Originally published as:

Yao, G., Zhang, S., Mahrhold, S., Lam, K.-H., Stern, D., Bagramyan, K., Perry, K., Kalkum, M., Rummel, A., Dong, M., Jin, R.
N-linked glycosylation of SV2 is required for binding and uptake of botulinum neurotoxin A (2016) Nature Structural and Molecular Biology, 23 (7), pp. 656-662.

DOI: 10.1038/nsmb.3245

N-linked glycosylation of SV2 sweetens the uptake of botulinum neurotoxin A1

Guorui Yao[1†], Sicai Zhang[2†], Stefan Mahrhold[3†], Kwok-ho Lam[1], Daniel Stern[4], Karine Bagramyan[5], Kay Perry[6], Markus Kalkum[5], Andreas Rummel[3*], Min Dong[2*], Rongsheng Jin[1*]

[1] Department of Physiology and Biophysics, University of California, Irvine, CA 92697, USA.

[2] Department of Urology, Boston Children's Hospital, Department of Microbiology and Immunobiology and Department of Surgery, Harvard Medical School, Boston, MA 02115, USA.

[3] Institut für Toxikologie, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30625 Hannover, Germany.

[4] Centre for Biological Threats and Special Pathogens - Biological Toxins (ZBS3), Robert Koch-Institut, Seestr. 10, 13353 Berlin, Germany.

[5] Beckman Research Institute of City of Hope, Department of Molecular Immunology, Duarte, CA 91010, USA.

[6] NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Building 436E, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA.

*Correspondence to: R.J. (r.jin@uci.edu), M.D. (min.dong@childrens.harvard.edu), or A.R. (rummel.andreas@mh-hannover.de).

† These authors contributed equally to this work.

ABSTRACT

Botulinum neurotoxin serotype A1 (BoNT/A1) is well adapted to invade motoneurons. Here we report a 2.0 Å resolution crystal structure of BoNT/A1 receptor-binding domain in complex with its neuronal receptor, the glycosylated human SV2C. We find that the remarkable neuronal tropism of BoNT/A1 is only achieved by directly recognizing both the peptide moiety and an N-linked glycan on SV2. This N-glycan—conserved in all SV2 isoforms across vertebrates—is essential for BoNT/A1 binding to neurons and its potent neurotoxicity. Remarkably, the glycan-binding interface is largely conserved in a newly identified mosaic toxin BoNT/HA (also known as mosaic BoNT/FA), and it is also the targeting site for a potent human BoNT/A1-neutralizing antibody that is currently licensed as an anti-botulism drug. Collectively, our studies reveal a new paradigm of host-pathogen interactions, in which pathogens exploit conserved host post-translational modifications to achieve highly specific receptor binding while also tolerating genetic changes across multiple isoforms of receptors.

BoNT/A1 is one of the most dangerous potential bioterrorism agents. Ironically, it is a licensed drug widely used in treating a variety of medical and cosmetic conditions [1]. According to a well-accepted dual-receptor model, the extreme potency of BoNT/A1 targeting motoneurons is mediated by its receptor-binding domain (H_CA), which synergistically binds to host protein receptors and gangliosides on the neuronal surface at neuromuscular junctions [2-4]. The synaptic vesicle glycoprotein 2 (SV2), a family of 12-transmembrane domain proteins that have three isoforms (SV2A, 2B, and 2C) in humans, are protein receptors for BoNT/A1 [5,6], as well as for BoNT/E [7], BoNT/D [8], and BoNT/F [9,10].

We have previously mapped the BoNT/A1-binding site to the fourth luminal domain of SV2s (SV2-L4) [5,6]. A crystal structure of H_CA in complex with the recombinant human SV2C-L4 expressed in *E. coli* (referred to as bSV2C with b indicating bacterial expression) has been reported recently [11]. It shows that H_CA-bSV2C recognition relies mostly on backbone-to-backbone interactions within a small interface (~596 $Å^2$), mediated by two β-strands in H_CA and one open edge of the quadrilateral β-helices of bSV2C [11]. This binding mode is in sharp contrast to BoNT/B, which recognizes its receptors synaptotagmin-I/II (Syt-I/II) through an extensive side-chain mediated protein-protein interaction network that ensures high binding affinity and specificity towards Syt-I/II [12,13]. So, how could BoNT/A1 possibly achieve extreme efficacy of targeting neurons using mostly backbone-mediated interactions for receptor recognition?

In this study, we determined the crystal structures of H_CA in complex with rat bSV2C-L4 and the physiologically more relevant glycosylated human SV2C-L4. We found that BoNT/A1 recognizes two distinct structural elements on SV2C: the protein moiety and an N-linked glycan that is conserved in all known SV2 homologs across vertebrates. Further biophysical, cellular and functional studies demonstrate that SV2 glycans are essential for BoNT/A1 binding to neuron and its extreme toxicity at its physiological site of action, the motor nerve terminals. Moreover, we found that the glycan-binding site of BoNT/A1 is also the target of a potent human neutralizing antibody, suggesting the potential for SV2 glycan as a novel target for developing BoNT inhibitors.

RESULTS

The crystal structure of H_CA in complex with rat bSV2C

Amino acid sequence analyses showed that even the few residues that mediate side-chain interactions in the H_CA and human bSV2C complex are not strictly conserved in SV2A and SV2B, or even SV2C from other species (e.g. rodents) (**Supplementary Fig. 1**). To gain a better insight into how BoNT/A1 can recognize SV2C from different species, we determined the crystal structure of H_CA in complex with rat SV2C-L4 expressed in *E. coli* (**Table 1**). The structure of the rat bSV2C-H_CA complex is virtually identical to that of the human bSV2C complex [root mean square deviation (RMSD) ~ 0.70 Å over 496 aligned Cα pairs]. Two major differences are observed. First, H_CA-R1294 forms hydrogen bonds with S519, C520, T521, and D539 of rat bSV2C (**Supplementary Fig. 2**), which are not observed in the structure of the human bSV2C-H_CA complex probably due to the different crystal packing modes [11]. Interestingly, R1294 only exists in two of the eight BoNT/A subtypes currently known (BoNT/A1 and A4). Second, a cation-π stacking interaction between BoNT/A1-R1156—exclusively existing in subtype BoNT/A1—and human SV2C-F563, previously thought to be critical for BoNT/A1-SV2C recognition [11], does not exist in the rat bSV2C-H_CA complex

because rat SV2C has a leucine (L563) in the place of human SV2C-F563. Leucine is also the homologous residue on SV2A and SV2B in both humans and rodents (**Supplementary Fig. 1c**). These findings suggest that the side-chain mediated interactions may vary significantly among different BoNT/A subtypes and SV2 isoforms, thus unlikely provide sufficient binding specificity and affinity between them. Therefore, some crucial BoNT/A1-SV2 interactions are missing in the crystal structures of H_CA-bSV2C complexes described here and previously [11].

SV2 glycosylation is crucial for BoNT/A1 binding to neurons

Native SV2s are glycosylated in neurons [14], and one of the three N-linked glycosylation motifs in L4—conserved in all SV2 isoforms across vertebrates—is located at the center of the BoNT/A1-binding interface of SV2 (e.g. N573 in SV2A, N516 in SV2B, and N559 in SV2C for human) (**Supplementary Fig. 1c**). To explore the functional role of this N-linked glycan, we utilized a molecule replacement approach to express either wild type (WT) SV2A-C, or the corresponding deglycosylation mutants (SV2A-N573A, SV2B-N516A, and SV2C-N559A) in hippocampal/cortical neurons cultured from SV2A/SV2B double knockout (KO) mice. Most hippocampal/cortical neurons do not express SV2C [14], and thus these neurons cultured from SV2A/SV2B double KO mice serve as a SV2-null neuron model. These deglycosylation mutations of SV2 do not affect protein-protein interactions with BoNT/A1, as e.g. bSV2C-N559A maintained a WT-like binding to H_CA [11].

The deglycosylation mutants of SV2A, 2B, and 2C all showed a lower molecular weight compared to WT SV2s, confirming that this Asn residue is indeed glycosylated in neurons (**Fig. 1 and Supplementary Fig. 3**). The expression level of SV2A-N573A was comparable to the WT SV2A, but there is a drastic reduction of SV2B-N516A expression and a mild reduction of SV2C-N559A expression (**Supplementary Fig. 3**). It suggests that glycosylation at this site is crucial for folding and/or stability of SV2B and to a lesser degree for SV2C—a known function of N-linked glycans [15].

We thus focused on SV2A-N573A and SV2C-N559A, and examined SV2-mediated toxin entry for BoNT/A1 and BoNT/D by analyzing the cleavage of synaptosomal-associated protein of 25 kDa (SNAP-25, the substrate of BoNT/A1) and synaptobrevin/vesicle-associated membrane protein 2 (VAMP2, the substrate of BoNT/D) after toxin exposure. BoNT/D serves here as an internal control for regular trafficking and sorting of SV2 mutants, as BoNT/D uses SV2s as receptors independent to N-glycosylation [8]. Furthermore, we have previously shown that BoNT/B1 that does not use SV2 as its receptor binds and enters the WT and SV2 KO neurons at similar levels [5,7]. We found that the SV2 KO neurons did not show BoNT/A1 entry at the conditions tested, which is likely due to BoNT/A1's low affinity for the ganglioside receptor that is unable to allow cell entry in the absence of SV2 [16,17]. Expression of SV2A-N573A and SV2C-N559A mediated significantly less entry of BoNT/A1 compared to the WT SV2s, whereas BoNT/D entry was not affected (**Fig. 1**). These results suggest that SV2 glycosylation at this strictly conserved site clearly contributes to BoNT/A1 binding and entry into neurons.

The SV2C glycans significantly enhances H_CA binding

Glycosylation is a common and highly diverse post-translational protein modification that profoundly alters the protein behavior [15]. Do the N-glycans contribute directly to BoNT/A1

binding? In this regard, we first carried out surface plasmon resonance (SPR) to examine how H_CA binds to the human bSV2C. We found that the binding displays a fast association rate (k_a: 1.5×10^6 $M^{-1}s^{-1}$) and a fast dissociation rate (k_d: 0.11 s^{-1}) (**Fig. 2a**), and the overall dissociation constant (K_D) of ~86 nM was comparable to a previously reported K_D of 260 nM determined by fluorescence anisotropy experiment [11]. The nature of this transient interaction is consistent with the pattern of backbone-to-backbone interactions revealed in the crystal structure, but difficult to reconcile with the extreme specificity of BoNT/A1 toward neurons.

We then expressed the human SV2C-L4 (residues V473-T567) as a secreted protein using human embryonic kidney 293 cells (HEK293), to mimic the physiologically relevant glycosylated receptor. HEK293 cells have been widely used to produce glycoproteins with human glycosylation patters [18]. The resulting protein (referred to as gSV2C, g stands for glycosylation) is glycosylated as evidenced by the appearance of multiple bands on SDS-PAGE that are bigger than its peptide mass and represent heterogeneous glycoforms. We characterized binding of H_CA to gSV2C by SPR, which revealed two binding components (**Fig. 2b-c and Supplementary Fig. 4**). A transient low affinity binding (K_D ~220 nM) closely resembles the binding of H_CA to bSV2C, which is likely due to heterogeneous glycosylation of SV2C under over-expression conditions [19]. Notably, a high affinity binding (K_D ~15 nM) of gSV2C displays a ~22-fold slower dissociation rate and a ~4-fold slower association rate compared to bSV2C. The slightly decreased association rate of gSV2C likely stems from restricted carbohydrate flexibility upon toxin binding, a physiological event that BoNT/A1 will encounter *in vivo* because the neuronal SV2s are glycosylated. Therefore, the relatively fast association rate displayed by bSV2C artificially increases the binding affinity of bSV2C. Taken together, these data demonstrate that the glycans of gSV2C stabilize the H_CA-gSV2C complex by significantly decreasing the dissociation rate.

The crystal structure of H_CA in complex with the glycosylated human SV2C

We next determined the crystal structure of H_CA in complex with gSV2C at 2.0 Å resolution (**Table 1**). The overall architecture of the H_CA-gSV2C complex is similar to that of the bSV2C complex (RMSD ~0.88 Å over 489 aligned Cα pairs). A complex-type N-linked glycan attached to gSV2C-N559 was observed with clear electron densities for the quadruple-saccharide core made up of two N-acetylglucosamine (NAG), a mannose (BMA), and a fucose (FUC) (**Fig. 3 and Supplementary Fig. 5**). There are two other putative N-linked glycosylation sites (N484 and N534) that are localized on SV2C-L4 and also conserved on SV2A and SV2B (**Supplementary Fig. 1c**). Only one NAG could be resolved in the electron density for the N534 glycan likely due to its high mobility, whereas no sugar could be identified at N484. SV2C N484 and N534 are located far away from BoNT/A1-binding interface, and thus unlikely to directly participate in toxin binding.

