

A Protein-Based Pentavalent Inhibitor of the Cholera Toxin B-Subunit**

Thomas R. Branson, Tom E. McAllister, Jaime Garcia-Hartjes, Martin A. Fascione, James F. Ross, Stuart L. Warriner, Tom Wennekes, Han Zuilhof, and W. Bruce Turnbull*

Abstract: Protein toxins produced by bacteria are the cause of many life-threatening diarrheal diseases. Many of these toxins, including cholera toxin (CT), enter the cell by first binding to glycolipids in the cell membrane. Inhibiting these multivalent protein/carbohydrate interactions would prevent the toxin from entering cells and causing diarrhea. Here we demonstrate that the site-specific modification of a protein scaffold, which is perfectly matched in both size and valency to the target toxin, provides a convenient route to an effective multivalent inhibitor. The resulting pentavalent neoglycoprotein displays an inhibition potency (IC_{50}) of 104 pM for the CT B-subunit (CTB), which is the most potent pentavalent inhibitor for this target reported thus far. Complexation of the inhibitor and CTB resulted in a protein heterodimer. This inhibition strategy can potentially be applied to many multivalent receptors and also opens up new possibilities for protein assembly strategies.

Many viruses, bacteria, and protein toxins adhere to their target cells by binding to specific cell-surface carbohydrates. Diarrheal diseases that are initiated in this way account for around two million deaths annually.^[1] For example, cholera is caused by an AB₅ protein toxin that has a single toxic A-subunit associated with five nontoxic B-subunits (CTB), which constitute a pentameric receptor for the GM1 glyco-

lipid found on the surface of intestinal epithelial cells.^[2] Multivalent binding between CTB and up to five copies of its GM1 ligand facilitates the entry of the toxin into the cell by endocytosis.

Multivalency is a common feature of cell-surface adhesion and has thus provided the primary focus for creating inhibitors of carbohydrate-binding proteins, including CTB.^[3] Polymeric structures and dendrimers with either galactosyl residues or the more strongly binding GM1 oligosaccharide (GM1os) ligands have demonstrated the potential improvements in activity that multivalency can provide.^[4] Pentavalent star-shaped structures that match the valency and positioning of the ligand groups to CTB have been shown to be some of the most successful inhibitors. A pentacyclen core with galactose ligands was used by Fan and co-workers to produce an inhibitor with a reported IC_{50} value of 1.4 μ M.^[5] Their modular approach showed that matching precisely the size and spacing of the ligands to the binding sites of CTB could optimize the inhibitory potential.^[5b,6] Other recent studies have used GM1os ligands on both corannulene^[7] and calixarene cores.^[8] These pentavalent structures gave IC_{50} values down to 5 nM and 450 pM, respectively. A pentameric carbohydrate core has also been used in the highly effective Starfish ligands for shiga-like toxin.^[9] While synthetic pentavalent ligands have proven to be effective inhibitors, the synthesis of large (> 5 nm diameter), precisely defined multivalent scaffolds is very challenging and hampers larger-scale applications.

Here we report a multivalent inhibitor for cholera toxin, based on an inactive mutant CTB protein that is modified with GM1os ligands (Figure 1). As the majority of the structure comes as a ready-made building block, the synthetic route to the inhibitor is relatively short. Previous attempts to introduce sugars at defined sites on an ankyrin repeat protein^[10] and the barstar protein^[11] led to ligands and inhibitors of plant lectins. However, as no attempts were made to match the size and spacing of the ligands to the binding sites, these multivalent compounds led to only modest enhancements in activity and/or lectin aggregation, which may be undesirable. In the example reported here, the pentavalent inhibitor is matched to the size and valency of the target CTB protein, which leads to defined complexes and potent inhibition.

