aid in the design of future optimized multivalent ligands targeting CT and LT.

2. Results and discussion

2.1. Initial SAR and focused library design

In an effort to further improve affinity while retaining good aqueous solubility and linkability, we have relied heavily on information gained from analysis of the large set of binding modes observed in co-crystal structures of compounds **1** and **2**, with CTB_5 and LTB_5 (Fig. 2).[10] These efforts have proved challenging as the multiplicity of binding modes makes it very difficult to establish a clear-cut structure activity relationship. Specifically, two issues presented themselves:

First, in earlier studies, special attention was given to the observation that the nitrophenyl moiety of MNPG is able to displace a crystallographically observed water (consensus water #2) to make a hydrogen bond with the main chain amide group of Gly 33 of the toxin pentamer in the process.[9] In some binding modes adopted by

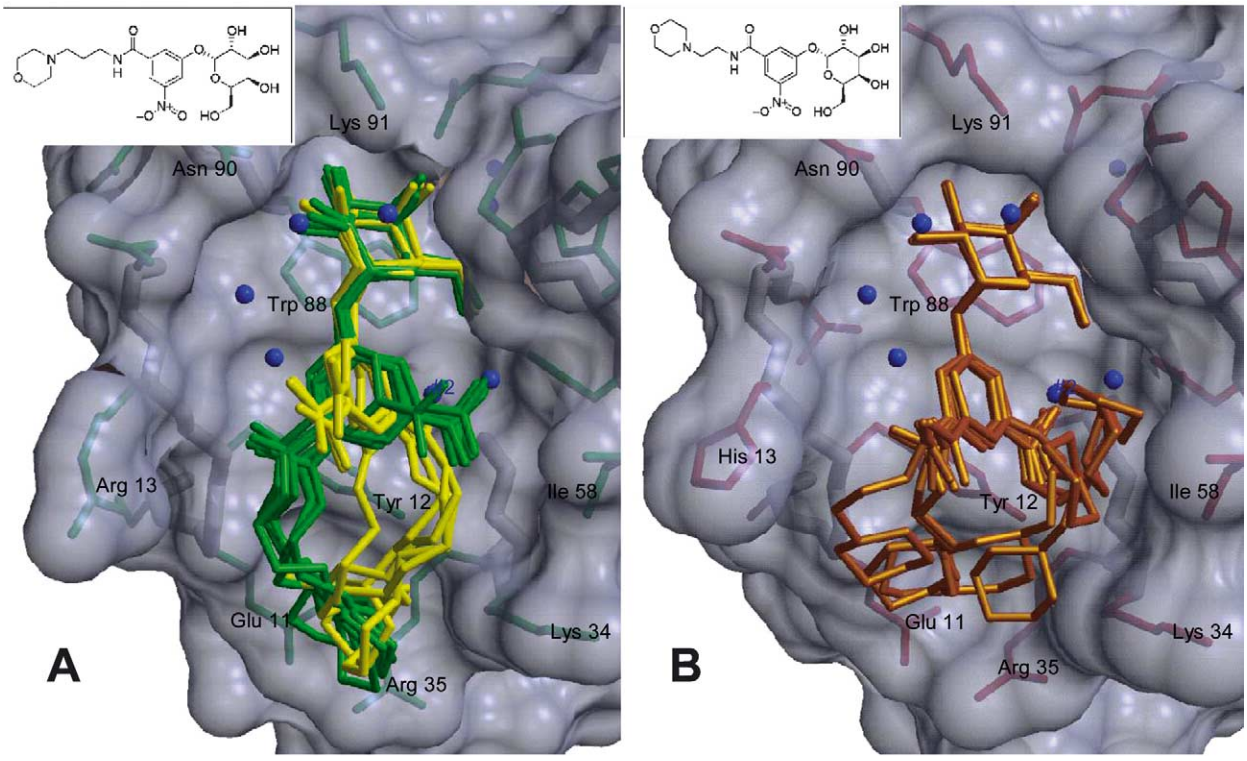
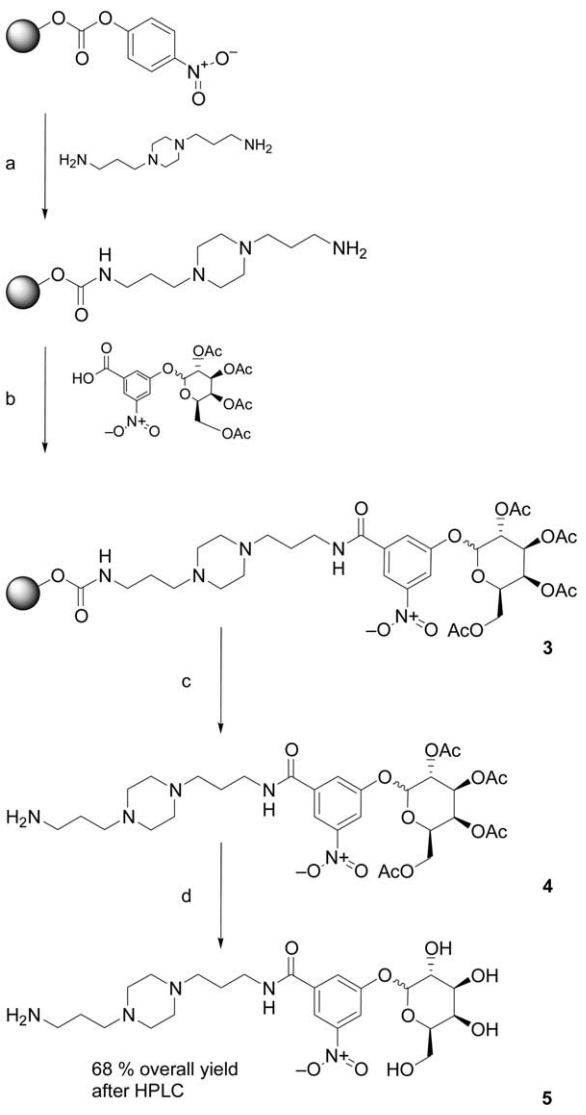


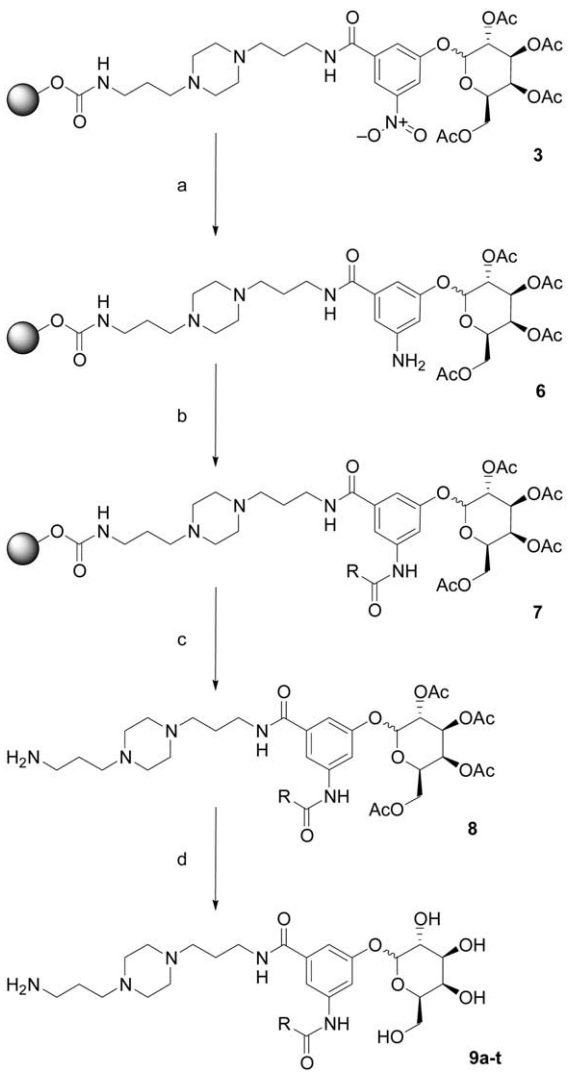
Figure 2. Previously determined co-crystal structures of compounds **1** (A, yellow and green) and **2** (B, gold) with LTB_5 and CTB_5 respectively.[10]

1 and **2**, however, the nitro group was not observed displacing water #2, but instead directed out of the binding site toward bulk solvent while the displacement of water #2 was accomplished by an amide group on the other side of the benzene ring. This brought into question the role of the nitro group as an essential component. Secondly, multiple binding modes allow the morpholine ring to make interactions in several places within the GM1 binding site thus complicating the interpretation of the crystal structures. The slightly longer linker of **1** allows the morpholine ring to occupy a consensus position near Arg 35 flanked by Lys 34 and the face of the aromatic ring of Tyr 12 (Fig. 2A). The morpholine ring of **2** also makes interactions there, as well as with a hydrophobic region of the binding site formed by Ile 58 and Lys 34 (Fig. 2B).