Remarkably the N559-glycan directly interacts with H_CA through a network of hydrogen bonds and van der Waals contacts. The most prominent are stacking interactions between residues F953 and H1064 of H_CA and the hydrophobic faces of the two NAG. In addition, ten well-defined water molecules act as a molecular "glue" in the glycan-H_CA interface to further strengthen the interactions (**Fig. 3d-e and Supplementary Table 1**). These interactions almost double the contact area between H_CA and SV2C from 557 to 925 $Å^2$. There are also weak electron densities beyond the mannose, suggesting that there might be more extensive H_CA-glycan interactions that

were not resolved, possibly due to the inherent flexibility of glycans and heterogeneous glycosylation. This structure unambiguously reveals that the N559-glycan of SV2C is recognized directly by BoNT/A1 as an integral part of the toxin-binding site.

BoNT/A1-glycan binding is conserved across SV2 isoforms and is critical for BoNT/A1 binding to neurons

We then sought to understand the functional role of protein-glycan interactions in BoNT/A1-SV2 recognition. Based on the crystal structure, we designed a set of single-site mutations on H_CA that selectively disrupt glycan binding: F953G and H1064G abolish the critical stacking of aromatic side chains against the sugar rings, and F953R, H1064R, G1292Q, and G1292R cause clashes between their bulky side chains and the N559-glycan. None of these mutations affected H_CA folding and stability as verified by thermal denaturation experiments, nor did they affect H_CA binding to bSV2C based on pull-down and SPR studies. As expected, these mutants markedly decreased H_CA binding to gSV2C, strongly supporting the direct involvement of N559-glycan of SV2C in BoNT/A1 binding (**Supplementary Figs. 6 and 7a-c**). Interestingly, SPR studies showed that the glycan-binding deficient H_CA-F953G binds to gSV2C much weaker than bSV2C (K_D: ~760 nM vs. ~160 nM) (**Supplementary Fig. 7d-f**). Therefore, the native neuronal SV2s that are always glycosylated initially impose steric hindrances for BoNT/A1 recognition. But BoNT/A1 manages to overcome this obstacle by directly employing SV2 glycans to strengthen binding.

We examined how these H_CA mutants bind to native SV2s in neurons. First, we analyzed binding of H_CA to endogenous SV2A and SV2B using cultured rat hippocampal/cortical neuron [14] (**Fig. 4a-b**). Mutations F953G, F953R, G1292Q, and G1292R largely abolished binding of H_CA to neurons; H1064G and H1064R also drastically reduced binding, suggesting that protein-glycan interactions are essential for BoNT/A1 binding to native SV2s on neurons. We further examined neurons that exclusively expressed SV2A, 2B, or 2C: hippocampal/cortical neurons cultured from SV2A(+/+)SV2B(-/-) mice served as neurons that only express SV2A, and neurons that only express SV2B or SV2C were created by infecting neurons cultured from SV2A(-/-)SV2B(-/-) mice with lentiviruses that express SV2B or SV2C, respectively. Mutating glycan-binding residues on H_CA (e.g. F953G, G1292R, and H1064G) reduced H_CA binding in all cases tested, demonstrating that BoNT/A1-glycan interactions are conserved and essential for all three SV2 isoforms (**Fig. 4c-e and Supplementary Fig. 8**).

The conserved SV2 N-glycan allows to tolerate genetic changes on various BoNT/A subtypes

Complementing our studies on protein-glycan interactions, we also examine the contribution of side-chain-mediated protein-protein interactions to SV2 recognition that could vary significantly among the eight different BoNT/A subtypes (**Supplementary Fig. 9**). In this regard, residues R1156 and R1294 of BoNT/A1 were mutated to their counterparts in other BoNT/A subtypes (e.g. R1156E and R1294S), which disrupt their interactions with the peptide moiety of SV2C but without affecting glycan binding. We used the double mutant H_CA-T1145A-T1146A as a control [11]. These two Thr residues, located at the core of protein-protein interface, are conserved in all BoNT/A subtypes (**Supplementary Fig. 9**). We found that binding of H_CA-T1145A-T1146A to

cultured rat hippocampal/cortical neurons was abolished (**Fig. 4a-b**). This is consistent with an earlier study showing that H_CA-T1145A-T1146A can no longer bind to bSV2C [11]. Interestingly, we found that H_CA-R1156E and R1294S still bound significantly to neurons (**Fig. 4a-b**), even though they showed decreased binding to both human bSV2C and gSV2C *in vitro* (**Supplementary Fig. 7**). These data suggest that loss of side-chain-mediated interactions at R1156 and R1294 of BoNT/A1 is tolerated on neuronal surfaces likely due to the presence of SV2 glycan and co-receptor gangliosides.

SV2 glycan-binding is essential for the extreme potency of BoNT/A1 at motor nerve terminals

To further establish the physiological relevance of protein-glycan interactions, we produced full-length BoNT/A1 containing single-site glycan-binding deficient mutations (F953G, F953R, H1064G, H1064R, G1292V, or G1292R) and examined their neurotoxicity at motor nerve terminals using an *ex vivo* mouse phrenic nerve hemi-diaphragm (MPN) assay (**Fig. 5**) [20]. Remarkably, all these mutations drastically reduced the potency of BoNT/A1. BoNT/A1-F953R had no detectable toxicity even at the maximal concentration tested (200 nM), which reflects a larger than 10^6-fold of toxicity reduction even though it displayed fully functional Zn^{2+}-endoprotease activity *in vitro*. Mutation G1292R also severely reduced the toxicity by 350-fold [21]. Mutations H1064G and H1064R displayed 3- and 7-fold reduction, respectively. These data demonstrate that glycan-binding is essential for the extreme potency of BoNT/A1 at the motor nerve terminals, its physiological site of action.

The SV2 glycan-binding mode is conserved in the newly identified BoNT/HA

We next examined whether glycan binding is conserved in natural variants of BoNT/A1, as BoNT genes are actively evolved and at least 40 different subtypes of BoNT have been reported [1]. In this regard, we focused on a newly reported mosaic toxin type HA (BoNT/HA, also known as BoNT/FA), which has a hybrid-like structure including a BoNT/A1-like H_C [22-26]. H_CHA is highly similar to H_CA (~83% identical) (**Supplementary Fig. 9**). Sequence alignment revealed that most of the glycan-binding residues, such as F953 and H1064, are conserved in H_CA and H_CHA, but residues R1156 and R1294 of H_CA are changed to M1148 and S1286 in H_CHA. H_CHA showed much weaker binding to human bSV2C as compared to H_CA, whereas its binding to human gSV2C was comparable to that of H_CA (**Fig. 6a**). It suggests that the loss of side-chain interactions due to genetic changes of M1148 and S1286 in H_CHA is compensated by SV2C glycan. In contrast, H_CHA-F943G (equivalent to H_CA-F953G) failed to bind to gSV2C based on a pull-down assay, nor did it bind to cortical neurons, likely due to the disruption of glycan binding (**Fig. 6a-b**). These data confirm that a similar glycan-binding mode is conserved in H_CHA and is critical for its binding to SV2s on neurons.

Furthermore, we found that the SV2 glycan-binding residues are largely conserved in seven out of the eight BoNT/A subtypes (BoNT/A1-A3 and A5-A8) that have been identified up to date (**Supplementary Fig. 9**) [27]. Notably, BoNT/A4 is the only one that has an Arg residue (R1292) at the equivalent position of BoNT/A1-G1292. Since BoNT/A1-G1292R mutant drastically decreases BoNT/A1 toxicity by blocking its binding to SV2 glycan (**Figs. 4-5 and Supplementary Figs. 7-8**) [21], we suggest that this is a major reason that causes the ~1,000-fold

reduced biological activity of BoNT/A4 compared to BoNT/A1 [28]. Taken together, our findings suggest that BoNT/A1 and variants utilizes the genetically invariable carbohydrates as surrogate amino acids to engage the host receptors. Synergistic binding to two distinct cell surface receptors, SV2 (including its peptide and glycan moieties) and ganglioside, provides a plausible explanation for the extreme potency of BoNT/A and its remarkable specificity for nerve terminals (**Fig. 6c**).

A new strategy for developing therapeutic antibodies against BoNT/A

The novel glycan receptor for BoNT/A presents a promising target for developing toxin inhibitors. Remarkably, we found that a BoNT/A1-neutralizing human monoclonal antibody family CR2/CR1, which is currently in clinical trial [29,30], directly targets the glycan-binding site on H_CA with its first antigen-binding loop of the light chain variable region occupying the glycan-binding site (**Fig. 7a**). The key epitopes for CR2/CR1 include precisely residues F953 and H1064 [29]. Interestingly, residue F36 of CR2/CR1 uses a π-stacking interaction to bind BoNT/A1-H1064, mimicking its interactions with SV2C N559-glycan. CR2/CR1 also blocks binding of BoNT/A1 to bSV2C [21], but this is due to the large size of CR2/CR1 causing side-to-side clash with bSV2C, as CR2/CR1 and bSV2C have non-overlapping binding sites on H_CA (**Fig. 7b-d**). Together our data reveal that the strong neutralization potency of CR2/CR1 is empowered by simultaneously blocking BoNT/A1 binding to the glycan and peptide moieties of SV2.

DISCUSSION

Host cell-surface glycans are crucial for pathogen recognition. For example, influenza hemagglutinins use carbohydrates in determining the host range (e.g. swine, avian or human) [31]. Our results reveal a novel host-recognition strategy, by which pathogens simultaneously recognize a protein segment and the neighboring glycans as a composite binding site. This unique strategy uses a conserved post-translational modification as an evolutionarily static recognition site, in addition to protein-protein interactions that encode the location and specificity information. Together, it provides a powerful solution to address the competing needs of achieving highly specific binding while also tolerating residue changes in receptors across multiple isoforms and species variants.

Intriguingly, a similar strategy may be utilized by some important broad-neutralizing human antibodies, which are capable of neutralizing multiple serotypes of targeted viruses, such as dengue viruses and immunodeficiency virus-1 (HIV-1), by simultaneously recognizing the protein components and the highly conserved glycans on virus proteins [32-34]. A similar model was also recently demonstrated for Notch1 receptor binding with its ligand Delta like 4 (DLL4) [35], where DLL4 binds to the fucose and glucose at the base of an O-glycosylation site located within the protein-protein interface. Therefore, our findings offer an important strategy for engineering ligand-receptor interactions and broadly neutralizing antibodies for therapeutic applications.

Accession codes. Atomic coordinates and structure factors for the H_CA-bSV2C and H_CA-gSV2C complexes will be deposited in the Protein Data Bank.

ACKNOWLEGEMENTS

This work was partly supported by National Institute of Allergy and Infectious Diseases (NIAID) grants R01AI091823 and R21AI123920 to R.J. and R01AI096169 to M.K; by National Institute of Neurological Disorders and Stroke (NINDS) Grant R01NS080833 to M.D.; and by the Bundesministerium für Bildung und Forschung grants FK031A212A to A.R. and FK031A212B to Brigitte G. Dorner (RKI). NE-CAT at the Advanced Photon Source (APS) is supported by a grant from the National Institute of General Medical Sciences (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). Use of the APS, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

AUTHOR CONTRIBUTIONS

G.Y. and S.M. performed the cloning and mutagenesis. G.Y., K.L. and R.J. carried out the protein expression, purification, characterization and crystallographic studies. K.P. collected the X-ray diffraction data. S.Z. and M.D. performed all experiments on cultured neurons. A.R. and S.M. generated the full-length BoNT/A1 mutants and performed the MPN assay with support of Nadja Krez. Dr. Jasmin Weisemann cloned H_CHA. D.S., K.B. and M.K. performed the SPR studies. R.J., M.D. and A.R. wrote the manuscript with input from other authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

1. Rossetto, O., Pirazzini, M. & Montecucco, C. Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat Rev Microbiol* **12**, 535-49 (2014).
2. Montecucco, C. How do tetanus and botulinum neurotoxins bind to neuronal membranes? *Trends Biochem Sci* **11**, 314-17 (1986).
3. Montecucco, C., Rossetto, O. & Schiavo, G. Presynaptic receptor arrays for clostridial neurotoxins. *Trends Microbiol* **12**, 442-6 (2004).
4. Rummel, A. Double Receptor Anchorage of Botulinum Neurotoxins Accounts for their Exquisite Neurospecificity. *Curr Top Microbiol Immunol* **364**, 61-90 (2013).
5. Dong, M. et al. SV2 is the protein receptor for botulinum neurotoxin A. *Science* **312**, 592-6 (2006).
6. Mahrhold, S., Rummel, A., Bigalke, H., Davletov, B. & Binz, T. The synaptic vesicle protein 2C mediates the uptake of botulinum neurotoxin A into phrenic nerves. *FEBS Lett* **580**, 2011-4 (2006).
7. Dong, M. et al. Glycosylated SV2A and SV2B mediate the entry of botulinum neurotoxin E into neurons. *Mol Biol Cell* **19**, 5226-37 (2008).
8. Peng, L., Tepp, W.H., Johnson, E.A. & Dong, M. Botulinum neurotoxin D uses synaptic vesicle protein SV2 and gangliosides as receptors. *PLoS Pathog* **7**, e1002008 (2011).

