LETTERS

Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin

Emma Byres[1]*, Adrienne W. Paton[2]*, James C. Paton[2], Jonas C. Löfling[3], David F. Smith[4], Matthew C. J. Wilce[1], Ursula M. Talbot[2], Damien C. Chong[2], Hai Yu[5], Shengshu Huang[5], Xi Chen[5], Nissi M. Varki[3], Ajit Varki[3], Jamie Rossjohn[1] & Travis Beddoe[1]

AB_5 toxins comprise an A subunit that corrupts essential eukaryotic cell functions, and pentameric B subunits that direct target-cell uptake after binding surface glycans. Subtilase cytotoxin (SubAB) is an AB_5 toxin secreted by Shiga toxigenic *Escherichia coli* (STEC)[1], which causes serious gastrointestinal disease in humans[2]. SubAB causes haemolytic uraemic syndrome-like pathology in mice[3] through SubA-mediated cleavage of BiP/GRP78, an essential endoplasmic reticulum chaperone[4]. Here we show that SubB has a strong preference for glycans terminating in the sialic acid *N*-glycolylneuraminic acid (Neu5Gc), a monosaccharide not synthesized in humans. Structures of SubB–Neu5Gc complexes revealed the basis for this specificity, and mutagenesis of key SubB residues abrogated *in vitro* glycan recognition, cell binding and cytotoxicity. SubAB specificity for Neu5Gc was confirmed using mouse tissues with a human-like deficiency of Neu5Gc and human cell lines fed with Neu5Gc. Despite lack of Neu5Gc biosynthesis in humans, assimilation of dietary Neu5Gc creates high-affinity receptors on human gut epithelia and kidney vasculature. This, and the lack of Neu5Gc-containing body fluid competitors in humans, confers susceptibility to the gastrointestinal and systemic toxicities of SubAB. Ironically, foods rich in Neu5Gc are the most common source of STEC contamination. Thus a bacterial toxin's receptor is generated by metabolic incorporation of an exogenous factor derived from food.

The B subunits of AB_5 toxins typically recognize cognate glycan receptors displayed on cell-surface glycoconjugates[5,6]. Receptor specificity is critical for the pathogenic process, as it determines host susceptibility, tissue tropism, and the nature and spectrum of the resultant pathology. Accordingly, we sought to understand the receptor specificity of SubAB. Glycan array analysis showed that Oregon-Green-labelled SubAB (OG–SubAB) had a high degree of binding specificity for glycans terminating with α2-3-linked residues of the non-human sialic acid Neu5Gc (Table 1). Much weaker binding was seen with those glycans that terminated in α(2-3)-linked *N*-acetylneuraminic acid (Neu5Ac), which differs by one hydroxyl group from Neu5Gc (Fig. 1a). Table 1 is a list of glycans selected from the microarray analysis of SubAB toxin and a mutant derivative $SubAB_{A12}$ (discussed later). This list represents the glycans on the array to which native SubAB has the highest apparent affinity and corresponding Neu5Ac derivatives, asialo- and sulphated-derivatives. Of all the glycans on the array, Neu5Gcα2-3Galβ1-4-GlcNAcβ (#260) bound SubAB best. The binding of SubAB to this glycan is reduced 20-fold if the Neu5Gc is changed to Neu5Ac (#237); over 30-fold if the Neu5Gc linkage is changed from α2-3 to α2-6 (#263); and 100-fold if the sialic acid is removed (#152). The high binding of SubAB to structures #258, #260 and #261 indicates that it has a high affinity for terminal α2-3-linked Neu5Gc, with little discrimination for the penultimate moiety. Surface plasmon resonance analysis (Supplementary Fig. 2a) showed an approximately tenfold higher SubAB binding response to Neu5Gcα2-3Lacβ than to Neu5Ac-α(2-3)-Lac-β. Competitive inhibition studies (Supplementary Fig. 2b) indicated that Neu5Gcα2-3Lacβ has an inhibition constant (K_i) of 2 mM, which is in the range reported for other monovalent sialic-acid–protein interactions[7,8]. This high specificity of SubAB for Neu5Gc-terminating glycans is unique among bacterial toxins.

Next we determined the structure of the *apo*-form of the SubB pentamer (Fig. 1 and Supplementary Table 1). As expected the SubB protomer adopted the common oligonucleotide/oligosaccharide-binding fold[9], typical of other AB_5 toxins. The SubB structure most resembled the S2, S3 and S5 subunits of pertussis toxin (PTX; Supplementary Table 2), where the S2/3 subunits of PTX contained a shallow binding site for sialylated glycoproteins[10]; SubB also contained a similar shallow pocket lined by similar residues. In all such previously studied AB_5 toxins, the sialic acid in question is Neu5Ac, a common sialic acid found in humans and other animals. We then determined the structure of SubB in complex with free Neu5Gc (Supplementary Table 1). Neu5Gc bound to SubB unambiguously (Fig. 1c), whereas identical experiments using Neu5Ac failed to show any binding. The shallow binding pocket of SubB bound Neu5Gc in the chair conformation (Fig. 1d). The Neu5Gc lay in the α-anomeric configuration, even though most free Neu5Gc (more than 90%) in solution is in the β-anomeric form; SubB is therefore highly selective for the small fraction of α-anomeric Neu5Gc present in solution. Neu5Gc is coordinated mainly by multiple polar interactions with SubB (Supplementary Table 3). The predominant van der Waals interaction arose from Phe 11, which swivelled by 90° when Neu5Gc was bound, stacking parallel with the sugar ring and forming a cap over the binding site. In addition, the Neu5Gc was sequestered by many direct and water-mediated hydrogen bonds, including interactions with the side chains of Asp 8, Ser 12, Glu 36 and Tyr 78. Neu5Gc differs from Neu5Ac by the addition of a hydroxyl on the methyl group of the *N*-acetyl moiety of Neu5Ac. This extra hydroxyl present in Neu5Gc made crucial interactions with SubB: namely, the extra hydroxyl interacts with $Tyr\,78^{OH}$ and hydrogen bonds with the main chain of Met 10. These key interactions could not occur with Neu5Ac, thus explaining the marked preference for Neu5Gc.