We postulated that by using an inactive mutant CTB protein as the scaffold, once modified with carbohydrates, the resulting pentavalent ligand should provide a precise fit of the ligand groups with the spacing and configuration of binding sites on wild-type CTB (Figure 1). The CTB homopentamer has five N-terminal threonine residues situated on the protein

[*] Dr. T. R. Branson, Dr. T. E. McAllister, Dr. M. A. Fascione, Dr. J. F. Ross, Dr. S. L. Warriner, Dr. W. B. Turnbull
School of Chemistry and Astbury Centre for Structural Molecular Biology, University of Leeds
Leeds, LS2 9JT (UK)
E-mail: w.b.turnbull@leeds.ac.uk

J. Garcia-Hartjes, Dr. T. Wennekes, Prof. H. Zuilhof
Laboratory of Organic Chemistry, Wageningen University
Dreijenplein 8, 6703 HB Wageningen (The Netherlands)
Prof. H. Zuilhof
Department of Chemical and Materials Engineering
King Abdulaziz University
Jeddah (Saudi-Arabia)

[**] We thank Dr. G. A. Posthuma-Trumpie for her assistance with inhibition assays. This work was supported by the EPSRC (EP/G043302/1), the Wellcome Trust (089308/Z/09/Z; 094232/Z/10/Z), the Royal Society (UF090025), the Interreg IV program, a Netherlands Organization for Scientific Research VENI-grant (722.011.006), and COST Action CM1102.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201404397>.

© 2014 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

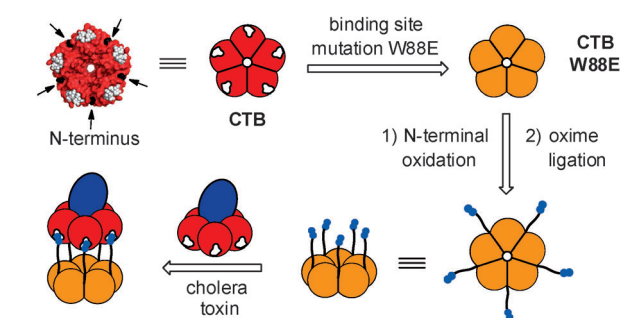
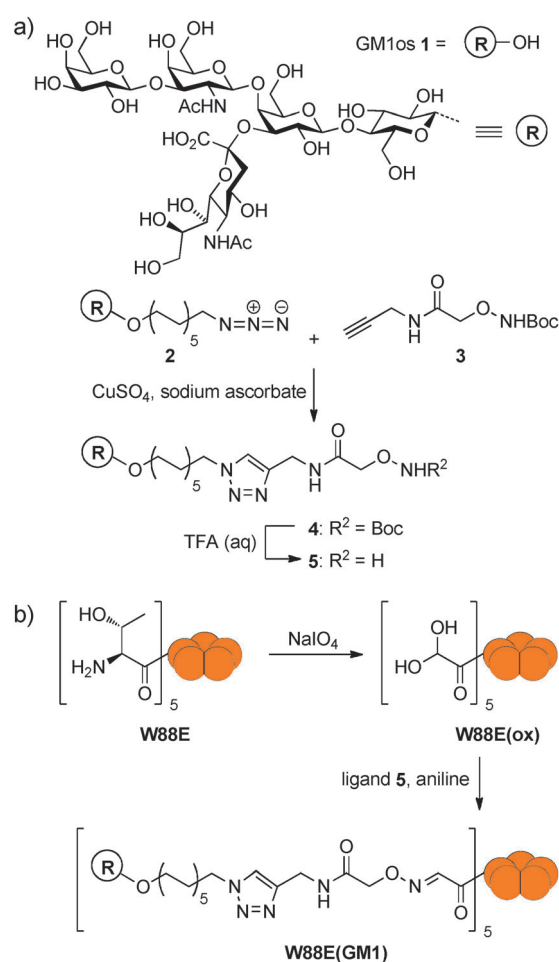