Based on this dual binding mode, we hypothesized that the additional interaction of a second morpholine ring might lead to a significant improvement in affinity leading us to design compound **2a**. However, we found this compound to have significantly lower affinity ($K_d = 60$ μM) than **2** (12 μM) for LTB_5.[10] Taken together, these issues presented a major challenge to formulating a strategy for the design of subsequent antagonists and underscore the difficulty in applying a classic structure-based drug design strategy to a system with a shallow binding site lacking well-defined pockets. Regardless, the observed binding modes of **1** and **2** provided a rough framework for designing a new series of compounds. Our first goal was to solve the crystal structure of **2a** bound to either CT or LT B_5 in order to find why it was a weaker binder than its parent compound **2**. Secondly, we set out to probe the effect of replacing the aforementioned nitro group with a set of larger and more diverse substituents that could potentially interact with unexploited regions of the GM1 binding site. Our rationale was that if the binding mode of the new ligands allowed water #2 to be displaced by the amide carbonyl *meta* to the new group, as seen in **1** of Fig. 2A (yellow conformers), then new favorable interactions might be made between this new group and residues lining either side of the binding site such as Arg/His 13, Ile 58, or Lys 34. Alternatively, if the binding mode were more like that of **1** where water #2 is not displaced (Fig. 2A, green conformers), then the new group would likely be interacting only with hydrophobic residues lining the inside wall of the binding pocket such as Ile 58 and its neighbors. In order to find out which scenario would dominate, we needed to first develop an efficient synthetic route for the synthesis of 3,5-substituted phenylgalactosides meeting our solubility and linkability requirements. To our knowledge, very few synthetic routes to 3,5-substituted phenylgalactosides have been reported, especially for compounds carrying a significant degree of diversity. Therefore, the creation of an efficient strategy for introducing further diversity into the non-galactoside portion of MNPG was considered a much-needed advance. To accomplish this, we chose to design the new library around compound **5** (Scheme 1), an analogue of compound **1** that employs a dialkyl piperazine moiety as a quasi-isosteric replacement for the propyl morpholine ring. The use of 1,4-bis(3-aminopropyl)piperazine allowed us to develop a convenient solid phase strategy for making linkable, 3,5-substituted phenylgalactosides in good yield. This route enabled us



Scheme 1. (a) excess 1,4-bis(3-aminopropyl)piperazine, CH_2Cl_2; (b) 5 equiv α/β (1:1) CNPG, DIC, HOBt, CH_2Cl_2/DMF; (c) TFA/CH_2Cl_2; (d) NaOMe/MeOH.



Scheme 2. (a) saturated $SnCl_2 \cdot H_2O$, NMP; (b) RCOCl, DIEA, CH_2Cl_2; (c) TFA/CH_2Cl_2; (d) NaOMe/MeOH.

to prepare a focused library of 20 compounds in parallel, where diversity was introduced upon acylation of an aromatic amine produced by on-bead reduction of the nitro group of **5**.

2.2. Focused library and precursor synthesis

The synthetic route to the focused library **9a–t** of Scheme 2 was first validated for steps leading to the nitro group-containing parent compound **5** of Scheme 1. In general, *p*-nitrophenylcarbonate Wang resin was first treated with an excess of 1,4-bis(3-aminopropyl)piperazine in either CH_2Cl_2 or THF to give the resulting immobilized diamine. Next, five equivalents of the previously described compound CNPG (3-carboxy-5-nitrophenyl-2,3,4,6-tetra-*O*-acetyl-galactopyranoside)[10,14] as a 1:1 anomeric mixture was activated using DIC/HOBt and coupled to the immobilized diamine to give the protected, resin-bound intermediate **3**. After cleavage with 1:1 TFA/CH_2Cl_2, and removal of acetate protecting groups using sodium methoxide in methanol, crude **5** was purified by HPLC yielding an inseparable 1:1 mixture of α and β anomers in roughly 70% isolated yield overall.

After validation through compound **5**, the methodology was extended to the parallel synthesis of focused library compounds **9a–t**, shown in Scheme 2. Efficient reduction of the nitro group of resin-bound **3** was performed using a saturated solution of $SnCl_2 \cdot H_2O$ in either *N*-methyl pyrrolidinone (NMP) or 1:1 DMF/CH_2Cl_2. Acylation of the resulting aromatic amine **6** was then carried out by treatment with a variety of acid chlorides in the presence of *N*,*N*-Diisopropylethylamine (DIEA). Cleavage, deprotection, and purification as described for compound **5** afforded **9a–t**. The identity and characterization of focused library members and control compounds are shown in Table 1. Isolated overall yields for the six steps range from 7 to 62 percent with an average around 22%. In general, these yields are substantially higher than those of previously reported compounds obtained using our solution phase method

Table 1. Physical properties and characterization of compounds[a–c]

Compd	*N*-acyl substituent[a]	ESI-MS Calcd. for $[M+H]^{+1}$	ESI-MS Found	Yield[b] (%)	RP-C18 R_t (min)	CLogP[c]
MNPG	None	n.d.[d]	n.d.[d]	n.d.[d]	n.d.[d]	−0.171
5	None	528.26	528.1	65	< 15 min	−1.51
9a	Propionyl	554.31	554.3	20	16.9	−2.08
9b	Butyryl	568.33	568.3	25	17.6	−1.55
9c	Pentenoyl	580.33	580.3	7	18	−1.51
9d	Isovaleryl	582.34	582.4	10	18.6	−1.15
9e	Ethylbutyryl	596.36	596.4	15	17.1	−0.714
9f	Phenylacetyl	616.33	616.4	62	19	−1.07
9g	Cyclopentylacetyl	608.36	608.4	17	19.6	−0.58
9h	2-Thiopheneacetyl	622.28	622.3	27	18.5	−1.43
9i	Cyclopentylpropionyl	622.37	622.4	17	21.2	0.01
9j	2-Ethylhexanoyl	624.39	624.4	16	21.2	0.34
9k	Hydrocinnamoyl	630.34	630.4	13	20.0	−0.51
9l	Phenoxyacetyl	639.32	632.3	18	19.6	−1.1
9m	Fluorophenylacetyl	634.31	634.3	37	19.6	−0.93
9n	Benzyloxyacetyl	638.31	638.4	16	19.9	−0.83
9o	4-Methoxyphenylacetyl	646.34	646.4	31	19.3	−1.15
9p	phenoxybutyryl	660.35	660.4	12	20.8	−0.29
9q	3,4-Dimethoxyphenylacetyl	676.35	676.4	37	18.6	−1.41
9r	(−)-Menthoxyacetyl	694.43	694.5	15	24.6	1.04
9s	2-Bromophenylacetyl	694, 696	694.5	27	20.1	−0.2
9t	Benzyloxyphenylacetyl	722.37	722.4	24	23.5	0.61

[a] Derived from the corresponding acid chloride.
[b] Overall isolated yield for six steps except for **5** (four steps).
[c] Calculated using ChemDraw Pro version 7.0.
[d] Not determined (commercially available compound).

that averaged less than 20% for only three steps.[10] The method also allows larger numbers of compounds to be prepared in parallel faster and more conveniently as well, although some limitations exist in the number of possible diamines and sufficiently reactive acid chlorides that could be practically employed.