[1]Protein Crystallography Unit and ARC Centre of Excellence for Structural and Functional Microbial Genomics, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. [2]School of Molecular and Biomedical Science, University of Adelaide, South Australia 5005, Australia. [3]Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California 92093-0687, USA. [4]Protein-Carbohydrate Interaction Core H, Emory University School of Medicine, Atlanta, Georgia 30322, USA. [5]Department of Chemistry, University of California, Davis, California 95616, USA.
*These authors contributed equally to this work.

Table 1 | Glycan array analysis of native SubAB and B subunit mutant $SubAB_{A12}$

Glycan number	Structure	SubAB		$SubAB_{A12}$	
		Mean relative fluorescence units	Coefficient of variation (%)	Mean relative fluorescence units	Coefficient of variation (%)
260	Neu5Gcα2-3Galβ1-4GlcNAcβ–Sp0*	37,606	4	590	10
261	Neu5Gcα2-3Galβ1-4Glcβ–Sp0	25,197	16	1,998	30
264	Neu5Gcα–Sp8*	25,114	19	172	87
258	Neu5Gcα2-3Galβ1-3GlcNAcβ–Sp0	24,209	12	327	54
15	α-Neu5Ac–Sp11	8,836	30	1,242	46
31	[3OSO3]Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	3,094	10	1,375	41
33	[3OSO3]Galβ1-3GlcNAcβ–Sp8	3,006	15	3,808	45
231	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	2,401	8	1,587	72
46	Neu5Acα2-3[6OSO3]Galβ1-4GlcNAcβ–Sp8	2,254	16	2,997	11
36	[3OSO3]Galβ1-4GlcNAcβ–Sp0	1,923	22	3,426	4
34	[3OSO3]Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	1,878	36	2,398	32
237	Neu5Acα2-3Galβ1-4GlcNAcβ–Sp8	1,862	31	4,247	8
240	Neu5Acα2-3Galβ1-4Glcβ–Sp8	1,799	10	1,127	25
243	Neu5Acα2-6GalNAcβ1-4GlcNAcβ–Sp0	1,730	18	1,871	67
117	Galβ1-3(Fucα1-4)GlcNAc–Sp0	1,549	42	1,395	55
68	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	1,385	30	1,204	53
225	Neu5Acα2-3Galβ1-3GlcNAcβ–Sp0	1,315	75	1,563	16
257	Neu5Gcα2-3Galβ1-3(Fucα1-4)GlcNAcβ–Sp0	1,244	36	2,250	14
259	Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	1,163	37	1,719	11
263	Neu5Gcα2-6Galβ1-4GlcNAcβ–Sp0	1,154	34	859	16
245	Neu5Acα2-6Galβ1-4GlcNAcβ–Sp0	1,048	34	1,989	38
262	Neu5Gcα2-6GalNAcα–Sp0	1,003	10	125	23
37	[3OSO3]Galβ1-4GlcNAcβ–Sp8	989	21	1,362	63
217	Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	922	14	335	42
10	GalNAcα–Sp8	857	39	938	43
230	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	798	26	344	24
134	Galβ1-3GlcNAcβ–Sp8	678	121	197	33
246	Neu5Acα2-6Galβ1-4GlcNAcβ–Sp8	658	30	755	26
154	Galβ1-4Glcβ–Sp0	644	41	789	48
118	Galβ1-3(Fucα1-4)GlcNAc–Sp8	604	31	411	59
242	Neu5Acα2-6GalNAcα–Sp8	564	16	434	24
133	Galβ1-3GlcNAcβ–Sp0	539	25	604	24
239	Neu5Acα2-3Galβ1-4Glcβ–Sp0	484	46	318	23
49	9-O-AcNeu5Acα2-6Galβ1-4GlcNAcβ–Sp8	478	42	204	59
4	Neu5Gcβ2-6Galβ1-4GlcNAc–Sp8	471	24	176	45
92	GalNAcβ1-4GlcNAcβ–Sp0	419	15	235	24
152	Galβ1-4GlcNAcβ–Sp0	374	40	755	27
67	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	326	52	273	103
155	Galβ1-4Glcβ–Sp8	299	32	164	73
153	Galβ1-4GlcNAcβ–Sp8	298	20	123	44

Data are presented for a selection of 40 of the 320 glycans present on the array. Complete data sets are available at: www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp. Data are for quadruplicate array spots.
* Sp0 and Sp8 designate $CH_2CH_2NH_2$ and $CH_2CH_2CH_2NH_2$ linkers, respectively.

Given the glycan array data, we also established how a trisaccharide that terminated in Neu5Gc (Neu5Gcα2-3Galβ1-3GlcNAc; #258 in the array) bound SubB (Fig. 1e). The mode of binding of the Neu5Gc moiety in the monosaccharide and the trisaccharide complex was identical. The remaining two sugar moieties present in the trisaccharide extended to solvent, but were nevertheless able to contact SubB. The small number of additional interactions between SubB and the tertiary sugar in the trisaccharide (Fig. 1e) is consistent with the Neu5Gc moiety driving the specificity and affinity for the interaction with SubB. Of note, the sequence Neu5Gcα2-3Galβ1-3GlcNAc is very common as a terminating structure of *N*-glycans in non-human animals.

Despite the commonality of oligonucleotide/oligosaccharide-binding fold, neither the receptor specificity nor the location of the receptor-binding site is conserved throughout the AB_5 toxin family. For Shiga and cholera toxins, whose receptors are glycolipids, the deep receptor-binding pockets are located on the membrane face of the toxin[11–13]. Like SubB[14], PTX binds sialylated glycoproteins; moreover, the PTX S2/3 sialic-acid binding site is also shallow and in the same location, halfway down the sides of the pentamer (Fig. 1f–h). Examination of the SubB–Neu5Gc structure superposed onto the structure of PTX in complex with the disaccharide, Neu5Acα2-3Gal, reveals that both toxins bind sialic acid in the same orientation, with similar residues interacting with the sialic-acid head group (Fig. 1d, f). However, the interactions between Tyr 78 and Asp 8 of SubB with the extra hydroxyl group of Neu5Gc have no such equivalents in PTX. At the same position as Tyr 78 in SubB is the small, non-polar Val 167 side chain in PTX S2/3, and there are no residues that overlay in the PTX S2/3 region equivalent to Asp 8 in SubB. The absence of these side chains in PTX S2/3 provides a basis for understanding why PTX, unlike SubB, is not specific for Neu5Gc.