Figure 1. Strategy for re-engineering the CTB protein to prepare a pentavalent neoglycoprotein inhibitor for cholera toxin. The N termini of a nonbinding CTB mutant **W88E** are oxidized to give aldehydes that undergo oxime ligation with a carbohydrate ligand that bears an aminoxy function. The resulting neoglycoprotein has ligand groups arranged with optimal spacing to bind to the cholera toxin protein. In the 3D structure of CTB (PDB code: 3CHB),^[24] the binding sites are marked in white, and N-terminal threonine residues are marked with black arrows.

surface toward the ligand-binding face and equally spaced around the protein between the GM1 binding sites. These residues bear unique vicinal amino alcohol groups, which can be selectively converted by oxidation to an aldehyde before reaction with an oxyamine.^[12] Modification of the protein scaffold at these sites with five GM1os ligands would result in a pentavalent ligand. As proteins of this class can be prepared in large quantities, the synthesis is potentially scalable.^[13]

The GM1os ligand was prepared by first synthesizing Boc-protected aminoxy alkyne **3** (Scheme 1) according to a reported procedure.^[14] GM1 azide **2** was prepared using a chemo-enzymatic approach^[15] before ligation to aminoxy alkyne **3** employing a copper-catalyzed alkyne azide cycloaddition (CuAAC) at room temperature with stirring for 48 hours. Attempts were made to use microwave-assisted CuAAC for the synthesis,^[16] but proved unsuccessful in this case. After purification by reverse-phase column chromatography, the GM1os ligand **4** was isolated. The Boc group was removed with TFA to form oxyamine **5**, directly prior to the attachment to the protein without further purification (to preserve the highly reactive aminoxy group).

To create the protein scaffold, a nonbinding mutant of CTB was required. Based on the work of Jobling and Holmes,^[17] a tryptophan residue (W88) in the GM1 binding pocket was changed to a glutamic acid residue by site-directed mutagenesis. The resulting CTB mutant protein (**W88E**) was no longer capable of binding to GM1os (Figure S3 in Supporting Information). The five N-terminal vicinal amino alcohol groups in **W88E** were then oxidized with NaIO₄ (Scheme 1) to give aldehydes, which were observed in their hydrated form **W88E(ox)** by ESI-MS (found: 11 557.8 Da; calcd: 11 557.9 Da). The oxidized protein was then allowed to react with deprotected oxyamine **5** in the presence of aniline at pH 7. Aniline is known to be an effective catalyst of oxime formation at pH 4.5,^[18] however, as the CTB protein denatures under acidic conditions, a neutral pH value was used and the reaction still proceeded to completion within 24 hours. Pentavalent ligand **W88E(GM1)** was purified by size-exclu-



Scheme 1. a) Synthesis of an oxime-modified GM1 oligosaccharide ligand **5**; b) N-terminal oxidation of the CTB **W88E** protein and ligation with ligand **5**.

sion chromatography (SEC) and ESI-MS confirmed that the modified protein had a mass of 12 844.5 Da (calcd: 12 844.4 Da; Figure S1). SDS-PAGE confirmed that the modification did not disrupt the pentameric form of the protein and no protein aggregation was observed (Figure S1).

An enzyme-linked lectin assay (ELLA) was performed to determine the inhibitory potential of the neoglycoprotein.^[7,8] The ability of CTB to bind to a GM1-coated microtiter plate was assessed across a wide range of inhibitor concentrations. Pentavalent ligand **W88E(GM1)** exhibited an exceptionally low IC₅₀ value of 104 pM (Figure 2 and Table 1), making it the most potent pentavalent ligand reported thus far. This IC₅₀ value corresponds to a 5100-fold enhancement compared to monovalent GM1os **1**, or 14 300-fold enhancement compared to monovalent GM1 azide **2**. The unmodified **W88E** protein as a control showed no inhibition of CTB binding to a GM1-coated microtiter plate.