9. Rummel, A. et al. Botulinum neurotoxins C, E and F bind gangliosides via a conserved binding site prior to stimulation-dependent uptake with botulinum neurotoxin F utilising the three isoforms of SV2 as second receptor. *J Neurochem* **110**, 1942-54 (2009).
10. Fu, Z., Chen, C., Barbieri, J.T., Kim, J.J. & Baldwin, M.R. Glycosylated SV2 and gangliosides as dual receptors for botulinum neurotoxin serotype F. *Biochemistry* **48**, 5631-41 (2009).
11. Benoit, R.M. et al. Structural basis for recognition of synaptic vesicle protein 2C by botulinum neurotoxin A. *Nature* **505**, 108-11 (2014).
12. Chai, Q. et al. Structural basis of cell surface receptor recognition by botulinum neurotoxin B. *Nature* **444**, 1096-100 (2006).
13. Jin, R., Rummel, A., Binz, T. & Brunger, A.T. Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. *Nature* **444**, 1092-5 (2006).
14. Janz, R. & Sudhof, T.C. SV2C is a synaptic vesicle protein with an unusually restricted localization: anatomy of a synaptic vesicle protein family. *Neuroscience* **94**, 1279-90 (1999).
15. Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-9 (2001).
16. Rummel, A., Mahrhold, S., Bigalke, H. & Binz, T. The HCC-domain of botulinum neurotoxins A and B exhibits a singular ganglioside binding site displaying serotype specific carbohydrate interaction. *Mol Microbiol* **51**, 631-43 (2004).
17. Pirazzini, M., Rossetto, O., Bolognese, P., Shone, C.C. & Montecucco, C. Double anchorage to the membrane and intact inter-chain disulfide bond are required for the low pH induced entry of tetanus and botulinum neurotoxins into neurons. *Cell Microbiol* **13**, 1731-43 (2011).
18. Swiech, K., Picanco-Castro, V. & Covas, D.T. Human cells: new platform for recombinant therapeutic protein production. *Protein Expr Purif* **84**, 147-53 (2012).
19. Croset, A. et al. Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *J Biotechnol* **161**, 336-48 (2012).
20. Bigalke, H. & Rummel, A. Botulinum Neurotoxins: Qualitative and Quantitative Analysis Using the Mouse Phrenic Nerve Hemidiaphragm Assay (MPN). *Toxins (Basel)* **7**, 4895-905 (2015).
21. Strotmeier, J. et al. Identification of the synaptic vesicle glycoprotein 2 receptor binding site in botulinum neurotoxin A. *FEBS Lett* **588**, 1087-93 (2014).
22. Barash, J.R. & Arnon, S.S. A novel strain of Clostridium botulinum that produces type B and type H botulinum toxins. *J Infect Dis* **209**, 183-91 (2014).
23. Gonzalez-Escalona, N. et al. Draft Genome Sequence of Bivalent Clostridium botulinum Strain IBCA10-7060, Encoding Botulinum Neurotoxin B and a New FA Mosaic Type. *Genome Announc* **2**(2014).
24. Kalb, S.R. et al. Functional characterization of botulinum neurotoxin serotype H as a hybrid of known serotypes F and A (BoNT F/A). *Anal Chem* **87**, 3911-7 (2015).
25. Maslanka, S.E. et al. A Novel Botulinum Toxin, Previously Reported as Serotype H, has a Hybrid Structure of Known Serotypes A and F that is Neutralized with Serotype A Antitoxin. *J Infect Dis* (2015).
26. Fan, Y. et al. Immunological Characterization and Neutralizing Ability of Monoclonal Antibodies Directed Against Botulinum Neurotoxin Type H. *J Infect Dis* (2016).
27. Kull, S. et al. Isolation and functional characterization of the novel Clostridium botulinum neurotoxin A8 subtype. *PLoS One* **10**, e0116381 (2015).

28. Whitemarsh, R.C. et al. Characterization of botulinum neurotoxin A subtypes 1 through 5 by investigation of activities in mice, in neuronal cell cultures, and in vitro. *Infect Immun* **81**, 3894-902 (2013).
29. Garcia-Rodriguez, C. et al. Molecular evolution of antibody cross-reactivity for two subtypes of type A botulinum neurotoxin. *Nat Biotechnol* **25**, 107-16 (2007).
30. Nayak, S.U. et al. Safety and pharmacokinetics of XOMA 3AB, a novel mixture of three monoclonal antibodies against botulinum toxin A. *Antimicrob Agents Chemother* **58**, 5047-53 (2014).
31. Skehel, J.J. & Wiley, D.C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* **69**, 531-69 (2000).
32. Rouvinski, A. et al. Recognition determinants of broadly neutralizing human antibodies against dengue viruses. *Nature* (2015).
33. Garces, F. et al. Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* **159**, 69-79 (2014).
34. Kong, L. et al. Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat Struct Mol Biol* **20**, 796-803 (2013).
35. Luca, V.C. et al. Structural biology. Structural basis for Notch1 engagement of Delta-like 4. *Science* **347**, 847-53 (2015).
36. Laskowski, R.A. & Swindells, M.B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-86 (2011).
37. Stenmark, P., Dupuy, J., Imamura, A., Kiso, M. & Stevens, R.C. Crystal structure of botulinum neurotoxin type A in complex with the cell surface co-receptor GT1b-insight into the toxin-neuron interaction. *PLoS Pathog* **4**, e1000129 (2008).
38. Matsumiya, S. et al. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J Mol Biol* **368**, 767-79 (2007).

FIGURE LEGENDS

Figure 1 SV2 glycosylation is critical for BoNT/A1 binding and entry into neurons. Lentivirus was used to infect mouse SV2A(-/-)SV2B(-/-) neurons to create neurons that express either the WT or the deglycosylation mutants of SV2A **(a)** or SV2C **(b)**. Neurons were simultaneously exposed to BoNT/A1 (1 nM) and BoNT/D (0.1 nM) in a high K^+ buffer for 5 minutes at 37°C. Unbound toxin was washed away, and the intoxicated cells were incubated for another 8 hour. Cell lysates were harvested and subjected to immunoblot analysis using mouse monoclonal antibodies against SNAP-25 (Cl 71.2), VAMP2 (Cl 69.1), and SV2 (pan-SV2). BoNT/D served as an internal control to confirm that SV2 mutants still sorted and localized correctly. Cleavage of SNAP-25 by BoNT/A1 generates a smaller fragment that is marked by an asterisk. Actin served as a loading control.

Figure 2 Glycosylation in SV2 luminal domain enhances BoNT/A1 binding. **(a)** Sensorgrams of H_CA binding to immobilized human bSV2C (blue) overlaid with a fit of 1:1 binding model (black). Since equilibrium binding was reached for all conditions tested, a steady state affinity was also determined (insert). **(b)** H_CA binding to immobilized human gSV2C (red) was best fit to a heterogeneous binding model (black), whereas a 1:1 binding model was inapplicable (**Supplementary Fig. 4**). Insert: individual contributions of the two binding events observed in the heterogeneous interaction (blue and green). **(c)** Kinetic binding rates and affinities for H_CA binding to human bSV2C or gSV2C. Shown values represent the mean ± S.D. (n = 2 for bSV2C, n = 3 for gSV2C).

Figure 3 Structure of H_CA in complex with human gSV2C. **(a)** Cartoon representation of the complex whereas H_CA is gold and gSV2C is green. The black oval highlights the interacting β-strands between them. gSV2C-N559 and the attached N-linked glycan are shown in the stick models. **(b)** A ~150° rotation of the complex about a vertical axis. **(c-d)** Close-up views of the protein-protein and protein-glycan association interfaces between H_CA and gSV2C. H_CA is in surface representation (gold), whereas H_CA residues that directly interact with the peptide moiety of gSV2C or the N559 glycan are colored purple and light blue, respectively. Well-defined water molecules that mediate H_CA-glycan binding are shown as green spheres. **(e)** Extensive interactions between gSV2C-N559 glycan and H_CA. A schematic representation of the glycan structure is shown. The plots were generated using LIGPLOT [36]. H_CA and gSV2C residues are labeled brown and green, respectively. Hydrogen bonds are indicated by dashed green lines. Key hydrogen bonding distances are listed in **Supplementary Table 1**. Residues involved in hydrophobic interactions are represented by an arc with spokes radiating towards the binding partners they contact.

Figure 4 Site-directed mutagenesis analysis of the SV2-binding site on H_CA. **(a-b)** Binding of H_CA variants (100 nM, 5 min, 37°C in a high K^+ buffer) to rat hippocampal/cortical neurons was analyzed by immunostaining (panel a) or immunoblot (panel b). Scale bar, 20 μm. Synaptophysin (Syp) served as a loading control. **(c-e)** Binding of H_CA variants (100 nM, 5 min) to neurons that express individual SV2A, 2B, or 2C was analyzed by immunoblot. Hippocampal/cortical neurons cultured from SV2A(+/+)SV2B(-/-) mice served as neurons that only express SV2A (panel c). Neurons that only express SV2B or SV2C were created by infecting neurons cultured from SV2A(-/-)SV2B(-/-) mice with lentiviruses that express SV2B

(panel d) or SV2C (panel e). The results of immunostaining analysis are shown in **Supplementary Fig. 8**.

Figure 5 Glycan binding-deficient BoNT/A1 displays drastic decreased neurotoxicity as examined by the mouse phrenic nerve (MPN) assay. BoNT/A1 F953R mutant showed no detectable toxicity even at the maximal concentration tested (200 nM), indicating a larger than 10^6-fold reduction on toxicity. Shown are mean ± S.D. of 3-6 technical replicates.

Figure 6 The SV2 glycan binding mode is conserved in BoNT/HA. **(a)** The SV2C glycan plays an indispensable role for H_CHA binding as demonstrated by a pull-down assay using human bSV2C (with a SUMO tag) or gSV2C as baits. "Mu" is H_CHA-F943G mutant. **(b)** The F943G mutation dramatically decreased H_CHA binding to rat cortical neurons as analyzed by immunostaining and immunoblot. **(c)** Proposed model for simultaneous binding of BoNT/A1 to two neuronal surface receptors: glycosylated SV2 and ganglioside. GT1b is modeled based on the structure of a GT1b-bound H_CA (PDB code 2VU9) [37]. A representative complex type N-linked glycan is modeled based on the structure of a glycan of human IgG1 (PDB code 3AVE) [38]. The glycan core that is observed in the H_CA-gSV2C complex is colored green, whereas the remaining carbohydrates (mannose and NAG colored gray) could potentially extend to the N-terminal sub-domain of H_CA ($H_{CN}A$).

Figure 7 The SV2 glycan-binding site on BoNT/A1 is the target for the neutralizing antibody CR1. **(a-b)** BoNT/A1-neutralizing therapeutic antibody CR1 (PDB code: 2NYY) [29] occupies the SV2C glycan-binding site on H_CA, but does not affect H_CA's SV2C peptide-binding site. An arrow in the schematic diagram indicates the side-to-side clash between SV2C and CR1. **(c)** A close-up view of the interface. gSV2C-N559 glycan (green), H_CA-F953 and -H1064 (gold), residues 30-36 of CR1 ligand chain (LC, red), and residues 101-103 of CR1 heavy chain (HC, gray) are shown in stick models. **(d)** Residues of H_CA that exclusively bind SV2C peptide, N559 glycan or CR1 are colored in purple, green or cyan, respectively, while the H_CA residues that are contacted by both N559 glycan and CR1 are in blue.

Table 1 Data collection and refinement statistics.

	HcA-bSV2C (rat)	HcA-gSV2C (human)
Data collection		
Space group	P 1 21 1	C 1 2 1
Cell dimensions		
a, *b*, *c* (Å)	88.66, 143.99, 110.92	109.00, 111.85, 126.25
α, β, γ (°)	90, 93.6, 90	90, 101.3, 90
Resolution (Å)	87.76-2.64 (2.73-2.64)	123.8-2.0 (2.03-2.00)
R_{meas}	0.172 (1.454)	0.159 (0.889)
CC1/2	0.991 (0.599)	0.984 (0.722)
I/σ(I)	9.1 (1.4)	6.7 (2.1)
Completeness (%)	99.44 (99.20)	99.6 (99.9)
Redundancy	3.4 (3.1)	2.9 (2.9)
Refinement		
Resolution (Å)	87.76-2.64	123.81-2.00
No. reflections	80,958	99,659
R_{work}/R_{free}	0.239/0.271	0.176/0.216
No. atoms		
Protein	16,825	8,645
Ligand/ion	-	128
Water	162	659
B-factors		
Protein	58.50	44.50
Ligand/ion	-	58.70
Water	55.40	51.50
r.m.s. deviations		
Bond lengths (Å)	0.012	0.009
Bond angles (°)	1.22	1.01

Statistics for the highest-resolution shell are shown in parentheses.