SubB residues that we considered critical for Neu5Gc binding were then mutated to determine their role in biologically relevant toxin–receptor interactions. The most critical residue was Ser 12, mutation of which abolishes interactions with the C1 carboxylate group of sialic acid; this reduced cytotoxicity for Vero cells by 99.98% (Supplementary Table 4). In addition, binding of labelled mutated toxin (OG–$SubAB_{A12}$) to Vero cells could not be detected by fluorescence microscopy, confirming an effect on receptor recognition (Fig. 2a). Further glycan array analysis using OG–$SubAB_{A12}$ indicated markedly reduced binding to the immobilized sialylated glycans (Table 1). The elimination of the hydroxyl of Tyr 78 in the $SubAB_{F78}$ mutant holotoxin also reduced Vero cell cytotoxicity by 96.9% (Supplementary Table 4), which highlights the importance of interactions with the hydroxyl group unique to Neu5Gc-containing glycans. The trace residual cytotoxicity of the $SubAB_{F78}$ mutant is probably attributable to SubB still being able to bind weakly to Neu5Ac-containing glycans, because Tyr 78 would not be required for binding Neu5Ac. Labelled $SubAB_{F78}$ also exhibited reduced binding to Vero cells (Fig. 2a). The Gln36Ala mutation resulted in

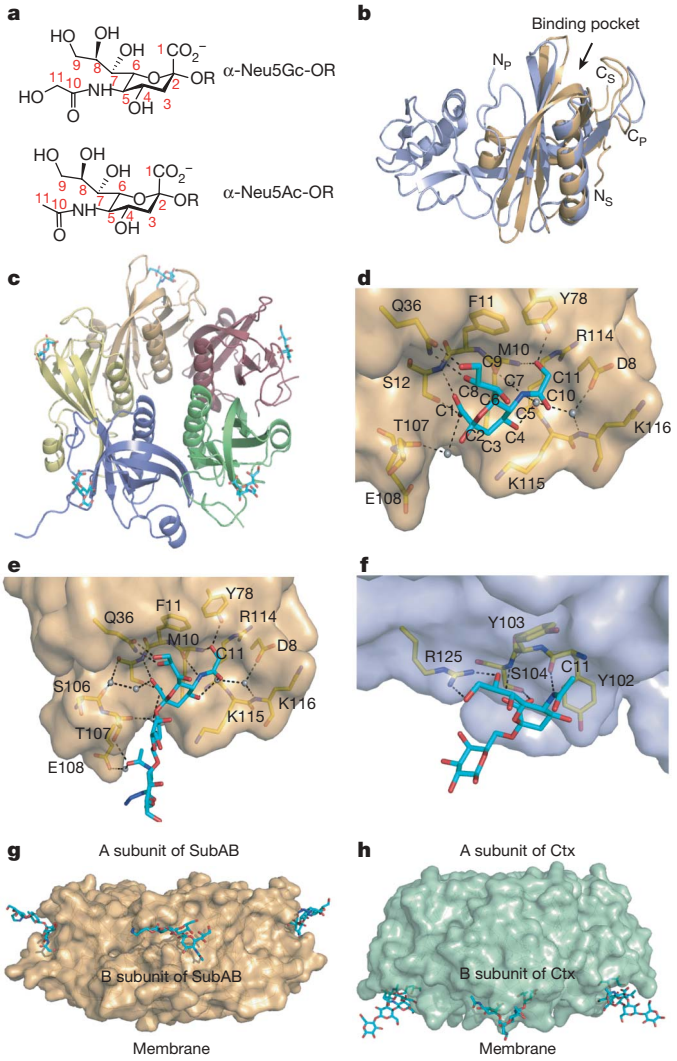

Figure 1 | Structural analysis of SubB–sialic acid interactions and comparison with other AB_5 toxins. **a**, Structures of Neu5Gc and Neu5Ac, showing additional O at C11 of the former. **b**, SubB protomer (orange) superimposed upon PTX S2 (blue). Amino (N) and carboxy (C) termini of SubB and PTX S2 are designated by subscripts S and P, respectively. **c**, Representation of the pentameric SubB–Neu5Gc structure, with each protomer colour-coded. Cyan sticks represent the sugar, blue sticks represent nitrogen atoms, and red sticks represent oxygen atoms. **d**, Neu5Gc in sialic-acid receptor-binding site of SubB. The extra hydroxyl of Neu5Gc interacts with $Tyr78^{OH}$, and hydrogen bonds with the main chain of Met 10. **e**, Trisaccharide Neu5Gcα2-3Galβ1-3GlcNAcβProN$_3$ binding to SubB. ProN3 refers to the linker used in the synthesis. **f**, Neu5Acα2-3Gal binding site of PTX S2/3, which shares similarity to the SubB binding site, namely: Ser 12 in SubB, Ser 104 in PTX S2/3; Gln 36 in SubB, Arg 125 in PTX S2/3; Phe 11in SubB, Tyr 103 in PTX S2/3. PTX does not have the equivalent of Tyr 78 and Asp 8. **g**, Side-on view of SubB:Neu5Gcα2-3Galβ1GlcNAcβProN$_3$. **h**, Side-on view of CtxB:GM_1. In **d**–**h**, cyan sticks represent ligands, dark blue sticks represent nitrogen atoms, and red sticks oxygen atoms. Yellow sticks represent key residues in the protein backbone. Black dotted lines represent hydrogen bonds.