Isothermal titration calorimetry (ITC) was used to analyze the interaction between pentavalent protein-based inhibitor **W88E(GM1)** and wild-type CTB. **W88E(GM1)** was titrated into a solution of CTB and the expected binding stoichiometry of one GM1os ligand per binding site was

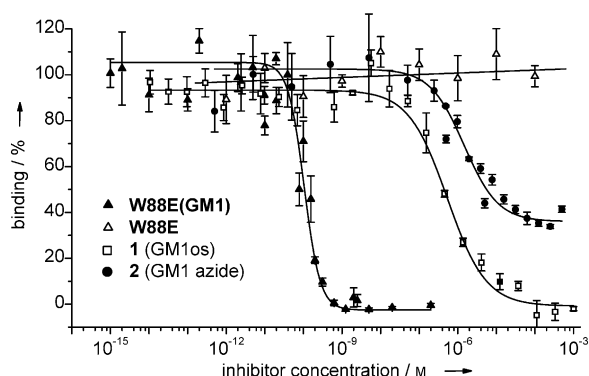


Figure 2. Enzyme-linked lectin assay (ELLA) indicates the inhibitory potential of **W88E(GM1)** and analogous monovalent ligands **1** and **2**. Error bars indicate the standard error of three measurements.

observed, which is consistent with a binding model in which **W88E(GM1)** forms a protein heterodimer complex with CTB. However, an apparent K_d of 30 nM was detected, similar to that of the monovalent GM1 azide **2**, which had a K_d of 56 nM (Figure S3 and Table S1), and that of GM1os **1** (43 nM), which was used in previous studies.^[20] The similarity between the affinities of mono- and pentavalent ligands contrasts with affinity enhancements reported for other multivalent systems.^[21] However, the result is in line with our previous studies of GM1-based dendrimers,^[22] in which it was concluded that a mismatch in valency between dimeric or tetrameric inhibitors and a pentavalent protein resulted in an aggregative mechanism of inhibition. In contrast, Fan and co-workers reported that their pentavalent ligands gave only 1:1 complexes in dynamic light scattering (DLS) and crystallization studies, and no larger aggregates were observed.^[23] Therefore, DLS and analytical ultracentrifugation (AUC) experiments were used to determine if 1:1 complexes or larger aggregates had formed between CTB and **W88E(GM1)**.

DLS was first used to determine the size of the assemblies (Figure 3a). The crystal structure of CTB shows the pentamer to have a diameter of 6.5 nm and a depth of 3.5 nm.^[24] These measurements are consistent with the experimental hydrodynamic diameter of 5.6 nm. At a 1:1 ratio of CTB and **W88E(GM1)**, DLS showed a single peak at 8.4 nm. A face-to-face dimer would have dimensions of 6.5 nm by at least 7 nm, which was in agreement with the result from DLS. Increasing the ratio of CTB to **W88E(GM1)** to 5:1 gave a similar result with no larger aggregates. The observations from the DLS experiment indicated that only a protein heterodimer of pentamers is formed.