METHODS

Construct design and cloning. The gene of H_CA (residues N872–L1296) was cloned into expression vector pQE30 with an N-terminal 6xHis-tag and a PreScission protease cleavage site. H_CHA (residues E860–L1286) was cloned into pGEX-4T-2 vector that has a thrombin cleavage site following GST. The core region of human SV2C-L4 (residues V473–T567) was cloned into two different vectors: pET28a vector for *E. coli* expression and a pcDNA vector for mammalian cell expression. For *E. coli* expression, a 6xHis/SUMO (*Saccharomyces cerevisiae* Smt3p) tag was introduced to the N-terminus of SV2C-L4 to facilitate protein expression and purification (SUMO-bSV2C). For mammalian cell expression, a human IL2 signal sequence (MYRMQLLSCIALSLALVTNS), a 9xHis-tag, and a Factor Xa cleavage site were added to the N-terminus of SV2C-L4 (gSV2C). A second SV2C mammalian expression construct was made based on gSV2C by inserting SUMO between the Factor Xa site and SV2C (SUMO-gSV2C). Rat SV2C-L4 (residues P455–Y577) was covalently linked to the C-terminus of H_CA through a peptide linker composed of a thrombin cleavage site (LVPRGS) and a PreScission protease cleavage site (LEVLFQGP). The covalently linked H_CA-rbSV2C was cloned into pET28a vector with an N-terminal 6xHis-tag and a thrombin cleavage site. All H_CA and H_CHA mutations were generated using QuikChange site-directed mutagenesis (Agilent). All pH6tBoNTA mutants were prepared using the GeneTailor method (Invitrogen GmbH, Karlsruhe, Germany) employing suitable primers and pH6tBoNTA as template DNA [21].

Protein expression and purification. H_CA, H_CHA, SUMO-bSV2C, and H_CA-rbSV2C were expressed in *E. coli* strain BL21-Star (DE3) (Invitrogen). Bacteria were cultured at 37°C in LB medium containing appropriate selecting antibiotics. The temperature was reduced to 18°C when OD_{600} reached 0.4. Expression was induced with 0.2 mM IPTG (isopropyl-b-D-thiogalactopyranoside) when OD_{600} reached 0.7, and continued at 18°C for ~16 hours. The cells were harvested by centrifugation and stored at -80°C until use.

Wild-type and mutated recombinant full-length neurotoxin H6tBoNTA were produced under biosafety level 2 containment (project number GAA A/Z 40654/3/123) utilizing the *E. coli* strain M15pREP4 (Qiagen, Hilden, Germany) during 16 h of induction at 22°C in the presence of 0.2 mM IPTG, and were purified on Co^{2+}-Talon matrix (Takara Bio Europe S.A.S., France). Full-length neurotoxins were eluted using 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 250 mM imidazole, subjected to size-exclusion chromatography (SEC; Superdex-200 16/60 column, GE Healthcare, Germany) in 100 mM Tris-HCl, pH 8.0, and 150 mM NaCl, frozen in liquid nitrogen and kept at -70°C.

The His-tagged proteins (H_CA, SUMO-bSV2C and H_CA-rbSV2C) were purified using Ni^{2+}-NTA (nitrilotriacetic acid, Qiagen, Hilden, Germany) affinity resins in a buffer containing 50 mM Tris, pH 8.0, 400 mM NaCl, and 40 mM imidazole. The proteins were eluted with a high-imidazole buffer (50 mM Tris, pH 8.0, 400 mM NaCl, and 300 mM imidazole) and then dialyzed at 4°C against a buffer containing 20 mM HEPES, pH 7.5, and 150 mM NaCl. The His-tag of H_CA and the His/SUMO-tag of SUMO-bSV2C were cleaved by PreScission and SUMO proteases, respectively. For the covalently linked H_CA-rbSV2C, both thrombin and PreScission protease were used to cut the linker between the two proteins in order to avoid potential conformational constraint. GST-H_CHA fusion protein was purified using glutathione Sepharose

4B affinity resin (GE Healthcare) in a buffer containing 20 mM HEPES, pH 7.5, and 150 mM NaCl. H_CHA was then released from the resins by on-column cleavage using thrombin.

Tag-cleaved H_CA and H_CHA were further purified by MonoS ion-exchange chromatography (GE Healthcare) in a buffer containing 50 mM MES, pH 6.0, and eluted with a NaCl gradient. The peak fractions were then subjected to Superdex-200 SEC (GE Healthcare) in a buffer containing 20 mM sodium phosphate, pH 6.0, and 50 mM NaCl. SUMO-bSV2C (His/SUMO-tag cleaved or un-cleaved) and the cleaved H_CA-rbSV2C were further purified by Superdex-200 SEC in a buffer containing 20 mM HEPES, pH 7.5, and 150 mM NaCl. H_CA-rbSV2C was concentrated to ~2 mg/ml for crystallization.

SUMO-gSV2C and gSV2C were expressed and secreted from HEK 293 cells (BioLegend) and purified directly from cell culture media using Ni^{2+}-NTA. The proteins were eluted from the resins with high concentration of imidazole and dialyzed against a buffer containing 50 mM Tris, pH 8.0, and 400 mM NaCl. gSV2C was then mixed with the purified H_CA at a molar ratio of ~1:2, and the H_CA-gSV2C complex was isolated by Ni^{2+}-NTA resins. After dialysis against a buffer containing 20 mM HEPES, pH 7.5, and 150 mM NaCl, the complex was further purified by Superdex-200 SEC using the same buffer. The complex was concentrated to ~10 mg/ml for crystallization.

Crystallization. Initial crystallization screens were carried out using a Gryphon crystallization robot (Art Robbins Instruments) with high-throughput crystallization screening kits (Hampton Research and Qiagen). Extensive manual optimization was then performed at 20°C using the hanging-drop vapor-diffusion method when proteins were mixed with reservoir solutions in 1:1 ratio. H_CA-rbSV2C was initially crystallized in a condition containing 100 mM sodium cacodylate, pH 6.5, 13% polyethylene glycol (PEG) 3,350, and 200 mM NaCl. The best crystals were obtained in the presence of 0.7% (v/v) 1-butanol, which was identified using an additive screen kit (Hampton Research). The crystals were cryo-protected in the original mother liquor supplemented with 20% (v/v) glycerol and flash-frozen in liquid nitrogen. The H_CA-gSV2C complex was originally crystallized as thin plates in a condition composed of 100 mM sodium acetate, pH 4.6, 20% PEG 3,350, and 200 mM ammonium phosphate monobasic. These crystals diffracted poorly. After extensive additive screening and optimization, the best crystals were obtained in the presence of 4% (w/v) polypropylene glycol P 400. The crystals were cryo-protected in the reservoir solution supplemented with 20% (v/v) ethylene glycol and flash-frozen in liquid nitrogen.

Data collection and structure determination. The X-ray diffraction data were collected at 100 K at the NE-CAT beamline 24-ID, Advanced Photon Source (APS). The data were processed with XDS [39]. The structure of H_CA-rbSV2C was determined by molecular replacement software Phaser [40] using the structure of H_CA (PDB code 3FUO) [10] as the search model. The structural model of rbSV2C was manually built, and the structural modeling and refinement were carried out iteratively using COOT [41] and Refmac from the CCP4 suite [42]. This structure was later used as the search model to determine the structure of the H_CA-gSV2C complex. All the refinement progress was monitored with the free R value using a 5% randomly selected test set [43]. The structures were validated through the MolProbity web server [44] and showed excellent

stereochemistry. Data collection and structural refinement statistics are listed in Table 1. All structure figures were prepared with PyMol (http://www.pymol.org).

Pull-down assay. The pull-down assay was performed using Ni^{2+}-NTA resins in 1 ml buffer containing 50 mM Tris, pH 8.0, 400 mM NaCl, 10 mM imidazole, and 0.1% Tween-20. SUMO-bSV2C or gSV2C served as the bait while H_CA or H_CHA variants were the preys. SV2C was pre-incubated with Ni^{2+}-NTA resins at 4°C for 1 hour, and the unbound protein was washed away. The resins were then divided into small aliquots (~ 5µg of bait) and mixed with the preys (~ 30 µg). The pull-down assay was carried out at 4°C for ~1.5 hours. The resins were washed twice, and the bound proteins were released from resins using 300 mM imidazole.

Protein melting assay. The thermal stability of H_CA or H_CHA variants was measured using a fluorescence-based thermal shift assay on a StepOne real-time PCR machine (Life Technologies). The protein (~5 µM) was mixed with the fluorescent dye SYPRO Orange (Sigma-Aldrich) immediately before the experiment. The samples were heated from 25°C to 90°C in ~ 45 min. The midpoint of the protein-melting curve (Tm) was determined using the analysis software provided by the instrument manufacturer. The data obtained from three independent experiments were averaged to generate the bar graph.

Surface plasmon resonance (SPR). Binding kinetics and affinity were determined on a Biacore X100 unit (GE Healthcare) at 25°C using HBS-EP+ (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% Tween-20) as running buffer at a flow rate of 30 µL/min. Using standard EDC/NHS amine coupling chemistry, SUMO-gSV2C or SUMO-bSV2C were coupled on flow cell (Fc) 2 of a CM5 sensor chip (GE Healthcare) to a surface density of ~83 resonance units (RUs) or ~105 RUs, respectively. Control Fc1 was blank immobilized by EDC/NHS activation before blocking with 1 M ethanolamine (GE Healthcare).

For kinetic measurements, H_CA was injected in 1:3 dilution series ranging from 1,200 nM to 4.94 nM for gSV2C or 2,000 nM to 2.74 nM for bSV2C. Each measurement was started and ended with injection of the highest analyte concentration to ensure retained binding capacity. Association was monitored for 120 seconds by analyte injection followed by 300 second injections of running buffer to monitor binding dissociation. Between measurements, the surface was regenerated by 60 second injections of 10 mM glycine-HCl (pH 1.7) at 10 µL/min. Binding kinetics were determined by fitting the double referenced [45] binding curves using the heterogeneous (gSV2C) or the 1:1 Langmuir binding models (bSV2C) with global R_{max} and RI set to zero (Biacore Evaluation Software 2.01). Due to the highly transient interaction of H_CA with bSV2C, equilibrium binding was reached for all H_CA concentrations tested. Therefore, steady state affinity was determined for bSV2C by fitting a four-parametric Hill equation to the binding responses 65 seconds after analyte injection over log-transformed H_CA-concentrations using Prism 5.04 (GraphPad) with bottom constraint set to zero. The good agreement between the binding affinity determined by kinetic analysis and the steady state affinity proved that kinetic binding rates for bSV2C were reliable despite being close to the measurement limits of the instrument. Kinetic binding rates were determined by n = 2 (bSV2C) or n = 3 (gSV2C) independent experiments. Shown values represent the mean ± standard deviation.

To compare the interaction of H_CA WT and its mutants with bSV2C (511 RUs) and separately with gSV2C (479 RUs), the reactions were performed in the “interactive manual run mode”. The

H_CA variants were injected at concentrations of 0.01, 1.0, 10, 25, 50, 75, 100 and 200 nM, each in triplicate, and at a flow-rate of 10 µL/min. The chip surface was regenerated by two injections of glycine-HCl (10 mM, pH 2.0, 30 sec contact time) allowing the removal of the analyte without changing the activity of the immobilized ligand. This was confirmed by the equal responses obtained from the binding assays before and after regeneration. Results were plotted as response (RU) versus concentration of H_CA using Prism 6 (GraphPad Software Inc., La Jolla, CA).

Cell biology materials and constructs. SV2A and SV2B knockout mice were obtained from the Jackson Laboratory. Rat cDNAs encoding SV2A, SV2B, and SV2C were generously provided by R. Janz (Houston, TX). They were cloned into a lentiviral vector (Lox-Syn-Syn) as we previously described [8]. This vector contains two separate neuronal-specific synapsin promoters. One promoter drives expression of SV2 and the other one drives expression of GFP as a marker. Human cDNA encoding SV2C was obtained from PlasmID repository of Harvard Medical School. Human monoclonal antibody against H_CA (RAZ-1) was generously provided by J. Marks (San Francisco, CA). Mouse monoclonal antibodies against VAMP2 (Cl 69.1), SNAP-25 (Cl 71.2), SV2 (pan-SV2), Syp (Cl 7.2) were generously provided by E. Chapman (Madison, WI) and are available from Synaptic Systems (Göttingen, Germany). The following antibodies were purchased from indicated vendors: mouse monoclonal antibody against actin (Sigma); rabbit polyclonal antibody against synapsin (Millipore); rabbit polyclonal antibody against GFP (Abcam). Purified BoNT/A1 of Hall-A strain and BoNT/D of D1873 strain were generously provided by E. Johnson (Madison, WI).