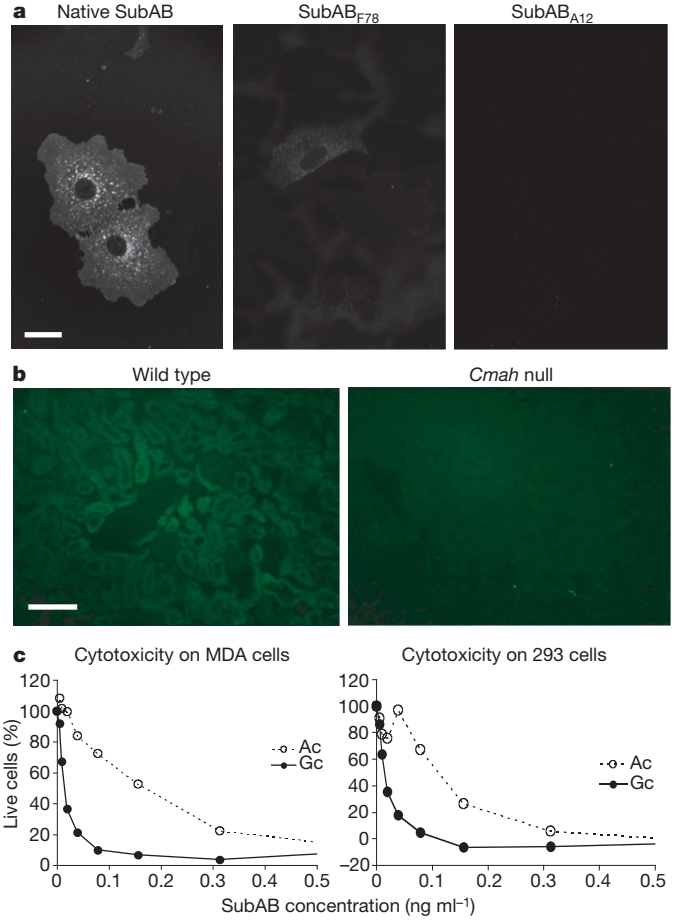

Figure 2 | Fluorescence microscopy and Neu5Gc-dependent cytotoxicity. **a**, Binding of SubAB mutants to Vero cells. Vero cells growing on coverslips were incubated with 1 μg ml^{-1} OG–SubAB, OG–SubAB$_{A12}$ or Texas Red-SubAB$_{F78}$ for 60 min at 37 °C (scale bar, 25 μm). **b**, Binding of SubAB to kidney sections from wild-type and *Cmah*-null mice. Frozen kidney sections were incubated with 1 μg ml^{-1} OG–SubAB for 60 min at 37 °C, washed, and examined by epifluorescence microscopy (scale bar, 100 μm). **c**, Neu5Gc-dependent cytotoxicity. Cytotoxicity of SubAB for MDA-MB-231 cells or 293 cells after growth in medium supplemented with 3 mM Neu5Gc (Gc) or Neu5Ac (Ac) was determined as described in the Methods. Data shown are from a single experiment. Analysis of data pooled from triplicate experiments yielded CD_{50} values of 3.92 ± 1.58 pg (mean ± standard error) and 13.17 ± 2.46 pg for MDA cells fed Neu5Gc or Neu5Ac ($P = 0.034$; two-tailed *t*-test). For 293 cells fed Neu5Gc or Neu5Ac the CD_{50} values are 1.33 ± 0.30 pg and 11.58 ± 3.30 pg, respectively ($P = 0.036$). The mean increase in SubAB susceptibility (± standard error) for cells fed Neu5Gc versus Neu5Ac was 5.56 ± 2.71-fold for MDA cells and 8.35 ± 0.88-fold for 293 cells.

reduction of cytotoxicity by 88% (Supplementary Table 4), which is consistent with the importance of the interactions with the C8 and C9 hydroxyl groups present on both Neu5Gc and Neu5Ac.

As further evidence of the biological significance of the high specificity of SubAB for Neu5Gc-terminating glycans, we showed that OG–SubAB bound to kidney tissue from wild-type mice, but not cytidine monophosphate (CMP)-*N*-acetylneuraminic acid hydroxylase (*Cmah*)-null mice[15] that had a human-like genetic defect in the ability to convert CMP-Neu5Ac to CMP-Neu5Gc (Fig. 2b). Humans lack this enzyme owing to a mutation in the *Cmah* gene that occurred after evolutionary separation of the hominin lineage from the great apes[16], suggesting the possibility of human genetic resistance to the toxin. However, human cells can metabolically assimilate Neu5Gc present in tissue culture media and incorporate it into cell-surface glycans[17]. We therefore manipulated the levels of Neu5Gc on the surface of human cell lines (MDA-MB-231 breast cancer cells and 293 embryonic kidney cells), which had been adapted for growth in human serum, resulting in undetectable amounts of Neu5Gc. These cells were then grown for 3 days in medium supplemented with 3 mM Neu5Gc or 3 mM Neu5Ac. In the former case, this increased Neu5Gc content to 50–75% of total sialic acid in membrane preparations (result not shown). Feeding MDA and 293 cells with Neu5Gc rather than Neu5Ac significantly increased their susceptibility to SubAB (50% cytotoxic dose (CD_{50}) values decreased 5.56- and

8.35-fold, respectively), rendering them as sensitive as Vero cells (Fig 2c).