2.3. Library screening and IC_{50} determinations

After preparative scale HPLC, library compounds **9a–t** were estimated to be greater than 95% pure based on analytical HPLC chromatography analysis. An initial library screening was performed using our previously reported Direct Enzyme-Linked Assay (DELA) at a concentration of 100 μM.[8] The results indicate that, as a series, almost all of the tested compounds are substantially better than the 1st generation lead MNPG (Fig. 3). Inhibition values range from 14 to 40% for the newly synthesized antagonists while values for MNPG remain less than 5%. The low solubility of the benzyloxyphenylacetyl derivative **9t**, indicated by severe cloudiness in the assay buffer solution, resulted in very poor inhibition. Although, these results likely represent a rather modest affinity gain for the series, some useful information can be obtained. Since the letter designation of the compounds in Figure 3 advance with increasing steric bulk, it can be seen that a slightly higher range of inhibition values is observed in compounds carrying larger *N*-acyl substituents. For example, **9k–s** as a group, clearly shows more inhibition than **9a–j** as a group. This is most likely a result of the larger substituents being able to make more favorable non-specific interactions with various regions of the GM1-binding site relative to the smaller ones. However, the potency of the compounds screened does not vary enough to easily establish a rank ordering. For this reason, we decided to pick four of the best compounds for more detailed studies that were also amenable to re-synthesis on a scale that would provide enough material to obtain full binding curves and also for co-crystallization with CTB_5 and LTB_5.

The chemical structures of these compounds, **9h**, **p**, **q**, and **s**, along with the parent compound **5** are shown in Figure 4. With the exception of **9p**, all four of these compounds incorporate specifically an *N*-phenylacetyl-type group, indicating a preference for this type of modification within the newly synthesized series. Binding curves and IC_{50} values for these compounds are shown in Figure 5. The IC_{50} values range from 200 μM to 790 μM indicating roughly a 2- to 6-fold improvement over MNPG alone, which has an IC_{50} regularly determined in this assay to be in the range of 1–1.4 mM.

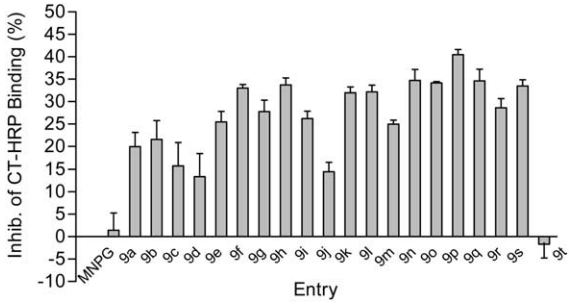


Figure 3. Library screening results. GD_{1b} Direct Enzyme Linked Assay (DELA) screening. Compounds were screened at 100 μM for the ability to block CT-HRP conjugate binding to GD_{1b}-coated microtiter plates. Error bars represent the standard error of the mean for three independent experiments (Calculated using the Prism software package, version 3.0, GraphPad Software Inc., San Diego, CA).

However, it is apparent that only one compound **9h**, containing the 2-thiophene acetyl group is better than the nitro group-bearing **5**. The determination of X-ray crystal structures of the key compounds of this study bound to CTB_5 or LTB_5 was undertaken to reveal binding modes and to obtain more insight into the underlying reasons for the differences in affinity between these compounds.

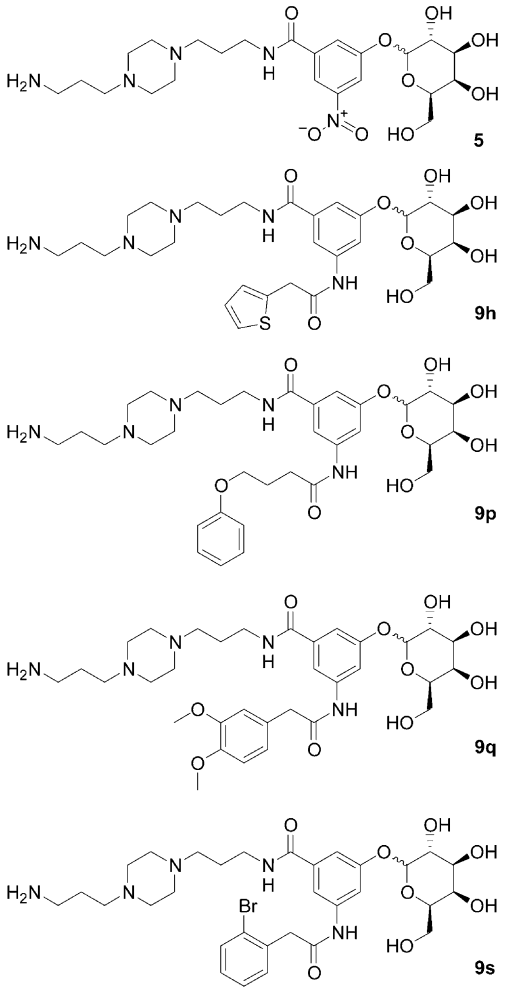


Figure 4. Structures of compounds selected for further study.

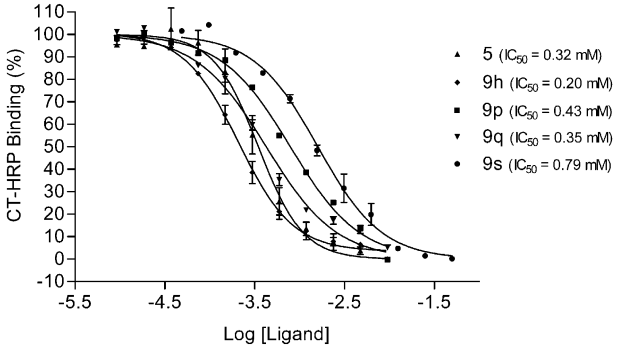


Figure 5. Binding curves for select compounds. IC_{50} values calculated using GraphPad Prism version 3.0, GraphPad Software Inc. San Diego, CA. Error bars represent the standard error of the mean for at least two replicate data sets.

3. Crystallographic studies

The X-ray crystal structures of **2a**, **5**, and **9h** bound to either CTB_5 or LTB_5 were solved at resolutions ranging from 1.35 Å to 2.0 Å (Table 2) with the goal of establishing a more clearly defined structure/activity relationship as well as for rationalizing observed trends in affinity. For all three compounds, the galactose moiety of the ligand is well defined and adopts the canonical binding mode within the GM1 binding site. However, in some cases, the entire phenyl galactoside aglycone is able to adopt various conformations leading to multiple binding modes in a manner similar to that reported for compounds **1** and **2**. Additionally, since 1:1 anomeric mixtures of **5** and **9h** were used for co-crystallization at concentrations well into the millimolar range, electron density corresponding to the β anomer is visible in a few of the binding sites. Thus, interpretation of the electron density maps is challenging in general with multiple conformations and two anomers present. The details of each new structure are presented below. Data collection and refinement statistics are shown in Table 2.

The LTB_5–**2a** complex was solved at a resolution of 2.0 Å. In all 5 binding sites, a single binding mode is observed containing well-ordered density from the galactose through the nitrobenzamide moiety (Fig. 6). In each case, canonical water #2 is displaced by the nitro group of the ligand, as observed in the LTB_5-MNPG structure. This is in contrast to the crystal structure of its parent compound **2** where two binding modes were observed differing by a torsion angle of 180° (Fig. 2B). Beyond the nitrobenzamide moiety, density corresponding to the rest of the molecule is less clear. In the D and H binding sites, there is no appreciable density corresponding to the dimorpholine portion of the ligand. The model of compound **2a** was truncated to match the connective density seen in two binding pockets D and H. In sites E and F, the electron density is discontinuous, but a preferred conformation among multiple binding modes can be discerned (not shown). In the remaining site G, a single conformation of the whole ligand including the two morpholine rings can be seen from the difference fourier density (Fig. 6).

The morpholine rings are situated in the binding pocket regions perpendicular to the plane of the nitrophenyl ring. There are relatively few interactions made by either of the morpholine rings with their own binding pocket. It is possible that the rings are associated with a local water mediated hydrogen-bond network, but no clear density for nearby waters is seen. The "lower" morpholine rings rests consistently, in all 5 subunits, in the space between Arg 13 and Tyr 12. In the F and G subunits, this same morpholine ring comes within 3.5 Å of Arg 13 (Fig. 7A). In the E subunit, this ring is instead making a number of Van der Waals contacts with a crystallographically related Thr 19. In the G subunit, the other morpholine ring makes Van der Waals interactions with a crystallographically related Gln 16. There are also Van der Waals contacts made to the carbon atoms of the same morpholine ring by a crystallographically related oxygen of Tyr 18. It appears that in

this site, the dimorpholine moiety is constrained to remain perpendicular to the plane of the MNPG by contact with a symmetry related copy of the toxin pentamer in the crystal lattice as a 90° rotation of the morpholine rings in either direction would bring one of them into steric conflict with this molecule as well as with Arg 13 of their own binding site. Thus, it is clear from the density that any potentially new favorable interactions coming about from the addition of the second morpholine ring to **2** are not strong enough to drive the nitrophenyl ring away from the canonical MNPG binding mode in addition to being sterically unfavorable. These observations lead to the conclusion that **2a** is a less viable candidate for continued optimization compared to **5** or **9h**.