Neuron culture and lentivirus transduction. Rat hippocampal/cortical neurons were prepared from E18-19 embryos. Mouse SV2A/B double knockout neurons were prepared from postnatal day 1 pups as previously described [8]. Dissected hippocampi and cortex were dissociated with papain following manufacture instructions (Worthington Biochemical, NJ). Cells were plated on poly-D-lysine coated coverslips. Experiments were carried out generally using DIV (days in vitro) 13-15 neurons. Lentivirus were prepared as described previously using HEK293FT cells [8]. Viruses were added to neurons at DIV5.

H_CA and H_CHA binding to neurons. Neurons were exposed to 100 nM H_CA or H_CHA in high K^+ buffer, which contains (mM): NaCl 140, KCl 3, KH_2PO_4 1.5, Na_2HPO_4 8, $MgCl_2$ 0.5, and $CaCl_2$ 1, for 5 minutes at 37°C. Cells were then washed three times with phosphate-buffered saline (PBS). Binding of H_CA or H_CHA was examined using two complementary approaches. (1) Immunostaining: neurons were fixed with 4% paraformaldehyde and permeabilized with 0.3% Triton X-100 in PBS solution. Images were collected using a Leica TCS SP8 confocal microscope with a 40X oil objective. (2) Immunoblot analysis: neurons were harvested in a lysis buffer (PBS with 1% Triton X-100, 0.05% SDS and protease inhibitor cocktail (Roche, CA), 100 µl per one well of 24-well plates). Lysates were centrifuged for 10 minutes at 4°C and the supernatants were subjected to SDS-PAGE and western blot analysis. Binding of H_CA or H_CHA were detected using a monoclonal human anti-H_CA antibody (RAZ-1), which recognizes both H_CA and H_CHA. It also recognizes mutants of H_CA and H_CHA examined in this study with similar sensitivity ("input" lanes in Figs. 4b and 6b)

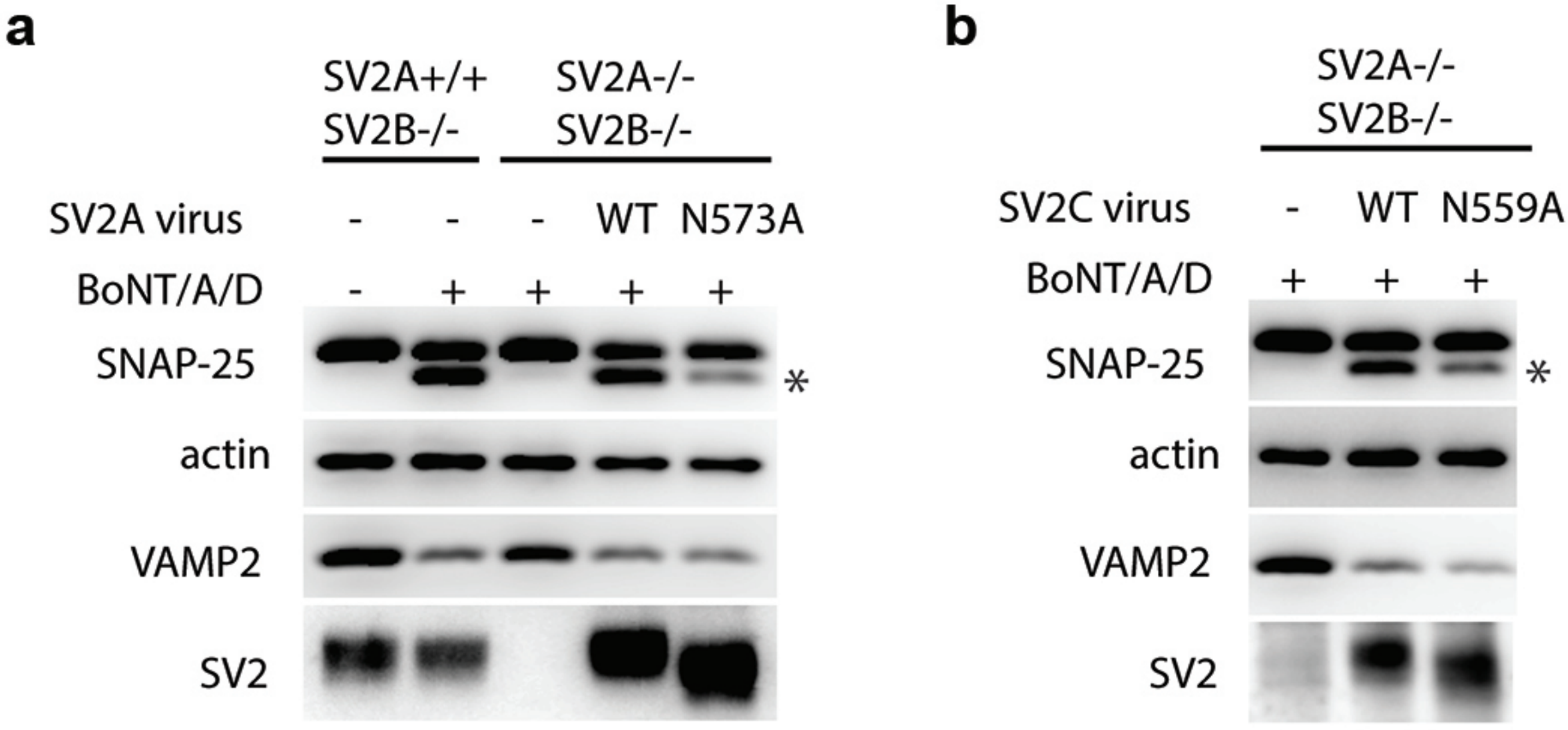
Entry of BoNTs into neurons. Neurons were exposed to 1 nM BoNT/A1 and 0.1 nM BoNT/D in high K^+ buffer for 5 minutes at 37°C. Cells were washed three times with PBS and further

incubated in toxin-free media for 8 hrs. Neuron lysates were then harvested and subjected to immunoblot analysis, detecting cleavage of toxin substrate SNAP-25 (for BoNT/A1) and VAMP2 (for BoNT/D). Cleavage of SNAP-25 by BoNT/A1 generates a smaller fragment that can be detected on immunoblot. Cleavage of VAMP2 by BoNT/D resulted in a loss of immunoblot signal of VAMP2.

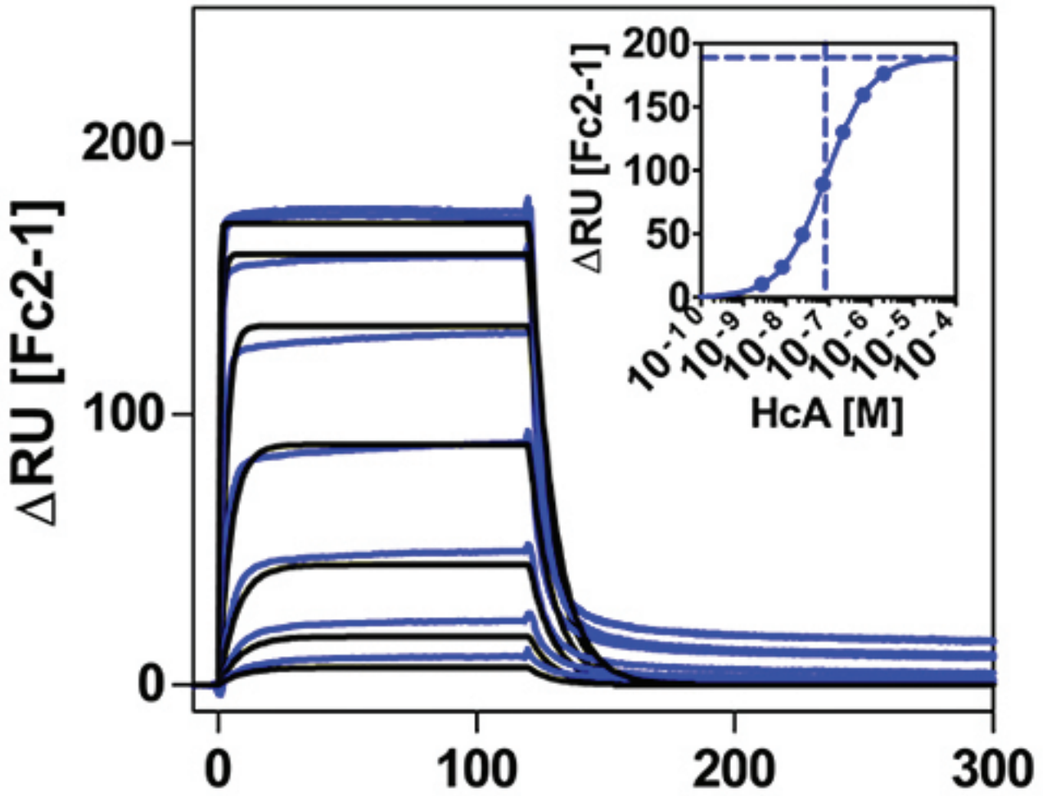
Mouse phrenic nerve hemidiaphragm (MPN) assay. The MPN assay was performed as described previously employing 20-30 g NMRI mice (Janvier SA, France) [20,21]. The phrenic nerve was continuously stimulated at 5-25 mA with a frequency of 1 Hz and with a 0.1 ms pulse duration. Isometric contractions were transformed using a force transducer and recorded with VitroDat Online software (FMI GmbH, Germany). The time required to decrease the amplitude to 50% of the starting value (paralytic half-time) was determined. To allow comparison of the altered neurotoxicity of mutants with H6tBoNTA wild-type, a power function (y(H6tBoNTA; 10, 30, 80 pM) = $139.6x^{-0.1957}$, $R^2 = 0.9991$) was fitted to a concentration-response-curve consisting of three concentrations determined minimum in technical triplicates. Resulting paralytic half-times of the H6tBoNTA mutants were converted to concentrations of the wild-type employing the above power functions and finally expressed as relative neurotoxicity.

REFERENCES

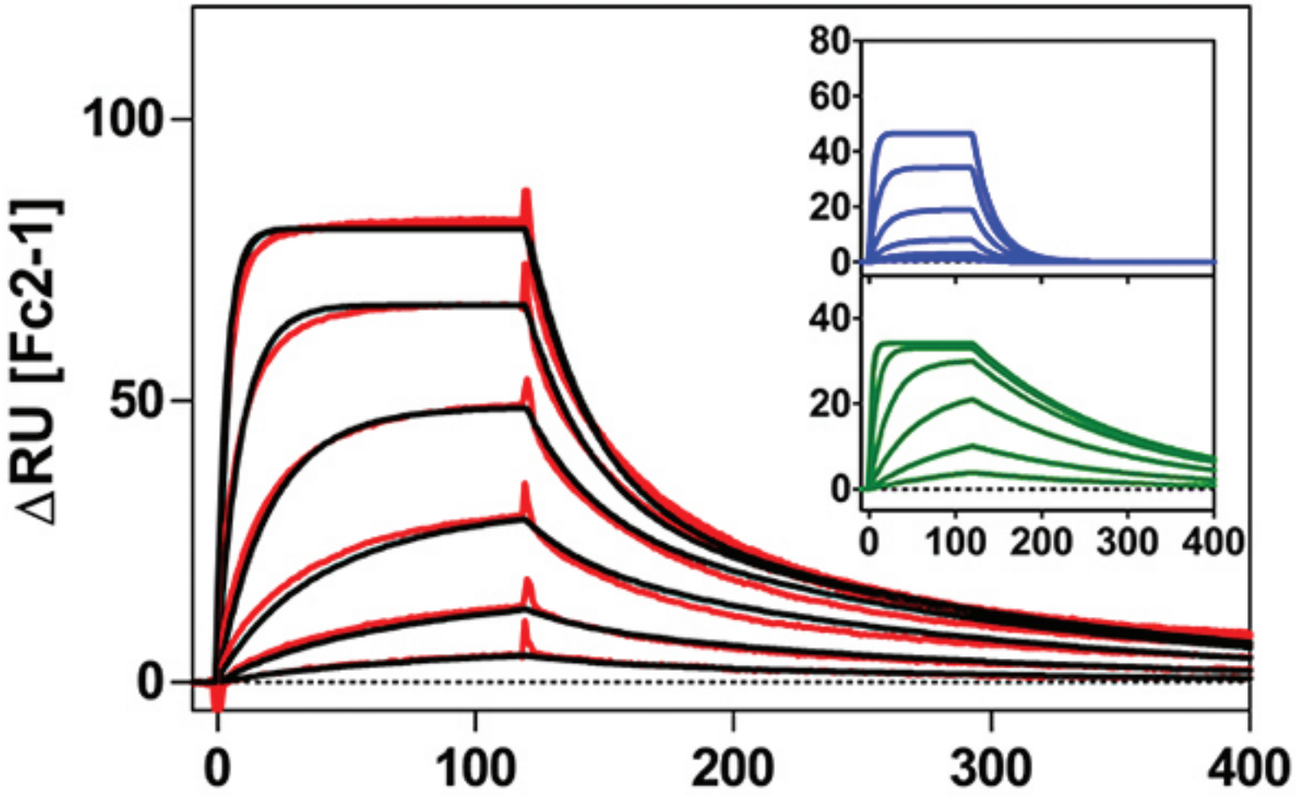
39. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
40. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Cryst* **40**, 658-74 (2007).
41. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).
42. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* **59**, 1131-7 (2003).
43. Brunger, A.T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-5 (1992).
44. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
45. Myszka, D.G. Improving biosensor analysis. *J Mol Recognit* **12**, 279-84 (1999).