We have previously shown that some normal human tissues contain small quantities of Neu5Gc[17]. No tissues in *Cmah*-null mice express Neu5Gc, as determined by a highly specific polyclonal chicken antibody, and by mass spectrometry[15]. Thus, human Neu5Gc must be derived from dietary sources, a mechanism confirmed by previous studies of human volunteers[17]. We therefore studied human colon sections by using an IgY antibody with absolute specificity for Neu5Gc (unpublished improvements over ref. 15), and noted that Neu5Gc was present on the epithelial surfaces and in the crypts (Fig. 3a). Similar regions of human colon sections also bound native SubAB, but binding of $SubAB_{A12}$ was markedly diminished (Fig. 3b). Importantly, pre-incubation with anti-Neu5Gc IgY, but not control IgY, substantially blocked binding of SubAB to human colon sections (Fig. 3c), demonstrating competition for the same epitope. Specific binding of SubAB to human kidney sections, particularly the glomerular endothelium, was also observed, but this was not seen for $SubAB_{A12}$ (Fig. 3d). Human kidney vasculature also shows staining with the anti-Neu5Gc antibody (data not shown). Furthermore, binding to kidney tissue was not seen with $SubAB_{F78}$, which is defective only in Neu5Gc-specific receptor interactions (Fig. 3e). Thus, binding of SubAB to human colon and kidney tissue is Neu5Gc dependent.

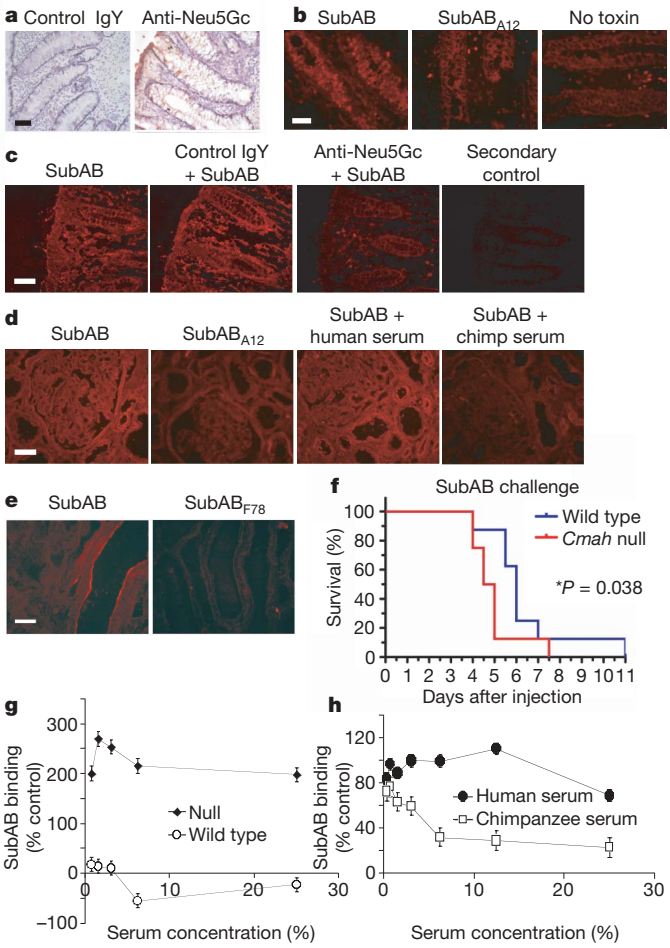

Figure 3 | Neu5Gc-dependent binding of SubAB to human tissues and toxicity of SubAB in wild-type and *Cmah*-null mice. **a**, Frozen sections of human colon were stained with chicken anti-Neu5Gc or control IgY at 5 μg ml^{-1}, followed by anti-chicken IgY–HRP conjugate, and examined by immunohistochemistry[17] (scale bar, 100 μm). **b**, Similar human colon sections were overlaid with or without 1 μg ml^{-1} SubAB or $SubAB_{A12}$ and bound toxin was detected using rabbit anti-SubA and Cy3-labelled goat anti-rabbit IgG and examined by epifluorescence microscopy (see Methods) (scale bar, 50 μm). **c**, Human colon sections were overlaid first with anti-Neu5Gc or control IgY at 5 μg ml^{-1}, followed by 1 μg ml^{-1} SubAB, and bound toxin was detected as for **b**. Background control sections received only rabbit anti-SubAB, followed by Cy3-labelled anti-rabbit IgG (scale bar, 100 μm). **d**, **e**, Human kidney sections were overlaid 1 μg ml^{-1} SubAB, $SubAB_{A12}$ or $SubAB_{F78}$, and in the presence or absence of 10% human or chimpanzee serum, as indicated. Bound toxin was detected as for **b** (scale bar, 50 μm). **f**, Wild-type and *Cmah*-null mice ($n = 8$ each) were injected intraperitoneally with 200 ng g^{-1} purified SubAB in a total of 100 μl PBS, and survival time was recorded. Kaplan–Meier survival curves were plotted and statistical analysis (Wilcoxon–Gehan test) was performed in GraphPad Prism. **g**, **h**, Inhibition of SubAB binding to immobilized Neu5Gc-α(2-3)-Lac-HSA by wild-type versus *Cmah*-null mouse serum (**f**), or human versus chimpanzee serum (**g**) was assayed by enzyme-linked immunosorbent assay (ELISA) as described in the Methods. Data are expressed as the percentage of a control with no serum and are the mean ± standard error of triplicate wells and representative of three independent experiments.

We also examined whether the inability of *Cmah*-null mice to express Neu5Gc affected *in vivo* susceptibility to injected SubAB. Surprisingly, the null mice had a slightly shorter median survival time (5 versus 6 days, $P = 0.038$; Fig. 3f). We reasoned that because normal mouse serum contains high levels of Neu5Gc-containing glycoproteins, these would compete with receptors on the surface of tissues for SubAB, thereby providing partial protection against toxicity. In contrast, human and *Cmah*-null mouse serum would not have any such protective activity. Thus, in the null mouse, lack of protective serum glycoproteins could counteract any benefit derived from lack of expression of Neu5Gc on cell-surface glycoconjugates. Indeed, wild-type but not *Cmah*-null mouse serum competitively inhibited binding of SubAB to immobilized Neu5Gc-glycan (Fig. 3g). Similar results were obtained for chimpanzee serum (which, like wild-type mouse serum, is rich in Neu5Gc-glycoproteins) versus human serum (Fig. 3h). Furthermore, chimpanzee serum, but not human serum, inhibited binding of SubAB to human kidney sections (Fig. 3d). Given that chimpanzee serum glycoproteins otherwise share near identity with their human counterparts, the effects can be attributed to the presence or absence of Neu5Gc competitors[18].