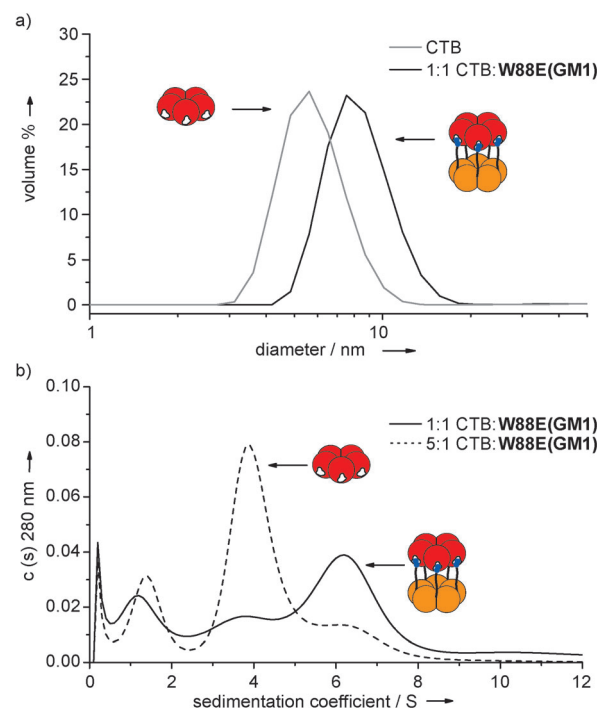


Figure 3. a) Dynamic light scattering (DLS) results showing the increase in the size of the particles when CTB is mixed at a ratio of 1:1 with **W88E(4)**. b) Sedimentation coefficient distributions [$c(s)$] of CTB mixed with **W88E(4)** at different ratios and the appearance of a single species corresponding to a protein heterodimer.

Sedimentation velocity AUC was used to further study the oligomerization state of the protein complexes being formed by ligand binding. At a ratio of 1:1 CTB and **W88E(GM1)**, AUC showed two peaks with sedimentation coefficients of 6.3 S and 3.6 S, corresponding to masses of approximately 110 kDa and 50 kDa, respectively (Figure 3b), which are consistent with the predicted masses of a dimer of protein pentamers and the single pentamer. The smaller peak at 3.6 S was likely due to the concentrations of the two components not being exactly equal. When the concentration of ligand **W88E(GM1)** was decreased, so that CTB was now in excess at a ratio of 5:1, the peak for the protein pentamer dominated, but the peak for a dimer of pentamers was still seen. This result showed that even with an excess of CTB, the only protein complex observed was a heterodimer of protein pentamers. No larger peaks were observed in the AUC experiment and there was negligible change in the protein absorbance after the AUC experiments, indicating that no protein was lost through the formation and pelleting of large

Table 1: Inhibitory potential results from the enzyme-linked lectin assay (ELLA).

Inhibitor	Valency	$\log(\text{IC}_{50})$	IC_{50} [nM] ^[a]	Relative potency (per GM1 group) ^[b]	Hill coefficient
2 (GM1 azide)	1	-5.83 ± 0.16	1460	1 (1)	1.0 ± 0.2
1 (GM1os)	1	-6.27 ± 0.04	530	2.75 (2.75)	0.9 ± 0.1
W88E(GM1)	5	-9.98 ± 0.08	0.104	14 300 (2860)	2.1 ± 0.5
W88E	0	— ^[c]	— ^[c]	— ^[c]	— ^[c]

[a] As curve fitting was performed with $\log[\text{inhibitor}]$ as x values, the fitting errors for IC_{50} values became asymmetric about the mean and were omitted for simplicity. [b] Potency was measured relative to monovalent GM1 azide **2**. [c] No inhibition detected.

aggregates.^[22] Nevertheless, it is still possible that at the inhibitor concentration used in these biophysical experiments (ca. 50 000 times higher than the IC₅₀ concentration for **W88E(GM1)** in the ELLA), random aggregates could initially form with modest binding affinities, as observed in our ITC experiments.^[25] If those aggregates were to rearrange more slowly to form the thermodynamically preferred 1:1 complexes observed by DLS and AUC,^[26] then this latter process would probably be undetectable during the ITC experiment, as the net enthalpy change for reorganization of the aggregates would probably be very small. It would be expected that 1:1 complexes should form directly at the much lower concentrations of **W88E(GM1)** inhibitor used in the ELLA.

SDS-PAGE performed on the AUC samples showed bands for the pentameric protein and a dimer of pentamers in the same ratios observed in the AUC experiments for the 1:1 and 5:1 mixtures of CTB and **W88E(GM1)** (Figure S1). Therefore, the interactions forming the protein heterodimers were sufficiently strong to survive the denaturing conditions of SDS-PAGE.