The CTB_5–**5** complex was solved at a resolution of 1.46 Å in a crystal form with one B pentamer per asymmetric unit providing five images of the binding mode of compound **5**. Sites E and F displayed density corresponding to primarily the α anomer (Fig. 8A). The three remaining sites contained density corresponding primarily to the β anomer (Fig. 8B). In most receptor binding sites there was at least some evidence of partial occupancy for both anomers. However, the density for one anomer was usually so much weaker than the other that it was disregarded in the final model for any particular site. Each of the α anomers retains the MNPG binding mode with the nitro group of the phenyl ring displacing canonical water #2 (Fig. 7B). In binding sites containing the β anomer, water #2 is not displaced. Instead, in

Table 2. Data collection and refinement statistics

Data collection statistics	LTB_5 + **2a**	CTB_5 + **5**	CTB_5 + **9h**
Source	APS 19ID	ALS 8.2.2	APS 19ID
Wavelength (Å)	0.97934	0.98000	0.97945
Resolution (Å)	1.99	1.46	1.35
Spacegroup	$P2_1$	C2	C2
Number of observations	124930	274226	342770
Unique reflections	34499	85036	104310
Completeness (%)	95.3 (84.1)	98.5 (93.2)	95.5 (68.4)
$<I/\sigma(I)>$	19.3 (7.1)	19.7 (8.6)	22.10 (1.9)
R_{merge}	6.1 (14.5)	4.8 (11.5)	5.4 (40.2)
Refinement statistics			
Number of residues	515	515	515
Water molecules	575	650	629
Ligand molecules	5	5	5
Ligand atoms — Site D	23	37 (full ligand)	21
Ligand atoms — Site E	39 (full ligand)	37 (full ligand)	21
Ligand atoms — Site F	39 (full ligand)	37 (full ligand)	21
Ligand atoms — Site G	39 (full ligand)	37 (full ligand)	21
Ligand atoms — Site H	28	37 (full ligand)	37
R_{cryst}	0.154	0.115	0.132
R_{free}	0.221	0.159	0.173
Figure of merit	0.869	0.926	0.895
RMSD from ideal			
Bond lengths (Å)	0.011	0.011	0.013
Angles	1.158	1.485	1.626
Chirality	0.067	0.095	0.112

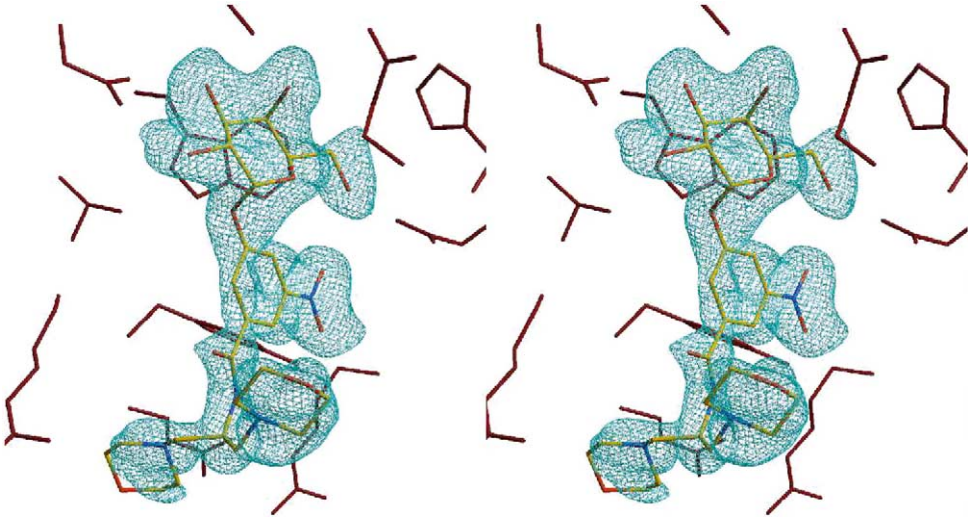


Figure 6. Electron density and model for **2a** in complex with subunit G of LTB_5. Stereopair showing experimental electron density is contoured at 2σ in a σ_A-weighted (mFo-DFc) difference map.

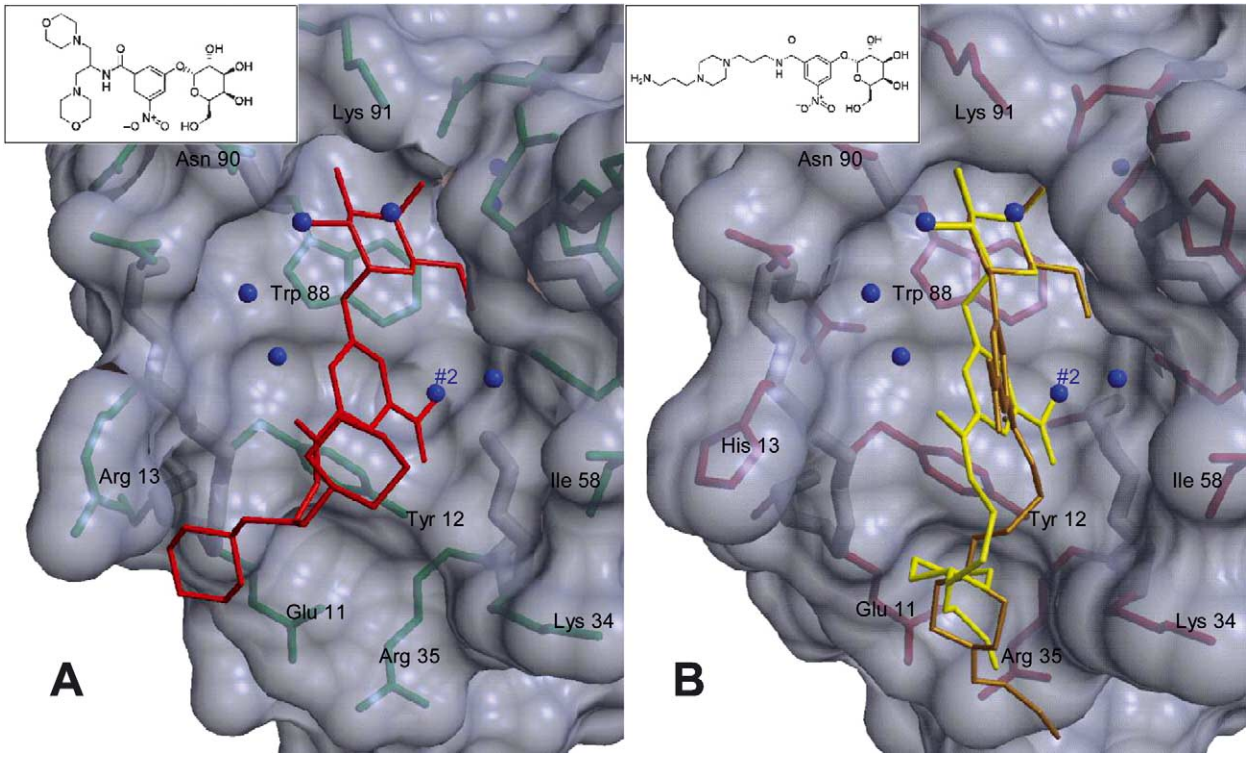


Figure 7. Representative crystallographic models for compound **2a** and anomers of **5**. A: Final crystallographic model of **2a** (red) bound in GM1 binding site G of LT. B: Representative models of the α (yellow) and β (gold) anomers of **5** superimposed within GM1 binding site G of CT.

binding site D, the nitro group comes within 3 Å of the terminal nitrogen from a symmetry related Lys 34 residue. In site H, the nitro group forms a hydrogen bond with the nitrogen of the primary amide from a symmetry related Gln 16 residue. In both cases, the amide carbonyl oxygen forms a hydrogen bond to water #2. The nitro group in site G does not appear to make any specific contacts.