This study presents the first example of a bacterial toxin showing a marked preference for Neu5Gc-containing glycans and is potentially significant in the context of host susceptibility to toxin-mediated disease. Indeed, glycans terminating in Neu5Gcα2-3Galβ1-3GlcNAc sequences are widely expressed on the cells of many mammals including livestock, suggesting an evolutionary reason for the emergence of this selective binding preference. Neu5Gc is not produced by bacteria or plants, is low or absent in poultry and fish, but abundant in red meats (lamb, pork and beef) and in bovine milk. However, humans are the known exception among mammals, and Neu5Ac predominates because of the *Cmah* mutation[16]. How then could the toxin mediate disease in humans? The answer probably lies in the diet, as human ingestion of red meats and milk products results in Neu5Gc incorporation into human tissues[17]. We detected significant levels of Neu5Gc on human colonic epithelium, and Neu5Gc-specific binding of SubAB at this site, as well as in the kidney vasculature, the major target organ for haemolytic–uraemic syndrome caused by STEC (Fig. 3). This Neu5Gc can only have originated from the diet, as there is no known alternative pathway for Neu5Gc biosynthesis[15]. Ironically, red meat and dairy products (the richest dietary sources of Neu5Gc) are the very foods that are most commonly contaminated with SubAB-producing STEC[2]. Thus, through regular dietary intake of red meats and milk, humans may pre-sensitize their tissues to a key virulence factor of a major pathogen that occurs sporadically in the same foods. Furthermore, because of the absence of protective Neu5Gc-bearing glycoproteins in their serum and other body fluids, humans who have consumed foods with high Neu5Gc content may actually be hyper-susceptible to the toxin, as illustrated in Supplementary Fig. 1.

METHODS SUMMARY

Purification and labelling of SubAB. SubAB holotoxin, and derivatives with B-subunit mutations, were purified by Ni-NTA chromatography and labelled with Oregon Green, as described previously[19,20].

Structure determination. SubB was crystallized, and the Se-Met-labelled SubB structure was solved using the multiple anomalous dispersion technique. Crystal soaks and difference Fourier analysis was used to locate the Neu5Gc-containing binding sites. Further details are provided in the Methods.

Glycan array analysis. Binding of OG-labelled SubAB and mutant derivatives thereof to immobilized glycans was investigated using a printed array of 320 glycan targets on version 3.0 of the glycan microarray of the Consortium for Functional Glycomics Core H (http://www.functionalglycomics.org/static/consortium/resources/resourcecoreh8.shtml). Further details are provided in the Methods.

DNA methods. Routine DNA manipulations were performed essentially as described previously[1]. DNA sequencing used dye-terminator chemistry and an ABI 3700 sequencer.

Cell culture and cytotoxicity assays. SubAB and mutant derivatives were assayed for cytotoxicity on Vero cells as previously described[1]. Cytotoxicity was also assayed on MDA-MB-231 (human breast cancer) and 293 (human embryonic kidney) cells, which had been adapted for growth in human serum to eliminate presence of Neu5Gc. Cells were then cultured for 3 days in 96-well plates in medium (RPMI and DMEM, respectively) supplemented with 3 mM Neu5Gc or Neu5Ac, and then exposed to serial dilutions of SubAB, as described for Vero cells[1]. Cell viability was assessed after incubation for a further 3 days by staining with 10% Alamar blue (Serotec), according to the manufacturer's instructions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 June; accepted 15 September 2008.
Published online 29 October 2008.

1. Paton, A. W., Srimanote, P., Talbot, U. M., Wang, H. & Paton, J. C. A new family of potent AB_5 cytotoxins produced by Shiga toxigenic *Escherichia coli*. *J. Exp. Med.* **200,** 35–46 (2004).
2. Paton, J. C. & Paton, A. W. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Clin. Microbiol. Rev.* **11,** 450–479 (1998).
3. Wang, H., Paton, J. C. & Paton, A. W. Pathologic changes in mice induced by subtilase cytotoxin, a potent new *Escherichia coli* AB_5 toxin that targets the endoplasmic reticulum. *J. Infect. Dis.* **196,** 1093–1101 (2007).
4. Paton, A. W. & Beddoe, T. Thorpe, C.M., Whisstock, J.C., Wilce, M.C.J., Rossjohn, J., Talbot, U.M. and Paton J.C. AB_5 subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. *Nature* **443,** 548–552 (2006).
5. Sandvig, K. & van Deurs, B. Membrane traffic exploited by protein toxins. *Annu. Rev. Cell Dev. Biol.* **18,** 1–24 (2002).
6. Lencer, W. I. & Tsai, B. The intracellular voyage of cholera toxin: going retro. *Trends Biochem. Sci.* **28,** 639–645 (2003).
7. Crocker, P. R., Paulson, J. C. & Varki, A. Siglecs and their roles in the immune system. *Nature Rev. Immunol.* **7,** 55–66 (2007).
8. Neu, U., Woellner, K., Gauglitz, G. & Stehle, T. Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc. Natl Acad. Sci. USA* **105,** 5219–5224 (2008).
9. Murzin, A. G. OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12,** 861–867 (1993).
10. Stein, P. E. *et al.* Structure of a pertussis toxin-sugar complex as a model for receptor binding. *Nature Struct. Biol.* **1,** 591–596 (1994).
11. Merritt, E. A., Sixma, T. K., Kalk, K. H., van Zanten, B. A. & Hol, W. G. Galactose-binding site in *Escherichia coli* heat-labile enterotoxin (LT) and cholera toxin (CT). *Mol. Microbiol.* **13,** 745–753 (1994).
12. Merritt, E. A. *et al.* Structural studies of receptor binding by cholera toxin mutants. *Protein Sci.* **6,** 1516–1528 (1997).
13. Ling, H., Bast, D. & Brunton, J. L. &. R. e. a. d. R. J. Structure of the Shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37,** 1777–1788 (1998).
14. Yahiro, K. *et al.* Identification and characterization of receptors for vacuolating activity of subtilase cytotoxin. *Mol. Microbiol.* **62,** 480–490 (2006).
15. Hedlund, M. *et al.* N-glycolylneuraminic acid deficiency in mice: implications for human biology and evolution. *Mol. Cell. Biol.* **27,** 4340–4346 (2007).
16. Varki, A. Multiple changes in sialic acid biology during human evolution. *Glycoconjugate J.* doi:10.1007/s10719-008-9183-z (7 September 2008).
17. Tangvoranuntakul, P. *et al.* Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl Acad. Sci. USA* **100,** 12045–12050 (2003).
18. Gagneux, P. *et al.* Proteomic comparison of human and great ape blood plasma reveals conserved glycosylation and differences in thyroid hormone metabolism. *Am. J. Phys. Anthropol.* **115,** 99–109 (2001).
19. Talbot, U. M., Paton, J. C. & Paton, A. W. Protective immunization of mice with an active-site mutant of subtilase cytotoxin of Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **73,** 4432–4436 (2005).
20. Chong, D. C., Paton, J. C., Thorpe, C. M. & Paton, A. W. Clathrin-dependent trafficking of subtilase cytotoxin, a novel AB_5 toxin that targets the ER chaperone BiP. *Cell. Microbiol.* **10,** 795–806 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff at the General Medicine and Cancer Institutes Collaborative Access Team Advanced Photon Source, Chicago, for assistance with data collection. This research was supported by a Program Grant from the National Health and Medical Research Council of Australia (NHMRC; to A.W.P. and J.C.P.), an NHMRC Project Grant (to T.B. and A.W.P.), a grant from the National Institute of General Medical Sciences to the Consortium for Functional Glycomics, RO1 grants from the National Institutes of Health (to A.W.P., J.C.P., J.R., X.C. and A.V.) and from the ARC Centre of Excellence in Structural and Functional Microbial Genomics (to J.R.). J.R. is supported by an Australian Research Council Federation Fellowship; T.B. by an NHMRC Career Development Award; J.C.P. by an NHMRC Australia Fellowship. We also thank C. J. Gregg for assistance with *in vivo* experiments and with collection of data, and L. Wiggleton for technical assistance with tissue sectioning and staining.