a
SV2A+/+
SV2B-/-
SV2A-/-
SV2B-/-
SV2A virus
-
-
-
WT
N573A
BoNT/A/D
-
+
+
+
+
SNAP-25
*
actin
VAMP2
SV2
b
SV2A-/-
SV2B-/-
SV2C virus
-
WT
N559A
BoNT/A/D
+
+
+
SNAP-25
*
actin
VAMP2
SV2

a

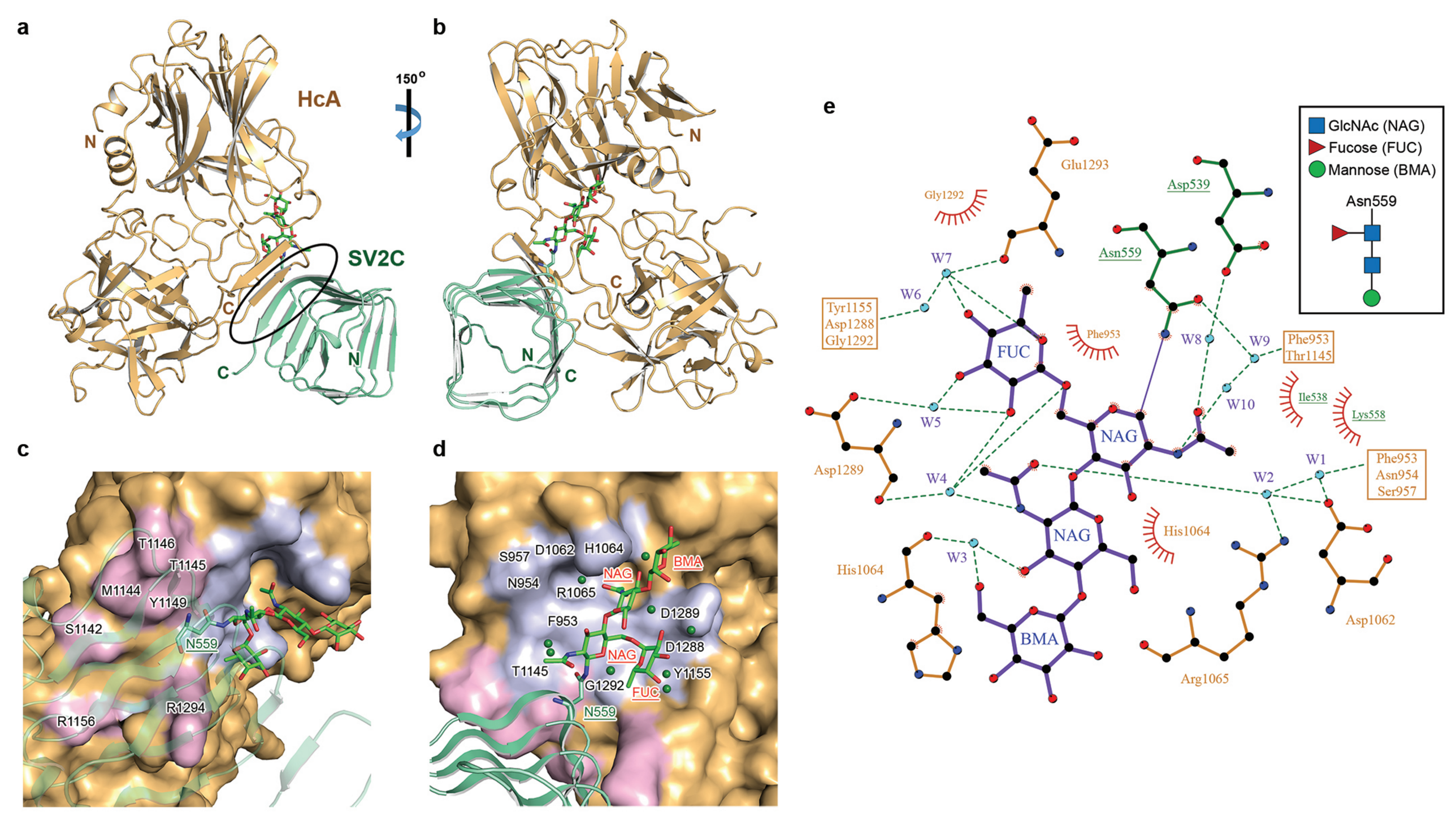


b

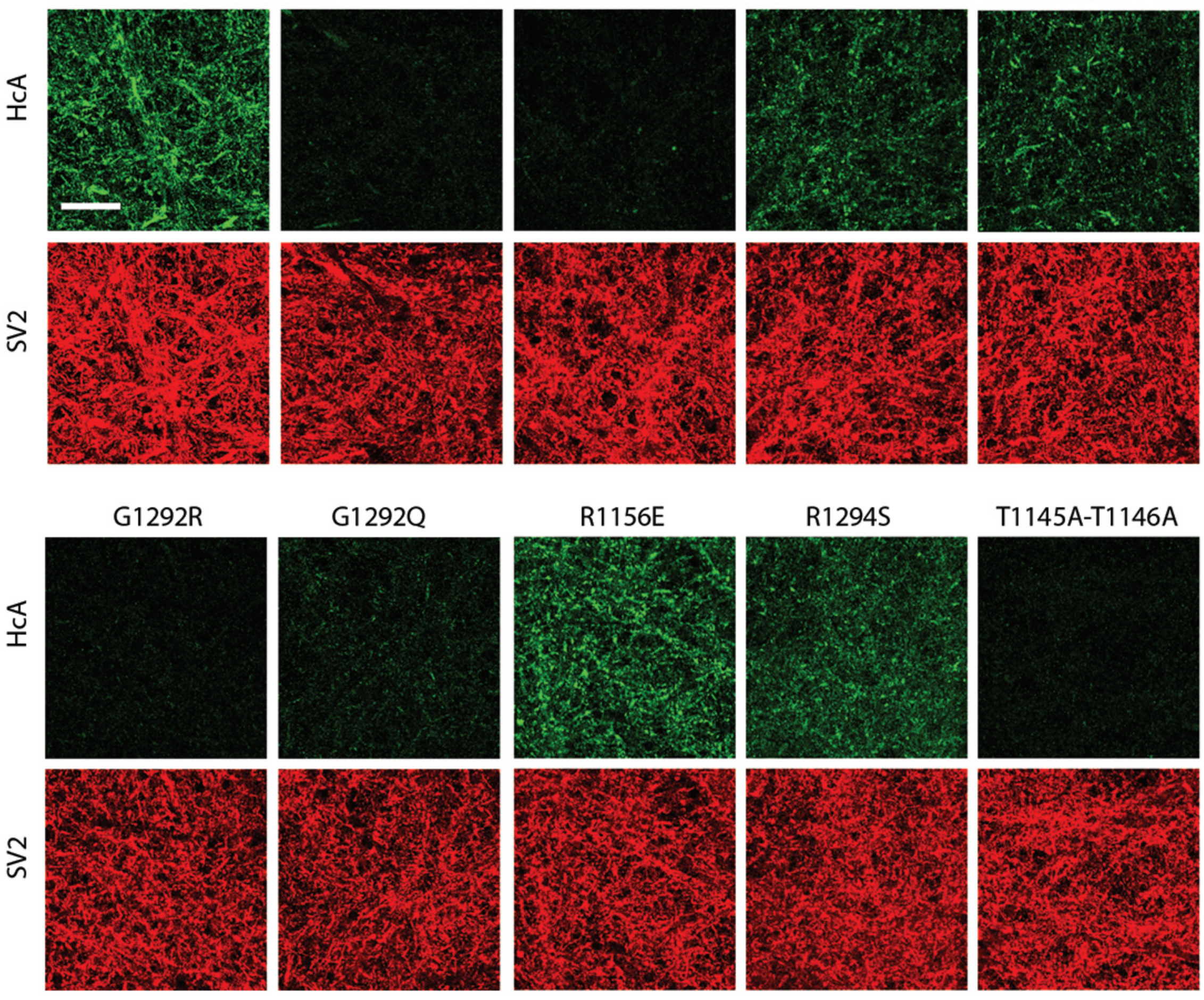


c

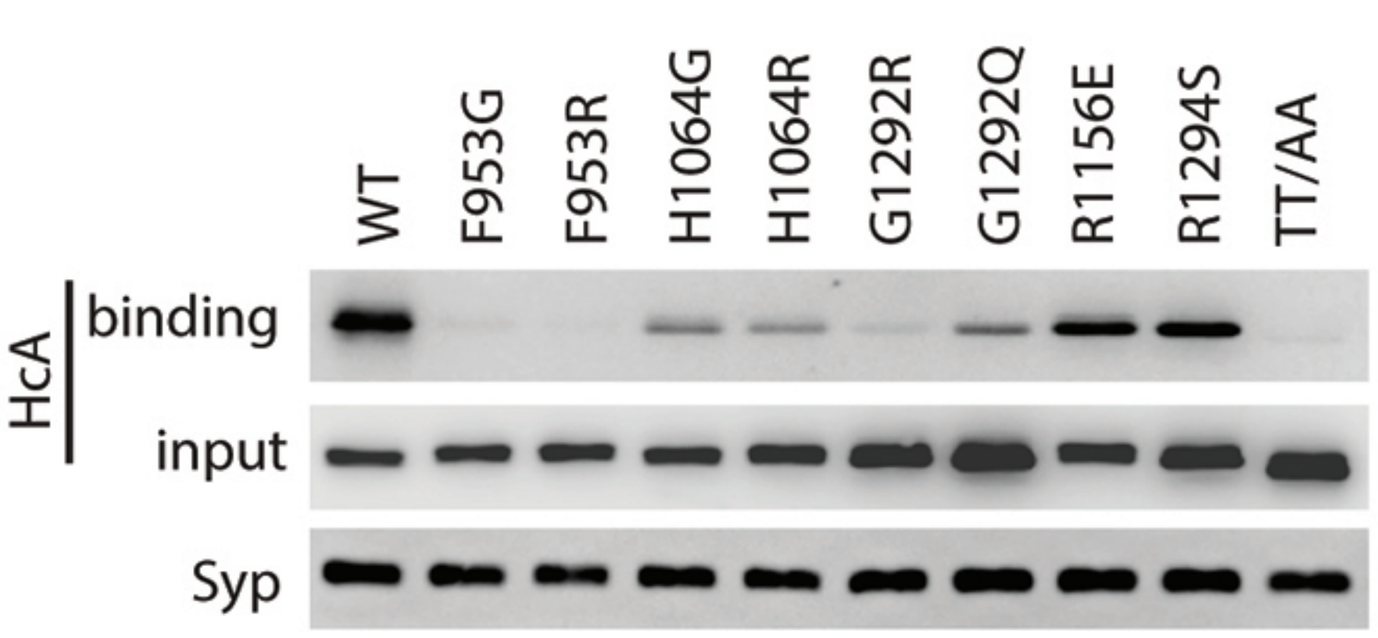
	k_a [$M^{-1}s^{-1}$]	k_d [s^{-1}]	$K_{D\ kinetics}$ [M]	$K_{D\ steady\ state}$ [M]
bSV2C	$1.5\pm0.0 \times 10^{6}$	$1.1\pm0.0 \times 10^{-1}$	$7.2\pm0.2 \times 10^{-8}$	$8.6\pm0.6 \times 10^{-8}$
gSV2C	$1.7\pm0.0 \times 10^{5}$	$3.7\pm0.0 \times 10^{-2}$	$2.2\pm0.0 \times 10^{-7}$	n.a.
	$3.4\pm0.1 \times 10^{5}$	$4.9\pm0.4 \times 10^{-3}$	$1.5\pm0.2 \times 10^{-8}$	


a
HcA
SV2C
N
C
N
C
150°
b
N
C
N
C
c
T1146
T1145
M1144
Y1149
S1142
N559
R1294
R1156
d
S957
D1062
H1064
BMA
N954
NAG
R1065
D1289
F953
D1288
NAG
T1145
Y1155
G1292
FUC
N559
e
GlcNAc (NAG)
Fucose (FUC)
Mannose (BMA)
Asn559
Glu1293
Asp539
Gly1292
Asn559
W7
W6
Tyr1155
Asp1288
Gly1292
FUC
Phe953
W8
W9
Phe953
Thr1145
Ile538
Lys558
W10
W5
NAG
Asp1289
W4
W1
W2
Phe953
Asn954
Ser957
His1064
NAG
W3
His1064
BMA
Asp1062
Arg1065