In all five binding sites of the CTB_5–**5** complex, the difference map showed clear density for the piperazine ring and distal linker terminating in a primary amine. Due to the anchoring of the MNPG scaffold, the remaining piperazine and linker are able to adopt essentially only one binding mode. This is in contrast to the structure of compound **1**, where two families of conformations differing by a 120° rotation of the nitrobenzamide ring were observed.[10] Specifically, the piperazine ring sits directly above and forms contacts with Tyr 12 and Glu 11, and continues with the primary amine tail stretching between residues Arg 35 and Lys 34, in a manner similar to that seen in the previous structure of **1**. In addition, for several binding sites, the primary amine forms hydrogen bonds with the extensive water network near Arg 35.

In the case of **9h** in complex with CTB_5, the structure was solved to 1.35 Å (Fig. 10A). All five binding sites present in the asymmetric unit again show clear density through the nitrobenzamide moiety. However, only in one site can the majority of **9h** be modeled into the electron density maps (Fig. 9). In the remaining binding sites, only the nitrobenzamide of **9h** could be modeled since the remainder most likely adopts a rather large number of conformations. The nitrobenzamide ring always sits essentially perpendicular to that of the canonical MNPG binding mode (Fig. 10B). The reason for this seems to arise from a steric clash with residues Gly 33 and Lys 34 in the water #2 region of the binding pocket when the *N*-acetyl thiophene moiety tries to occupy this position formerly occupied by the much smaller nitro group. As a result of this steric hinderance, the benzene ring of **9h** may direct the *N*-acetyl thiophene toward several hydrophobic residues lining the upper rim of the binding site. However, since there is no density corresponding to the thiophene ring, or its adjacent linker, it has not been modeled. The carbonyl oxygen of the amide *meta* to the *N*-acetyl thiophene forms a hydrogen bond with a backbone amide nitrogen from His 13 as well as a water-mediated hydrogen bond to Trp 88.

While only binding site H contains relatively clear density for most of **9h**, all 5 binding sites have density in the difference fourier map corresponding to the piperazine ring resting on Tyr 12 of the binding site, as in the case of **5**. The observation of this consensus position for the piperazine-containing linker indicates that favorable, albeit non-specific interactions are being made. Unfortunately, the steric bulk of the *N*-acetyl thiophene group appears to prevent highly ordered binding of the ligand, as indicated by a lack of continuous electron density in most binding sites. Nevertheless, the fact that **9h** is the most potent compound presented here, if only by a factor of two over compound **5**, may indicate that further optimization of this lead would be worthwhile.

4. Conclusions

We have developed a robust solid-phase synthesis of 3,5-substituted phenyl galactosides providing a convenient route to a potentially important class of antagonists of CT and LT receptor-binding. The synthetic strategy provides compounds with good aqueous solubility in fairly good yields, an improvement over previously reported syntheses. The general synthetic method could be easily adapted to incorporate a wider array of pyranose sugars or even more complex oligosaccharides thus making it possible to generate libraries of compounds targeting other AB_5 bacterial toxins or potentially other proteins that recognize and bind to carbohydrate receptors. The primary amine unmasked upon cleavage from the solid support provides a convenient point for further derivatization or attachment to multivalent scaffolds or other macromolecules.

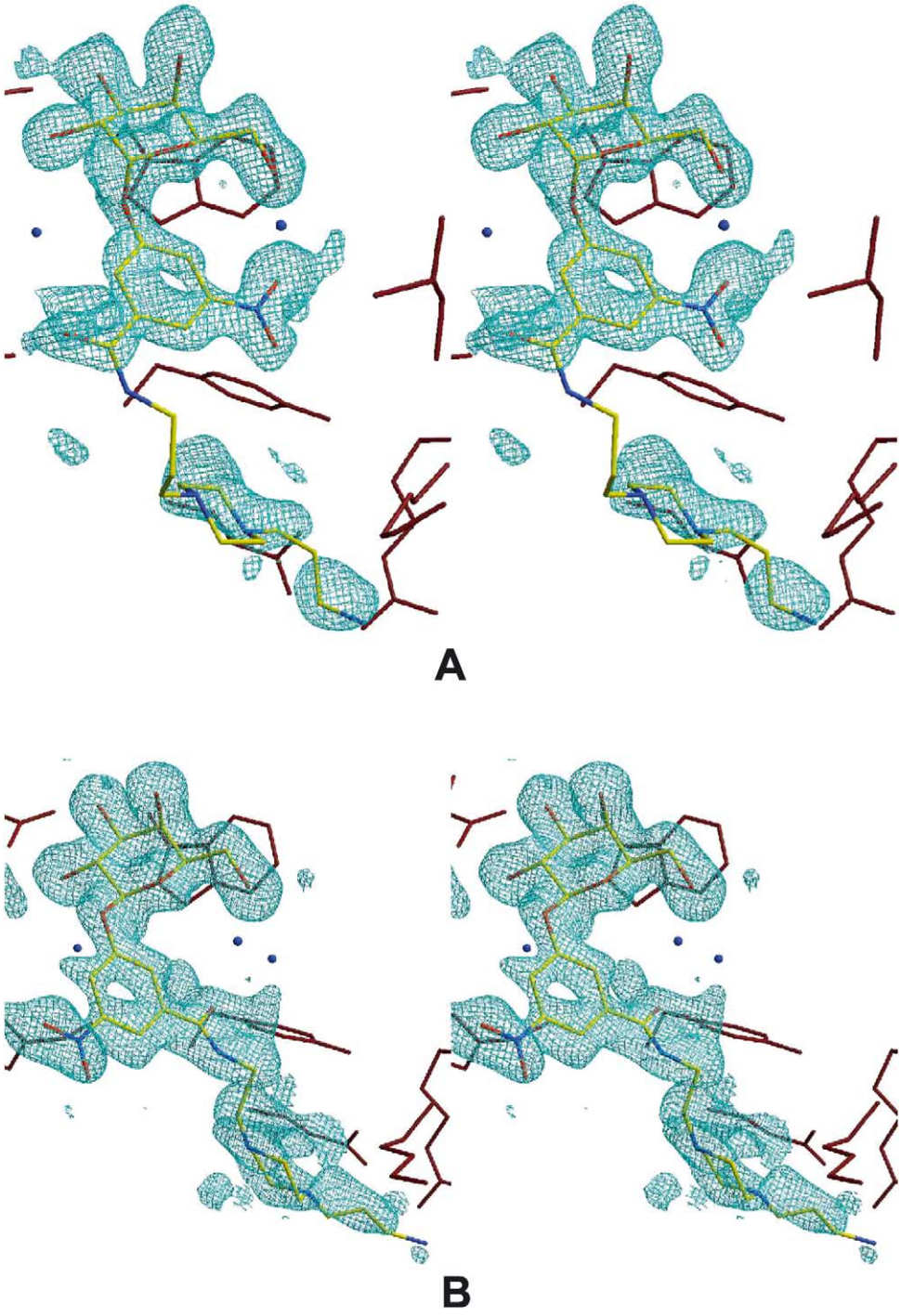


Figure 8. Electron density and models for compound **5** in complex with CTB_5. Experimental electron density showing the α anomer in receptor binding site F (stereopair A) and the β anomer in receptor binding site D (stereopair B) of **5** is contoured at 2σ (blue) in a σ_A-weighted (mFo-DFc) difference map.

Taken together, the assay and crystallographic results indicate that the *m*-nitrophenyl moiety is important for affinity in these compounds. More specifically, it appears that replacing the nitro group, which functions as a hydrogen bond acceptor and displaces a water molecule, with hydrophobic groups doesn't provide enough favorable interactions within the series presented here to produce a large overall affinity gain despite retaining the ability to displace water #2 by the amide carbonyl *meta* to the *N*-acyl group. This is likely one reason that even the most potent compound presented here, **9h** is only slightly better than the nitrophenyl-containing compound **5** in terms of IC_{50} values. Furthermore the crystal structure of **9h** bound to CTB_5 shows that the bond between the central benzene ring and the sugar ring is forced to adopt an unusual torsion angle in order that the binding site can accommodate the steric bulk of the *N*-acetylthiophene moiety which seems slightly too short to make optimal contact with the upper rim of the binding site. This binding mode is different than either of the two predicted possibilities envisioned when the series was designed. To further