Author Contributions E.B. and A.W.P. contributed equally. E.B, T.B. and M.C.J.W. crystallized SubB, solved the structure and contributed to manuscript preparation. A.W.P. constructed mutants and contributed to design and interpretation of experiments, project management and writing of the manuscript. J.C.P. and J.R. contributed to design and interpretation of experiments, project management and writing of the manuscript. D.C.C. and U.M.T. performed experiments. D.F.S. performed and interpreted glycan array experiments. J.C.L., N.M.V. and A.V. designed, performed and interpreted experiments relating to Neu5Gc on cells and tissues and to cytotoxicity *in vivo*, and contributed to manuscript preparation. H.Y., S.H. and X.C. synthesized oligosaccharides. A.V., T.B. and J.R. are joint senior and corresponding authors.

Author Information The coordinates and structure factors for the SubB structures are deposited in Protein Data Bank under accession numbers 3DWA, 3DWP and 3DWQ. Raw glycan array data are available at www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.V. (a1varki@ucsd.edu), J.R. (jamie.rossjohn@med.monash.edu.au) or T.B. (travis.beddoe@med.monash.edu.au).

METHODS

Glycan array analysis. Labelled proteins were diluted to 0.1 mg ml^{-1} in Tris-buffered saline (TBS: 20 mM Tris, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, pH 7.4) containing 1% BSA and 0.05% Tween-20, and an aliquot (70 µl) was applied to separate microarray slides and incubated under a coverslip for 60 min at room temperature. Coverslips were then removed and slides were washed by dipping four times in successive washes of TBS containing 0.05% Tween-20, TBS and deionized water. Slides were then spun for approximately 15 s to dry and immediately scanned in a PerkinElmer ProScanArray MicroArray Scanner using an excitation wavelength of 488 nm and ImaGene software (BioDiscovery) to quantify fluorescence. The data are reported as average relative fluorescence units of four of six replicates (after removal of the highest and lowest values) for each glycan represented on the array.

Purification and crystallization of SubB. Purification of SubB was performed as detailed previously for the holotoxin[4] with the following modifications. SubB was eluted from an Ni-NTA column using wash buffer containing 500 mM imidazole and further purified by gel filtration using an S200 16/60 column (GE Healthcare) pre-equilibrated in 20 mM Tris HCl, 150 mM NaCl, 2 mM EDTA, pH 7.0. Fractions containing SubB were pooled and purity was assessed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometry (data not shown). SubB was concentrated to 2 mg ml^{-1} and 40 mM 3-(decyldimethylammonio) propanesulphonate was added before storing at 4 °C. Crystals of SubB grew in drops containing 1 µl protein (2 mg ml^{-1}) and 1 µl reservoir solution, with the reservoir solution consisting of 500 µl of 16% (w/v) PEG 3350, 100 mM sodium cacodylate, pH 6.2, and 200 mM ammonium fluoride. Selenomethionine-substituted SubB was purified and crystallized under the same conditions.

X-ray data collection, structure determination and refinement. X-ray data were collected for SeMet-labelled SubB at the GMCA-CAT beamline at the Advanced Photon Source, Chicago. Data were collected from a single crystal at three wavelengths: inflection point, peak and high energy remote. Native SubB crystals were soaked by the addition of 1 mM of Neu5Gc, Neu5Ac or the trisaccharide Neu5Gcα2-3Galβ1-3GlcNAcβProN_3 for 1 h at 20 °C. Datasets from the soaked crystals were collected in-house using an R-axis IV^{++} detector. X-ray data were autoindexed and processed with MOSFLM[21], SCALA[22] and the CCP4 suite of programs[23].