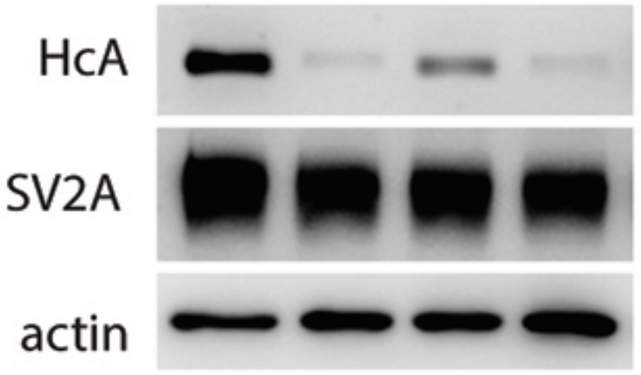
a


WT
F953G
F953R
H1064G
H1064R
HcA
SV2
G1292R
G1292Q
R1156E
R1294S
T1145A-T1146A

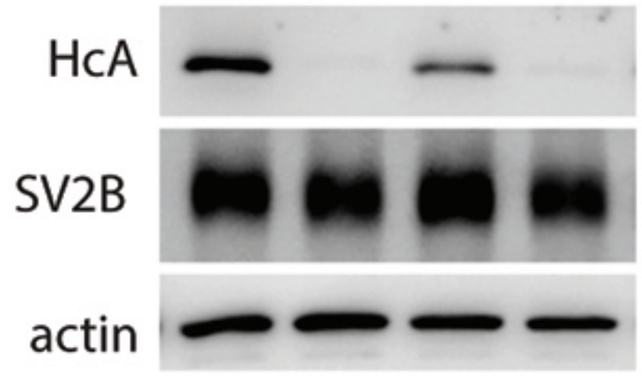
b


WT
F953G
F953R
H1064G
H1064R
G1292R
G1292Q
R1156E
R1294S
TT/AA
HcA
binding
input
Syp