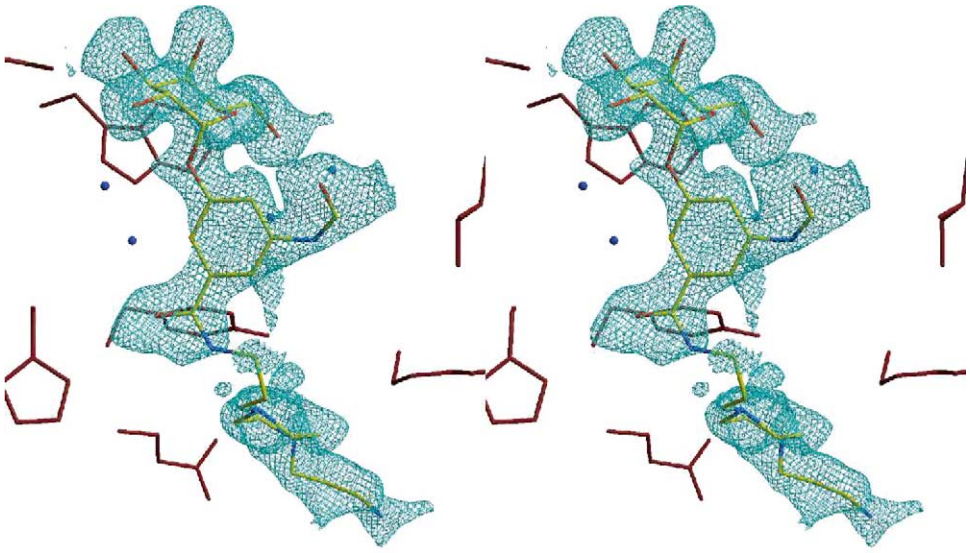


Figure 9. Electron density and model for compound **9h** in receptor binding site H of CTB_5. Stereopair showing experimental electron density is contoured at 1.5σ in a σ_A-weighted (mFo-DFc) difference map.

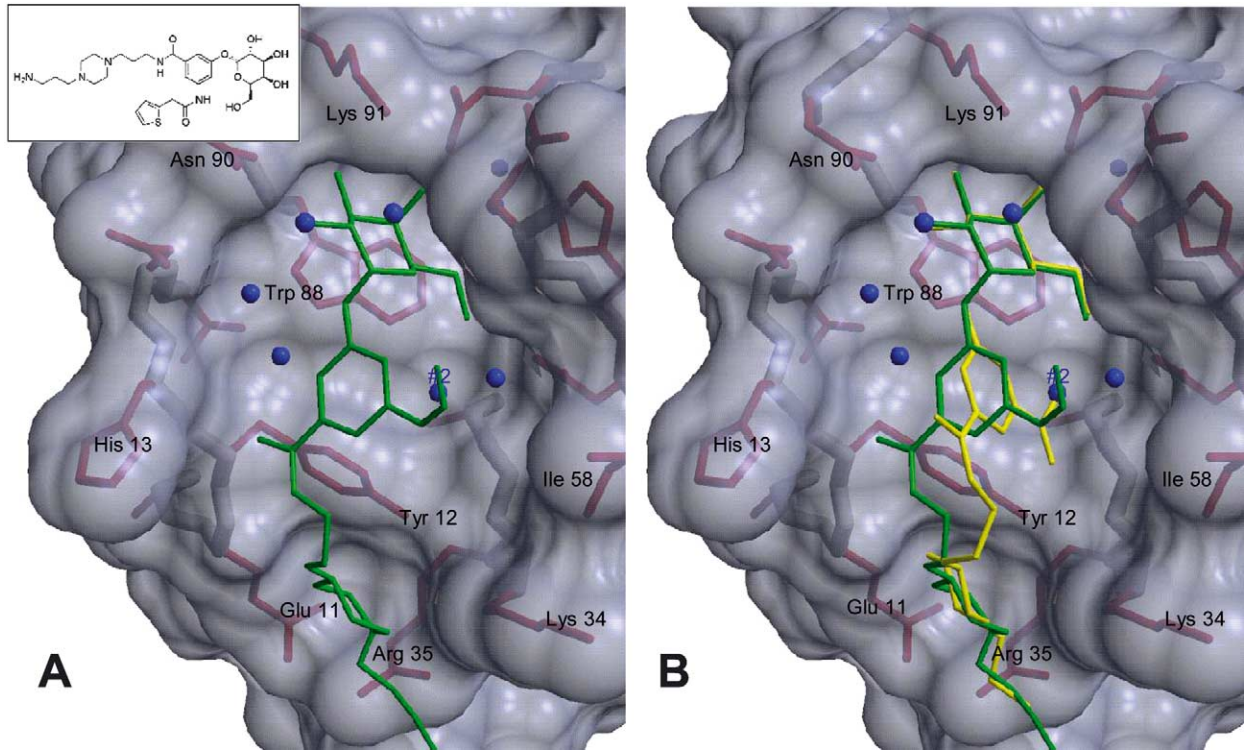


Figure 10. Comparison of models for compound **5** and **9h**. A: Representative crystallographic model for compound **9h** (green) bound in receptor binding site G of CT (A). B: Superimposition of α anomers of **5** (yellow) and **9h** (green) within the GM1 binding site (B). Note: only part of **5** and **9h** visible in electron density are depicted.

improve affinity, optimization of the non-nitrophenyl portions of compound **5** may prove more useful than further elaboration of **9h** or similar *N*-acylated derivatives. Alternatively, future ligand design efforts that seek to retain the hydrogen bonding properties of the nitro group as well as adding new hydrophobic interactions may result in more dramatically improved affinity.

Despite variations in the orientation of the aryl ring within the series of compounds examined crystallographically here, the structures of **5**, **9h** and to a lesser extent **2a**, reveal that the piperazine or morpholine ring prefers to reside in a consensus pocket near the face of Tyr 12, the guanidine group of Arg 35, and aliphatic portion of Lys 34, in agreement with the binding mode observed for the quasi-isosteric lead compound **1**.[10] Thus, targeting this 'hot spot' with a propylmorpholine or propylpiperazine derivatives will be a key feature of continued lead optimization efforts.

5. Experimental

5.1. Synthetic chemistry general

Unless otherwise noted, commercially available reagents and solvents were used as purchased without further purification. All acid chlorides were purchased from Aldrich Chemical Co. with the exception of 4-phenoxybutyryl chloride and cyclopentylacetyl chloride, which were purchased from Lancaster Synthesis Inc. ^{1}H NMR spectra were recorded using a Bruker AC-300 operating at 300 MHz. Samples were dissolved in methanol-d_4 and not subjected to spinning during spectra acquisition. For compound **9h**, the ^{1}H NMR peak assignments are reported in full, including those for the core structure. Since core structure peak positions were essentially identical for all lead compounds, the peak listings for **9h**, **p**, **q**, and **s** refer to the *N*-acyl substitutent only. LC-MS characterization was performed using an Agilent 1100 series instrument equipped with a variable wavelength detector coupled to an Agilent 1100 MSD-Trap SL electrospray ion trap mass spectrometer operating in positive ion mode. Reported values for the molecular ion $[M+H]^+$ of target compounds refer to those obtained during final LC-MS analysis after purification. HPLC detection and extinction coefficients employed a wavelength of 250 nm. An extinction coefficient of either 1738 M^{-1} cm^{-1} or 2985 M^{-1} cm^{-1} was used to calculate yields and concentrations prior to the initial screening assay. Columns used for both preparative C18 reverse-phase HPLC purification and LC-MS characterization were purchased from Vydac (21×250 mm, 10–15 μm) or Agilent Technologies (4.6×150 mm).

5.2. Synthesis of *N*-{3-[4-(3-Amino-propyl)-piperazin-1-yl]-propyl}-3-nitro-5-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-benzamide (5)

The synthesis of the nitrophenyl galactopyranoside **5** (Scheme 1) was carried out in either a single 10 mL BioRad biospin™ column or in parallel using an Argonaut Technologies Quest210 synthesizer under inert atmosphere. A general procedure is as follows: To a 10 mL column, 250 mg of *p*-nitrophenyl carbonate Wang resin (Novabiochem, loading = 0.6 mmol g^{-1}) was added. The resin was allowed to swell for 15 min in 4 mL CH_2Cl_2. 1,4-bis-(3-aminopropyl)piperazine (500 μL, ~20 equiv) was then added. Agitation of the resin was carried out for 1.5 h followed by washing twice with CH_2Cl_2. The coupling step was then repeated once more.