The *apo*-form of SubB was phased by the single-wavelength anomalous dispersion (SAD) method using SeMet inflection point data and the PHENIX Autosol suite[24]. An initial model of 460 out of 630 residues was built by RESOLVE[25], and this was used as input for ARP/wARP[26]. The model was completed by rounds of manual building in COOT[27] followed by refinement in REFMAC5 (ref. 28) with fivefold non-crystallographic symmetry maintained until the final stages of refinement. Waters and PEG 400 molecules were added in COOT to give a final SubB model consisting of a homopentamer of 588 residues (with each B subunit between 115 and 120 residues). The apo-SubB structure was used, through PHASER[29], to solve the sugar-bound structures, with residues rebuilt and refined as described for the *apo* structure. Ligands were constructed using the PRODRG2 server[30]. Clear density was observed for the various sugars in the SubB binding sites and these were further confirmed by constructing F_o–F_c omit maps. Data collection and refinement statistics are shown in Supplementary Table 1. The stereochemistry and overall quality of each of the structures was confirmed by CCP4 programs.

Site-directed mutagenesis of *subB*. Derivatives of SubAB with either S12A, Q36A or Y78F substitutions in the B subunit were constructed using overlap extension PCR mutagenesis. For the S12A mutant, this involved high-fidelity PCR amplification (Expand High Fidelity PCR kit; Roche Molecular Diagnostics) of *subAB*-positive O113:H21 STEC DNA using primer pairs pETsubAF/SubBA12R and SubBA12F/pETsubBR (Supplementary Table 5). This generated two fragments with the necessary mutation in the S12 codon of *subB* incorporated into the overlapping region by the SubBA12R and SubBA12F primers. The two separate PCR products were purified, mixed together and the complete *subAB* region re-amplified using primer pair pETsubAF/pETsubAR. These primers incorporate BamHI and XhoI restriction sites, enabling direct cloning into pET-23(+) (Novagen), followed by transformation into *E. coli* BL21(DE3) (Novagen). The other mutations were constructed in an analogous manner, using first-round primer pairs pETsubAF/SubBA36R and SubBA36F/pETsubBR for Q36A, and pETsubAF/SubBF78R and SubBF78F/pETsubBR for Y78F. Mutations were confirmed by sequencing, and expression of intact SubAB protein was confirmed by SDS–PAGE and western blotting. Holotoxins carrying confirmed B-subunit mutations were purified as for native toxin.

Tissue immunohistochemical studies. Human colon or kidney sections were overlaid with or without 1 µg ml^{-1} SubAB, $SubAB_{A12}$ or $SubAB_{F78}$, with or without 10% human or chimpanzee serum, as indicated, for 1 h at room temperature. Slides were then washed and fixed in 10% buffered formalin for 15 min and washed again. The sections were overlaid with rabbit anti-SubA at 1:5,000 in 1% BSA/PBS and incubated for 1 h at room temperature, washed again and overlaid with Cy3-labelled goat anti-rabbit IgG (Jackson Immunoresearch) for 1 h at room temperature. Slides were washed, mounted using aqueous mounting medium and viewed using epifluorescence with a Zeiss Axiophot microscope with appropriate excitation and barrier filters. Digital photomicrographs were taken using a Sony CCD (charge-coupled device) camera and National Institutes of Health image software and photo panels constructed using Adobe Photoshop and Illustrator.

SubAB binding inhibition ELISA. Inhibition of SubAB binding to immobilized Neu5Gcα2-3LacHSA by wild-type versus *Cmah*-null mouse serum, or human versus chimpanzee serum, was assayed as follows. All reagents were diluted and blocking was done in 1% fish gelatin. Costar 96-well ELISA plates were coated with saturating amounts of Neu5Gcα2-3LacHSA at 4 °C overnight, and then blocked at 4 °C overnight. Serum was added at the indicated concentrations, followed by 15 ng ml^{-1} SubAB. After 2 h incubation at room temperature, the plates were washed four times in PBS. Rabbit anti-SubAB serum diluted 1:5,000 was added. The plates were incubated for 2 h at room temperature, followed by another washing step, and then horseradish peroxidase conjugated to goat anti-rabbit IgG, 1:20,000, for 2 h at room temperature. After washing, 3,3′,5,5′-tetramethylbenzidine (Sigma) was added to the plates; the reaction was stopped after 15 min by 2 M sulphuric acid and absorbance was read at 450 nm.

Surface plasmon resonance. Experiments were conducted at 25 °C on a Biacore 3000 instrument using HBS buffer (10 mM HEPES–HCl, pH 7.4, 800 mM NaCl and 0.005% surfactant P20; supplied by the manufacturer). Approximately 200 resonance units of Neu5Acα2-3Lacβ-biotin and Neu5Gcα2-3Lacβ-biotin was immobilized onto streptavidin-coupled sensor chips (Biacore). SubAB was passed over all flow cells at 20 µl min^{-1} for 1 min. The final response was calculated by subtracting the response of the control surface from the glycan surface. For inhibition studies, SubAB (100 nM) was incubated with increasing concentrations of Neu5Gcα2-3Lacβ-biotin (3 µM to 4 mM) for 1 h at room temperature before being passed over the chip at 5 µl min^{-1} for 3 min. After each injection, the surface was regenerated with three injections of 10 mM glycine, pH 2.0. The experiments were performed in duplicate. The amount of SubAB bound at equilibrium was used to generate the inhibition curve that was analysed by nonlinear regression using PRISM software (version 3.0).

21. Leslie, A. G. W. in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography* 26 (SERC, 1992).
22. Evans, P. R. in *Proc. CCP4 Study Weekend on Recent Advances in Phasing* 97–102 (CCLRC, 1997).
23. Collaborative Computational Project. Number 4. *Acta Crystallogr. D* **50,** 760–763 (1994).
24. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).
25. Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59,** 38–44 (2003).
26. Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. Warp: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy models. *Acta Crystallogr. D* **30,** 551–554 (1997).
27. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
28. Murshodov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
29. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).
30. Schuettelkopf, A. W. & van Aalten, D. M. F. PRODRG – a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. D* **60,** 1355–1363 (2004).