In conclusion, the combination of DLS, AUC, and SDS-PAGE confirmed that protein-based pentavalent ligand **W88E(GM1)** bound to CTB in a 1:1 ratio, forming protein heterodimers. With an IC₅₀ value of 104 pM, this inhibitor is the most potent pentavalent structure described thus far. By using a protein scaffold that matches the size and the spacing between binding sites and ligands, it is possible to control the structure of the complexes that are formed.^[3c] With the growing interest in ligand-mediated protein assembly, it is possible that tailored glycoproteins of this type could also be used for the preparation of nanostructured protein materials.^[27] Furthermore, the protein-based inhibitor reported here also has the advantage that it is relatively easy to synthesize. Proteins of this type can be produced on an industrial scale^[13] and the reaction to modify the protein is simple. While the synthesis of the carbohydrate moiety itself is not trivial,^[15] the combined use of protein modification and enzymatic oligosaccharide synthesis provides an attractive strategy for biopharmaceutical synthesis. This work demonstrates a general strategy for the creation of multivalent inhibitors of protein/carbohydrate interactions.

Received: April 16, 2014

Published online: July 2, 2014

Keywords: carbohydrates · glycoproteins · multivalency · protein modifications · protein structures

- [1] M. Kosek, C. Bern, R. L. Guerrant, *Bull. W. H. O.* **2003**, *81*, 197–204.
- [2] D. Vanden Broeck, C. Horvath, M. J. S. De Wolf, *Int. J. Biochem. Cell Biol.* **2007**, *39*, 1771–1775.
- [3] a) T. R. Branson, W. B. Turnbull, *Chem. Soc. Rev.* **2013**, *42*, 4613–4622; b) A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K.-E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penades, F. Peri, R. J. Pieters, O. Renaudet, J.-L. Reymond, B. Richichi, J.