After the second washing, 4 mL of CNPG (3-carboxy-5-nitrophenyl-2,3,4,6-tetra-*O*-acetyl-galactopyranoside) stock solution was dispensed into the reaction vessel. The CNPG stock solution consisted of: CNPG (α:β = 1:1) 1.0 g (1.9 mmol) in 4 mL of 1:1 DMF/CH_2Cl_2, HOBt 400 mg (3 mmol), and 1,3-diisopropylcarbodiimide (DIC) 400 μL (2.5 mmol). The mixture was agitated for 3–4 h and then washed with CH_2Cl_2 and EtOAc in an alternating fashion five times. Cleavage of acetate-protected compound **3** from the solid support was achieved by suspending the resin in 4 mL a solution of 1:1 TFA/CH_2Cl_2 and agitating for 30 min. The solvent was then removed by rotary evaporation to give the crude acetate-protected product **4** which was then purified by HPLC (solvent A, 0.1% TFA in water; solvent B, CH_3CN gradient 0–80% B over 30 min) affording 72 mg of pure **4**. Yield: 68% isolated overall yield.

Aliquots of **4** were quantitatively deprotected by dissolving the sample in 2 mL of 100 mM NaOMe/MeOH and then stirred for 30 min. The samples were then brought to pH 3–4 by adding a drop of glacial acetic acid and desalted using either a Waters Sep-Pak™ C18 cartridge or re-purification by HPLC. After removal of the HPLC solvent, the pure compound **5** was characterized using LC-MS and ^{1}H NMR. The approximate anomeric ratio of the product was 1:1 as determined by ^{1}H NMR. ESI-MS *m/z*: 528.0 $[M+H]^+$. ^{1}H NMR: δ 8.35 (s, 1H), 8.18 (s, 0.5H, α), 8.11 (s, 0.5H, β), 7.99 (s, 0.5H, α), 7.92 (s, 0.5H, β), 5.69, 5.68 (d, 0.5H, α anomeric proton), 5.04, 5.01 (d, 0.5H, β anomeric proton), 3.97–3.63 (m, 5H), 3.52 (t, 4H), 3.17 (t, 4H), 3.03 (t, 4H), 2.68 (t, 4H), 2.05 (quintet, 2H), 1.89 (quintet, 2H).

5.3. Library synthesis (9a–t)

Parallel solid-phase synthesis of compounds **9a–t** was carried out using an Argonaut Technologies Quest210 synthesizer under inert atmosphere. To start, 50 mg of *p*-nitrophenyl carbonate Wang resin (Novabiochem, loading = 1.6 mmol g^{-1}) was loaded into each reaction vessel. The resin was allowed to swell for 15 min in 1 mL CH_2Cl_2. 1,4-bis-(3-aminopropyl)piperazine (100 μL, ~20 eq., Aldrich) was then added to each vessel. Agitation of the resin was carried out for 1.5 h where the resin was washed twice with CH_2Cl_2. The coupling step was then repeated once more.

After the second washing, 1 mL of CNPG (3-carboxy-5-nitrophenyl-2,3,4,6-tetra-*O*-acetyl-galactopyranoside) stock solution was dispensed into each reaction vessel. The CNPG stock solution consisted of: CNPG 1.25 g (2.4 mmol) in 10 mL of 1:1 DMF/CH_2Cl_2, HOBt 400

mg (3 mmol), and 1,3-diisoropylcarbodiimide (DIC) 400 μL (2.5 mmol). The mixture was agitated for 3 h and then washed with CH_2Cl_2 and EtOAc in an alternating fashion five times. The resin-bound nitrophenyl galactopyranoside **3** in each reaction vessel was treated with 1 mL of a saturated solution of $SnCl_2$ in 1:1 DMF/CH_2Cl_2 and agitated for 3 h. The resin carrying the immobilized aromatic amine **6** was washed with CH_2Cl_2 and EtOAc in an alternating fashion five times. Acylation of **6** was achieved through the addition to each vessel, a solution containing one of 20 possible acid chlorides containing the appropriate acid chloride in excess (250 mg), 750 mL of CH_2Cl_2, and 250 μL *N*,*N*-diisopropylethylamine (DIEA). The resin was agitated for 4 h and then washed five times with CH_2Cl_2. Cleavage of the protected, acylated library **7a**–**t** from the solid support was achieved by suspending the resin samples in 1 mL a solution of 1:1 TFA/CH_2Cl_2 and agitating for 30 min. The solvent was then removed by rotary evaporation and dried for an h under high vacuum to give the penultimate series **8a**–**t**. The acetate groups of **8a**–**t** were removed by dissolution of each sample in 2 mL of 100 mM NaOMe/MeOH with stirring for 30 min. The samples were then brought to pH 3–4 by adding a drop of glacial acetic acid and purified by preparative HPLC (solvent A, 0.1% TFA in water; solvent B, CH_3CN gradient 0–80% B over 30 min). After removal of the solvent, purified library compounds **9a**–**t** were characterized by LC-MS. Yields were determined spectrophotometrically and are reported along LC-MS data in Table 1.

5.4. Characterization of lead compounds 9h, p, q, and s

5.4.1. *N*-{3-[4-(3-Amino-propyl)-piperazin-1-yl]-propyl}-3-(2-thiophen-2-yl-acetylamino)-5-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-benzamide (9h). ESI-MS *m*/*z*: 622.3 $[M+H]^+$. ^1H NMR: δ 7.84 (s, 1H), 7.45 (s, 0.5H), 7.40 (s, 0.5H), 7.30–7.27 (m, 1.5H), 7.0–6.95 (m, 2H), 5.55 (d, 0.5H, α anomeric proton), 3.90 (s, 2H), 3.9–3.56 (m, 5H), 3.48 (t, 4H), 3.15 (t, 4H), 3.02 (t, 4H), 2.62 (t, 4H), 2.03 (quintet, 2H), 1.86 (quintet, 2H).

5.4.2. *N*-{3-[4-(3-Amino-propyl)-piperazin-1-yl]-propyl}-3-(4-phenoxy-butyrylamino)-5-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-benzamide (9p). ESI-MS *m*/*z*: 660.3 $[M+H]^+$. ^1H NMR: *N*-acyl substituent only, 7.32–7.21 (m, 2H), 6.91–6.87 (m, 3H), 4.05 (t, 2H), 2.71–2.57 (m, 4H, including 2H from overlapping resonance of core structure), 2.16, (quintet, 2H).

5.4.3. *N*-{3-[4-(3-Amino-propyl)-piperazin-1-yl]-propyl}-3-[2-(3,4-dimethoxy-phenyl)-acetylamino]-5-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-benzamide (9q). ESI-MS *m*/*z*: 676.3 $[M+H]^+$. ^1H NMR: *N*-acyl substituent only, 6.97–6.86 (m, 3H), 3.83 (s, 3H), 3.81 (s, 3H), 3.62, (s, 2H).

5.4.4. *N*-{3-[4-(3-Amino-propyl)-piperazin-1-yl]-propyl}-3-[2-(2-bromo-phenyl)-acetylamino]-5-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-benzamide (9s). ESI-MS *m*/*z*: 694, 696.2 $[M+H]^+$. ^1H NMR: *N*-acyl substituent only, 7.42–7.17 (m, 4H), 3.9 (s, 2H).

5.5. Library screening and IC_{50} measurements

The GD1b Direct Enzyme Linked Assay (DELA) was used as previously reported for both the initial library screening and IC_{50} measurements for lead compounds **5**, **9h**, **p**, **q**, and **s**.[8] Samples consisted of 6 ng/mL CTB_5 pentamer conjugated to horseradish peroxidase (CTB-HRP, List Biological Labs, Campbell, California) incubated for 2 h with synthesized compounds. Screening was carried out in duplicate for each anomeric mixture at a single concentration of 100 μM. Inhibition values reported in Figure 3 are reported as the average of three independent determinations. IC_{50} values were calculated from either duplicate or triplicate data sets of at least ten different concentrations of antagonists by nonlinear regression using the Prism software package (version 3.0, GraphPad Software, Inc.). Reported values are the average of at least two independent determinations.

5.6. Protein expression and purification

Toxin B pentamers were obtained essentially as described previously.[8] Briefly, the general procedures follow: Wild-type cholera toxin B subunit was expressed by *E. coli* strain Top10 containing a pBAD/CTB_5 vector. The vector is controlled by a pBAD promoter and contains an ampicillin resistance marker. Heat-labile enterotoxin B_5 was expressed by a pROFIT vector in JM105 cells. This vector is controlled by a λP_R-promoter and contains a kanamycin resistance marker. Cells were grown at 30 °C in Luria-Bertani Broth (LB) to OD_{600} of ~0.6 before overnight induction, accomplished by addition of 0.2% arabinose for CTB_5 and heat shock at 42 °C for LTB_5 protein. Whole-cell lysates were prepared by resuspending the cell pellet in lysis buffer (20 mM Tris–HCl pH 7.5, 0.1 mM DTT, 0.2 mM EDTA, 200 mM NaCl) and lysing the cells by French press. The supernatant from centrifuged lysate was collected and batch-bound to immobilized D-galactose resin (Pierce) for 30 min to 16 h at 4 °C, washed with Buffer G (50 mM Tris–HCl pH 7.4, 200 mM NaCl, 1 mM EDTA, 3 mM NaN_3), and eluted with Buffer G + 300 mM D-galactose (Fluka). Gal-column purified fractions of B_5 were further purified using a Superdex 75 sizing column (Pharmacia) eluting with Buffer G.

5.7. Crystallization

Crystals of LTB_5 complexed with **2a** grew from sitting drops containing 2 μL of 32% PEG 5000, 100 mM Tris–HCl pH 7.8, 50 mM NaCl, with 2 μL of pure LTB_5 at 5 mg/mL. Glycerol (10%) in mother liquor was used as a cryoprotectant for flash-freezing. Crystals grew in space group $P2_1$.

Crystals of CTB_5 complexed with **5** grew from sitting drops containing 2 μL of 42% PEG 350, 50 mM NaCl, 100 mM Tris–HCl pH 7.5, mixed with 2 μL of pure CTB_5 at 5 mg/mL. Mature crystals formed within 24 h of sitting drop preparation and were flash-frozen for collection at the synchrotron light source. No cryoprotectant was added. Crystals of space group C2 were

isomorphous to the previously determined structure of CTB_5 complexed with G_{M1}-OS (PDB code 3CHB).[15]

Crystals of CTB_5 complexed with **9h** grew from sitting drops containing 2 μL of 36% PEG 550 MME, 50 mM NaCl, 100 mM Tris–HCl pH 7.3 with 2 μL of toxin at 5 mg/mL. No cryoprotectant was added before flash-freezing the crystal. Crystals of space group C2 were isomorphous to the previously determined structure of CTB_5 complexed with G_{M1}-OS (PDB code 3CHB).[15]

5.8. Structure determination and refinement

Crystal diffraction data was collected primarily at synchrotron radiation level beamlines. Principle sources included the Advanced Light Source (ALS) at Berkeley National Lab, CA, and the Advanced Photon Source (APS) in Argonne National Laboratory, IL. X-ray diffraction data for LTB_5-**2a** and CTB_5-**5** were measured at a wavelength of 0.979 Å at the APS beamLine 19ID. X-ray diffraction data for CTB_5-**9h** was measured at a wavelength of 0.979 Å at ALS. Diffraction images were integrated and scaled with HKL2000 and TRUNCATE.[16,17] Phases were determined using the molecular replacement program molrep.[18] CTB_5 complexed with G_{M1}-OS (PDB code 3CHB) with ligand and waters removed, was used as a search model for CTB_5 data sets. LTB_5 complexed with *m*-carboxyphenyl-α-D-galactopyranoside, (PDB code 1DJR) with ligand and waters removed, was used as a search model for LTB_5 data sets.[19]

Rigid body, TLS, isotropic, and anisotropic model refinement were carried out in REFMAC,[20] and Arp/wArp[21] within the CCP4 suite of programs.[17] Individual anisotropic temperature factors were applied for higher resolution structures **9h** and **5** near the end of refinement. Individual anisotropic values varied markedly between subunits. Real-space manual modeling and automated water-picking was done with Xfit.[22] A general refinement scheme began with a rigid body refinement, and subsequent restrained refinements while increasing the resolution. Iterative building of the water shell involved automated and manual water placement with Xfit and Arp/wArp,[21] with SFCHECK[23] and WATCHECK[24] being used, along with manual inspection, to discern incorrectly modeled waters. TLS parameters and anisotropic refinement parameters for high-resolution structures were then employed during the final rounds of refinement. Partial ligand subtracted maps were used to help refine more 'flexible' regions of the ligands. MNPG subtracted maps gave good indications of proper piperazine and morpholine ring placement. R_{free}, established after scaling, was the quality determinant for all refinement steps. The geometric description of the ligands for refmac (the 'cif' file) was generated by inputting a ligand-coordinate file generated by the program MOE[25] to the website http://davapc1.bioch.dundee.ac.uk/prodrg. The output cif file was then manually inspected with reference to the Cambridge Structural Database before being used in REFMAC5. Figures were made with the aid of software programs MSMS and Raster3d.[26,27]

Accession numbers

Crystal structure coordinates for **2a**, **5**, and **9h** bound to toxin pentamers have been deposited in the Protein Data Bank (PDB entry codes 1PZI, 1PZJ, and 1PZK respectively).

Acknowledgements

This work was supported in part by the National Institutes of Health (grant AI44954 to EF and grant AI34501 to WGJH). We also would like to thank Stewart Turley for assistance with crystallographic data collection and Drs. Ethan Merritt and Christophe Verlinde for stimulating discussions and advice. We gratefully acknowledge the staff of beamline 8.2.2 at the Advanced Light Source, and the staff at the SBC beamline 19ID for assistance in data collection. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source were supported by the US Department of Energy, Office of Biological and Environmental Research, under contract number W-31-109-ENG-38. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the US Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

References and notes

1. WHO. (2003). World Health Organization, [http://www.who.int/csr/disease/cholera/impactofcholera/em].
2. Widdus, R. *Bulletin of the World Health Organization* **2001**, *79* (8), 713.
3. Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; Wartna, E. S.; van Zanten, B. A. M.; Witholt, B.; Hol, W. G. J. *Nature* **1991**, *351*, 371.
4. Spangler, B. D. *Microbiol. Rev.* **1992**, *56*, 622.
5. De Wolf, M. J. S.; Dams, E.; Dierick, W. S. H. *Biochim. Biophys. Acta* **1994**, *1223*, 296.
6. Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; van Zanten, B. A. M.; Berghius, A. M.; Hol, W. G. J. *Nature* **1992**, *355*, 561.
7. Minke, W. E.; Hong, F.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Biol. Chem.* **1999**, *274* (47), 33469.
8. Minke, W. E.; Roach, C.; Hol, W. G. J.; Verlinde, C. L. M. J. *Biochemistry* **1999**, *38* (18), 5684.
9. Merritt, E. A.; Sarfaty, S.; Feil, I. K.; Hol, W. G. J. *Structure* **1997**, *5* (11), 1485.
10. Pickens, J. C.; Merritt, E. A.; Ahn, M.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *Chem. Biol.* **2002**, *9*, 215.
11. 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.
12. Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 8818.
13. Zhang, Z.; Merritt, E. A.; Ahn, M.; Roach, C.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 12991.
14. Fan, E.; Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Roach, C.; Ahn, M.; Hol, W. G. J. *Acta. Cryst.* **2001**, *D57*, 201.

15. Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. J. *J. Mol. Biol.* **1998**, *282*, 1043.
16. Otwinowski, Z.; Minow, W. *Methods Enzymol.* **1997**, *276*, 307.
17. Collaborative Computational Project 4 *Acta Cryst.* **1994**, *D50*, 760.
18. Vagin, A.; Teplyakov, A. *J. Appl. Cryst.* **1997**, *30*, 1022.
19. Minke, W. E.; Pickens, J.; Merritt, E. A.; Fan, E.; Verlinde, C. L. M. J.; Hol, W. G. J. *Acta Cryst.* **2000**, *D56*, 795.
20. Murshodov, G. N.; Vaguine, A. A.; Dodson, E. J. *Acta Cryst.* **1997**, *D53*, 240.
21. Perrakis, A.; Morris, R. J. V. S.L. *Nat. Struct. Biol.* **1999**, *6*, 458.
22. McRee, D. E. *J. Struct. Biol.* **1999**, *125*, 156.
23. Vaguine, A. A.; Richelle, J.; Wodak, S. J. *Acta Cryst.* **1999**, *D55*, 191.
24. Hooft, R. W. W.; Sander, C.; Vriend, G.; Abola, E. E. *Nature* **1996**, *381*, 272.
25. MOE: Chemical Computing Group, Inc., Montreal Canada.
26. Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38* (3), 305.
27. Merritt, E. A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505.