- Rojo, F. Sansone, C. Schaffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof, A. Imberty, *Chem. Soc. Rev.* **2013**, *42*, 4709–4727; c) R. J. Pieters, *Org. Biomol. Chem.* **2009**, *7*, 2013–2025; d) V. Wittmann, R. J. Pieters, *Chem. Soc. Rev.* **2013**, *42*, 4492–4503.
- [4] a) B. D. Polizzotti, K. L. Kiick, *Biomacromolecules* **2005**, *7*, 483–490; b) S.-J. Richards, M. W. Jones, M. Hunaban, D. M. Haddleton, M. I. Gibson, *Angew. Chem.* **2012**, *124*, 7932–7936; *Angew. Chem. Int. Ed.* **2012**, *51*, 7812–7816; c) C. L. Schengrund, N. J. Ringler, *J. Biol. Chem.* **1989**, *264*, 13233–13237; d) A. V. Pukin, H. M. Branderhorst, C. Sisu, C. A. G. M. Weijers, M. Gilbert, R. M. J. Liskamp, G. M. Visser, H. Zuilhof, R. J. Pieters, *ChemBioChem* **2007**, *8*, 1500–1503.
- [5] a) Z. Zhang, J. C. Pickens, W. G. J. Hol, E. Fan, *Org. Lett.* **2004**, *6*, 1377–1380; b) E. Fan, Z. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, *J. Am. Chem. Soc.* **2000**, *122*, 2663–2664.
- [6] Z. S. Zhang, E. A. Merritt, M. Ahn, C. Roach, Z. Hou, C. Verlinde, W. G. J. Hol, E. Fan, *J. Am. Chem. Soc.* **2002**, *124*, 12991–12998.
- [7] M. Mattarella, J. Garcia-Hartjes, T. Wennekes, H. Zuilhof, J. S. Siegel, *Org. Biomol. Chem.* **2013**, *11*, 4333–4339.
- [8] J. Garcia-Hartjes, S. Bernardi, C. A. G. M. Weijers, T. Wennekes, M. Gilbert, F. Sansone, A. Casnati, H. Zuilhof, *Org. Biomol. Chem.* **2013**, *11*, 4340–4349.
- [9] P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Armstrong, H. Ling, N. S. Pannu, R. J. Read, D. R. Bundle, *Nature* **2000**, *403*, 669–672.
- [10] J. J. Hollenbeck, D. J. Danner, R. M. Landgren, T. K. Rainbolt, D. S. Roberts, *Biomacromolecules* **2012**, *13*, 1996–2002.
- [11] L. M. Artnier, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Denedde, N. Budisa, C. P. R. Hackenberger, *Chem. Commun.* **2012**, *48*, 522–524.
- [12] J. H. Chen, W. G. Zeng, R. Offord, K. Rose, *Bioconjugate Chem.* **2003**, *14*, 614–618.
- [13] M. Lebens, S. Johansson, J. Osek, M. Lindblad, J. Holmgren, *Nat. Biotechnol.* **1993**, *11*, 1574–1578.
- [14] E. Trévisiol, E. Defrancq, J. Lhomme, A. Laayoun, P. Cros, *Eur. J. Org. Chem.* **2000**, 211–217.
- [15] A. V. Pukin, C. A. G. M. Weijers, B. van Lagen, R. Wechselberger, B. Sun, M. Gilbert, M.-F. Karwaski, D. E. A. Florack, B. C. Jacobs, A. P. Tio-Gillen, A. van Belkum, H. P. Endtz, G. M. Visser, H. Zuilhof, *Carbohydr. Res.* **2008**, *343*, 636–650.
- [16] J. A. F. Joosten, N. T. H. Tholen, F. Ait El Maate, A. J. Brouwer, G. W. van Esse, D. T. S. Rijkers, R. M. J. Liskamp, R. J. Pieters, *Eur. J. Org. Chem.* **2005**, 3182–3185.
- [17] M. G. Jobling, R. K. Holmes, *Mol. Microbiol.* **1991**, *5*, 1755–1767.
- [18] a) A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem.* **2006**, *118*, 7743–7746; *Angew. Chem. Int. Ed.* **2006**, *45*, 7581–7584.
- [19] Attempts to use CuAAC to attach GM1 azide **2** to a CTB **W88E** protein bearing alkyne functions led to incomplete derivatization of the protein pentamers.
- [20] W. B. Turnbull, B. L. Precious, S. W. Homans, *J. Am. Chem. Soc.* **2004**, *126*, 1047–1054.
- [21] a) S. Cecioni, S. Faure, U. Darbost, I. Bonnamour, H. Parrot-Lopez, O. Roy, C. Taillefumier, M. Wimmerová, J.-P. Praly, A. Imberty, S. Vidal, *Chem. Eur. J.* **2011**, *17*, 2146–2159; b) T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387–429.
- [22] C. Sisu, A. J. Baron, H. M. Branderhorst, S. D. Connel, C. Weijers, R. de Vries, E. D. Hayes, A. V. Pukin, M. Gilbert, R. J. Pieters, H. Zuilhof, G. M. Visser, W. B. Turnbull, *ChemBioChem* **2009**, *10*, 329–337.
- [23] E. A. Merritt, Z. S. Zhang, J. C. Pickens, M. Ahn, W. G. J. Hol, E. K. Fan, *J. Am. Chem. Soc.* **2002**, *124*, 8818–8824.

- [24] E. A. Merritt, P. Kuhn, S. Sarfaty, J. L. Erbe, R. K. Holmes, W. G. J. Hol, *J. Mol. Biol.* **1998**, 282, 1043–1059.
- [25] Curiously, a reverse titration (i.e., CTB added to a solution of **W88E(GM1)**), gave the same K_d as before (32 nM) but a binding stoichiometry of 0.5 which is consistent with the occurrence of more complex processes under the conditions and timescale of the ITC experiment.
- [26] Whitesides and co-workers have reported that upon mixing a divalent antibody with a trivalent ligand, large aggregates are formed initially which rearrange over a period of 15–20 minutes to form discrete complexes: B. Bilgiçer, D. T. Moustakas, G. M. Whitesides, *J. Am. Chem. Soc.* **2007**, 129, 3722–3728.
- [27] K. Matsuura, *RSC Adv.* **2014**, 4, 2942–2953.
-