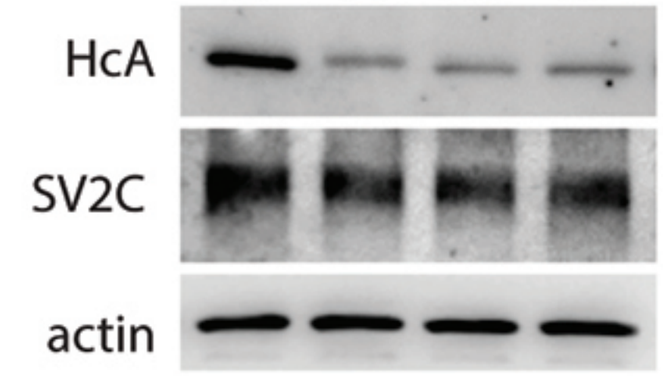
c

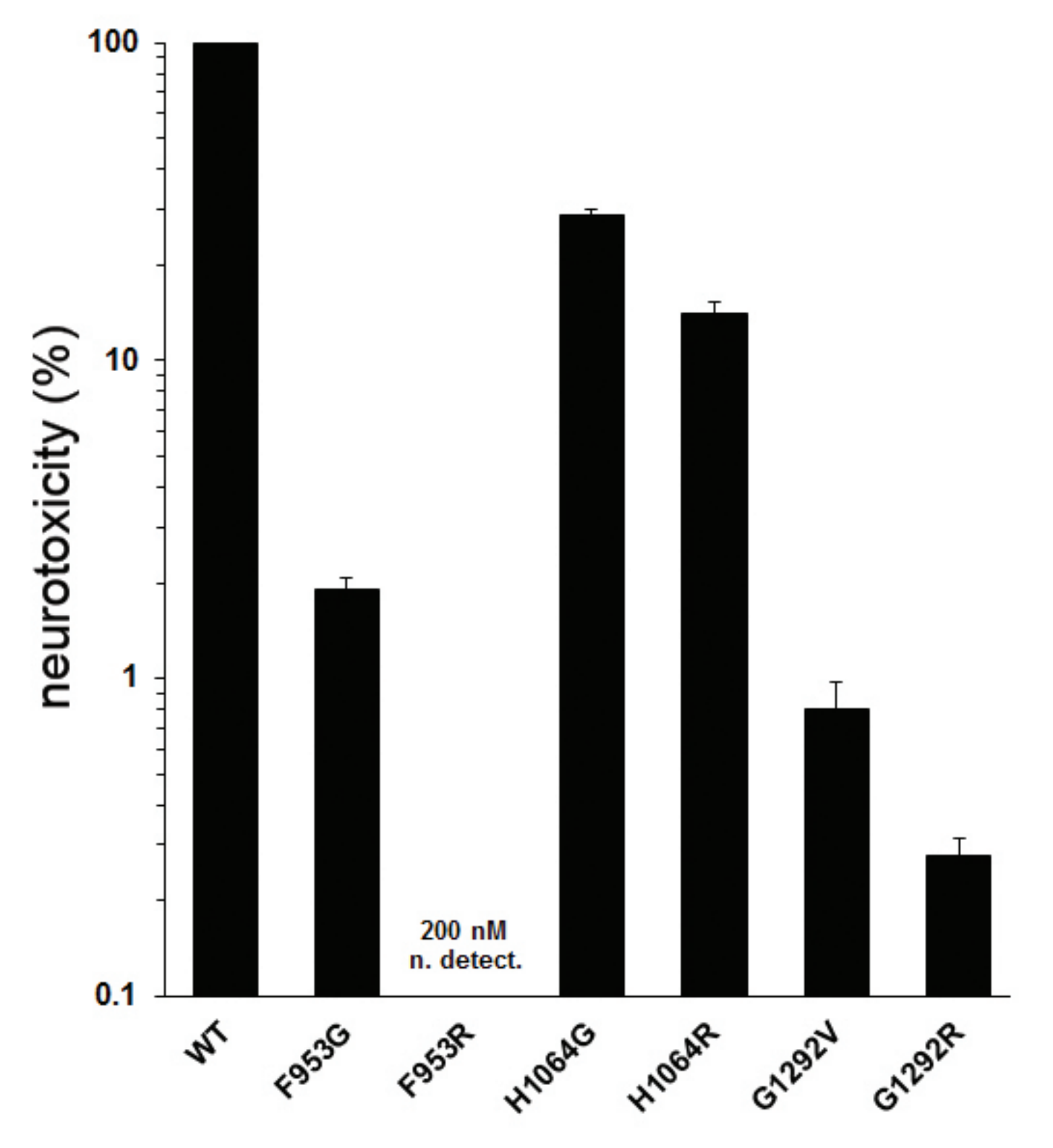

WT
F953G
H1064G
G1292R
HcA
SV2A
actin

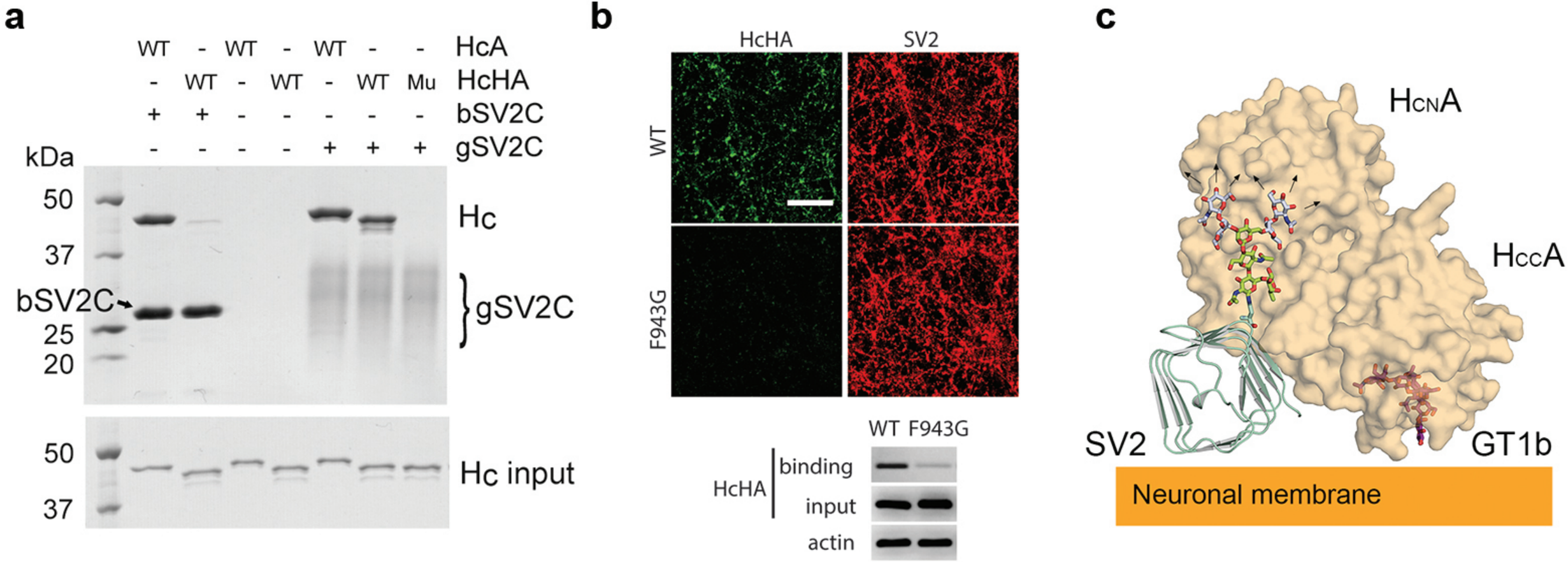
d

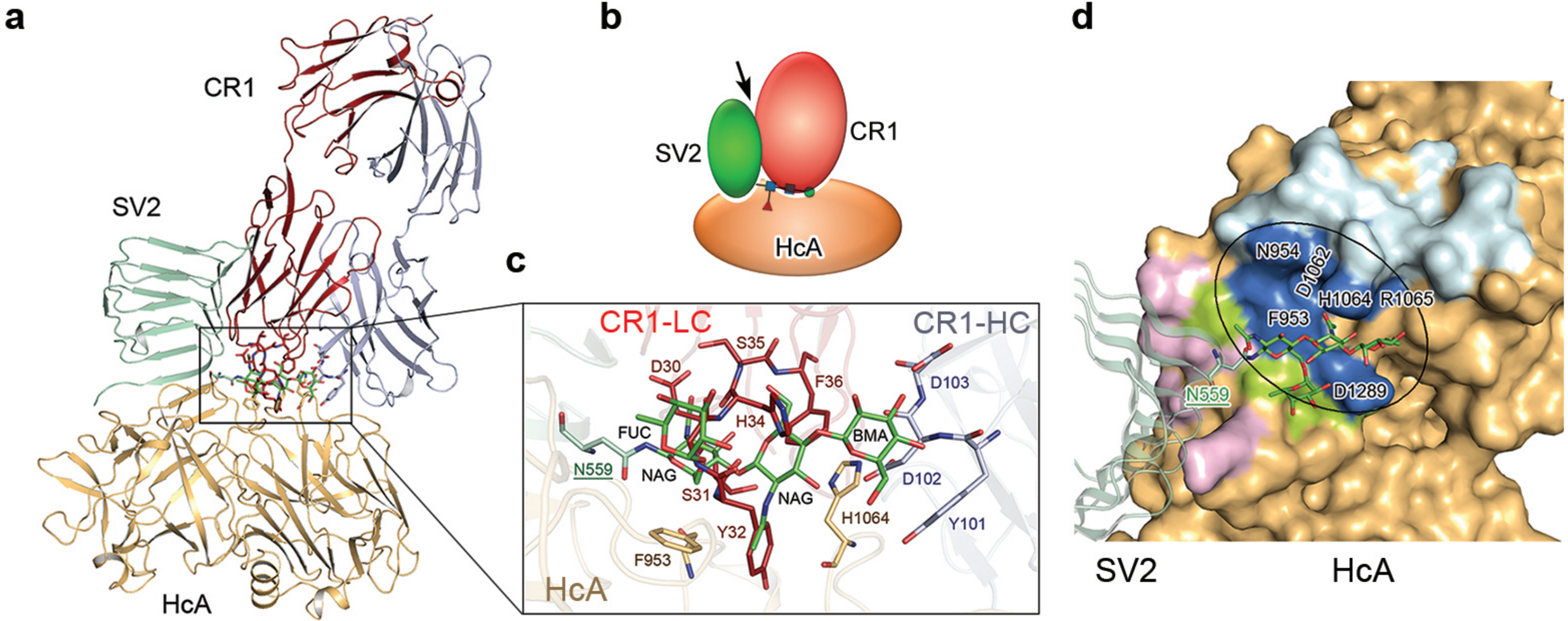

WT
F953G
H1064G
G1292R
HcA
SV2B
actin

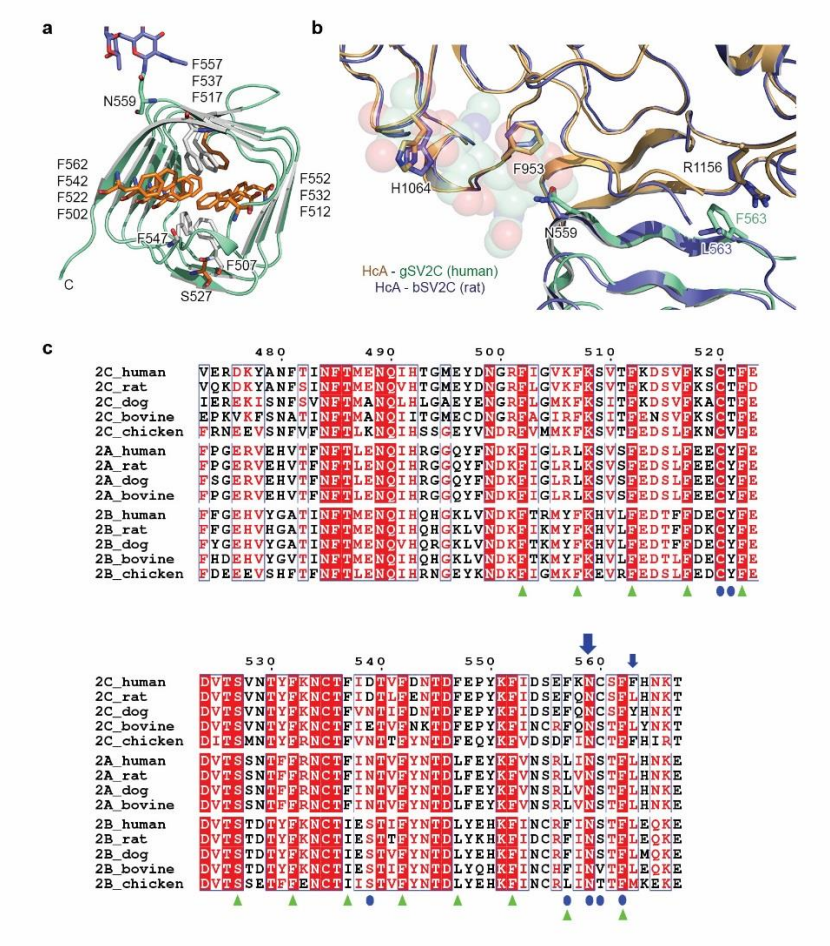
e


WT
F953G
H1064G
G1292R
HcA
SV2C
actin


100
10
1
0.1
neurotoxicity (%)
200 nM
n. detect.
WT
F953G
F953R
H1064G
H1064R
G1292V
G1292R


a
WT - WT - WT - - HcA
- WT - WT - WT Mu HcHA
+ + - - - - - bSV2C
- - - - + + + gSV2C
kDa
50
37
bSV2C
25
20
Hc
gSV2C
50
37
Hc input
b
HcHA
SV2
WT
F943G
WT F943G
HcHA
binding
input
actin
c
HcNA
HccA
SV2
GT1b
Neuronal membrane

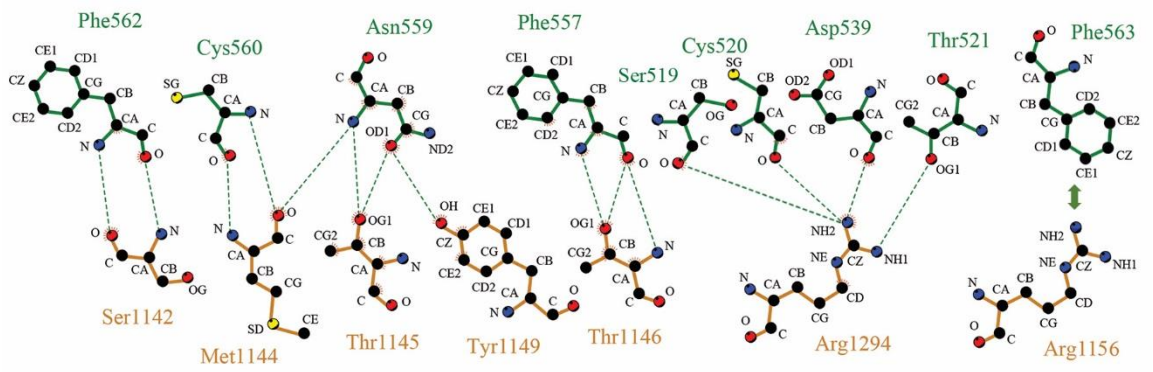

a
CR1
SV2
HcA
b
SV2
CR1
HcA
c
CR1-LC
CR1-HC
S35
D30
F36
D103
H34
FUC
BMA
N559
NAG
D102
S31
NAG
H1064
Y101
Y32
F953
HcA
d
N954
D1062
H1064
R1065
F953
N559
D1289
SV2
HcA



Supplementary Figure 1

SV2 has a unique structure.

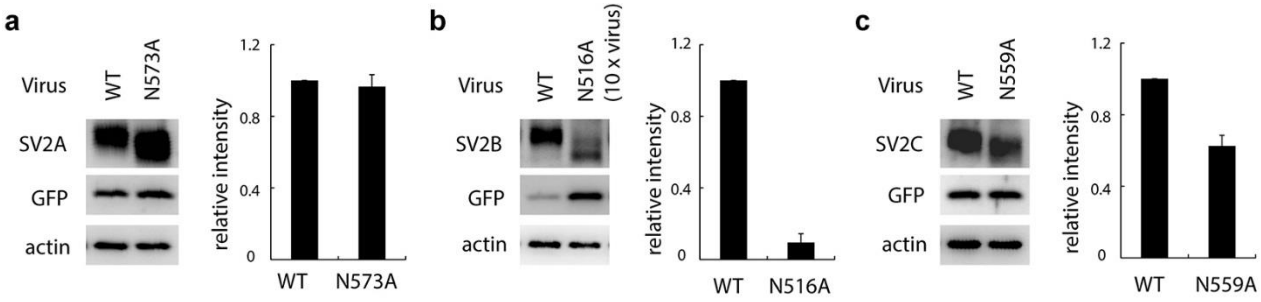
(a) The structure of human gSV2C displays a unique pentapeptide-repeat motif, where phenylalanine residues, except S527, spaced 5 residues apart provide important stacking effect to stabilize the structure. These residues are shown in sticks, with the ones that are conserved in all three SV2 isoforms across different species are colored gold. **(b)** The structures of H_CA in complex with the rat bSV2C or human gSV2C are superimposed. The N559 glycan of gSV2C is shown as a transparent sphere model. Residue F563 of human SV2C is replaced by L563 in rat SV2C, which abolishes the cation-π stacking interaction. **(c)** Representative sequences of SV2A, SV2B, and SV2C were selected from different species: human SV2A (NP_055664.3), 2B (CAG33367.1), 2C (AAI00828.1); rat SV2A (NP_476558.2), 2B (NP_476555.1), 2C (NP_113781.1); dog SV2A (XP_003639668.1), 2B (XP_005618386.1), 2C (XP_546060.2); bovine SV2A (NP_776387.1), 2B (NP_001076917.1), 2C (NP_001178948.1); chicken SV2B (XP_425081.2), 2C (XP_429151.3). Sequence alignments were made using Clustal Omega [1] and ESPript [2]. Only the truncated loop 4 of SV2 is shown for clarity. Identical residues are indicated with white letters on a red background, similar conserved residues are in red letters, varied residues are in black letters. The residues of SV2C that directly bind to H_CA are indicated by blue ovals; the pentapeptide-repeat is labeled by green triangles. The N-linked glycosylation site, N559, is highly conserved, but not F563 (blue arrows).



Supplementary Figure 2

The protein-protein interactions between H_CA and the human gSV2C.

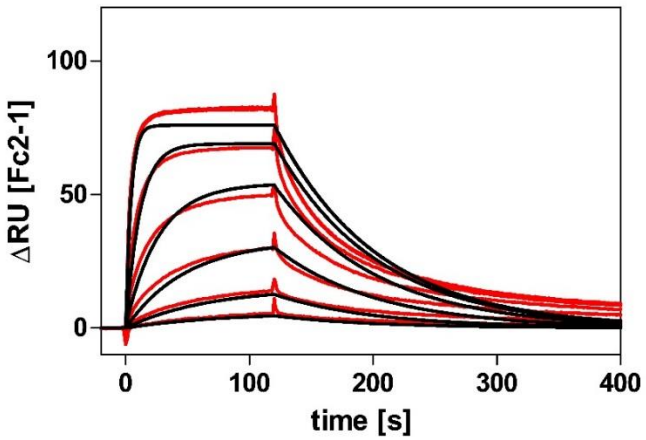
The plots were generated using LIGPLOT [3]. BoNT/A and SV2C residues are labeled brown and green, respectively. Hydrogen bonds are indicated by dashed green lines. A similar interaction network is observed in the structure of H_CA in complex with the rat bSV2C, except that the cation-π stacking interaction (double arrow) is unique for human SV2C.



Supplementary Figure 3

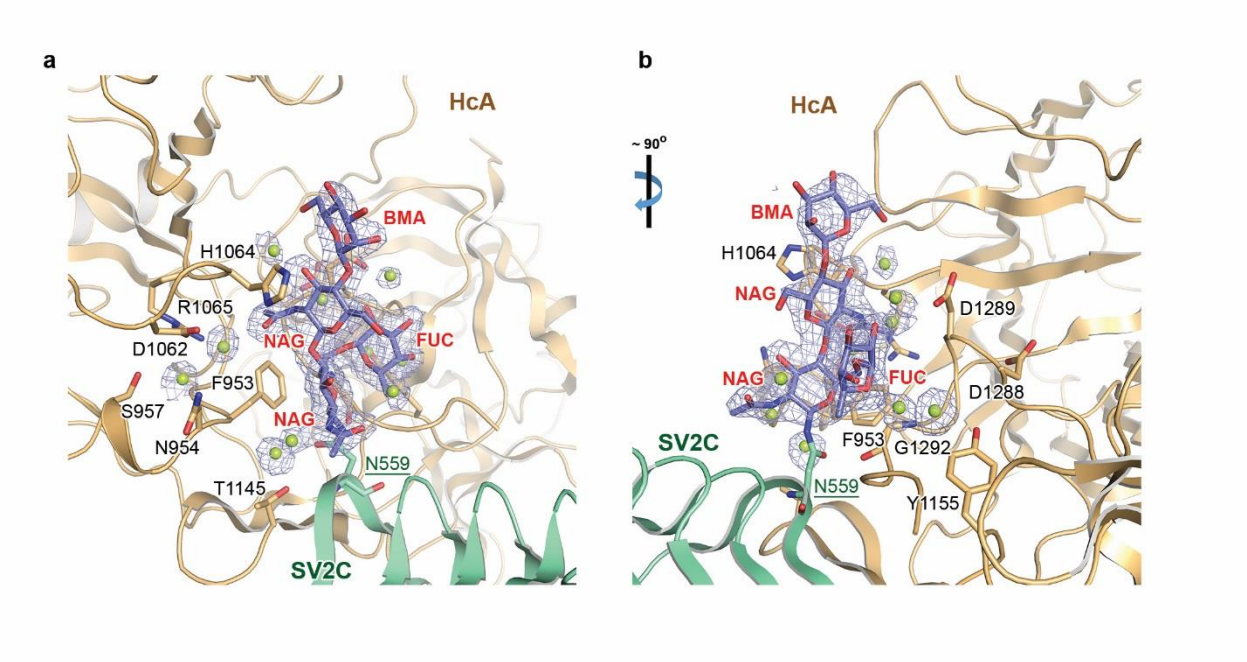
Expression levels of the deglycosylation mutants of SV2A, 2B, and 2C in neurons.

Hippocampal/cortical neurons cultured from SV2A(-/-)SV2B(-/-) mice were infected with lentiviruses that express either WT SV2A, 2B, and 2C, or indicated deglycosylation mutants. Cell lysates were harvested and subjected to immunoblot analysis. Actin served as a loading control. The lentiviral vector contains two separated synapsin promoters, with one driving expression of SV2 and the other driving expression of GFP. Thus, GFP served as an internal control for viral infection. Immunoblot signals of SV2 were quantified, normalized using GFP signals, and compared between WT and deglycosylation mutants. The same amounts of viruses were used for WT SV2A and SV2A-N573A (panel **a**), and for WT SV2C and SV2C-N559A (panel **c**). The deglycosylation mutation has no effect on SV2A and modestly reduced the expression level of SV2C in neurons. However, it severely reduced the expression level of SV2B. As shown in panel **b**, even with 10-fold more viruses, SV2B-N516A expression was still drastically lower than WT SV2B. The data are presented as mean ± S.D., n = 3.



Supplementary Figure 4

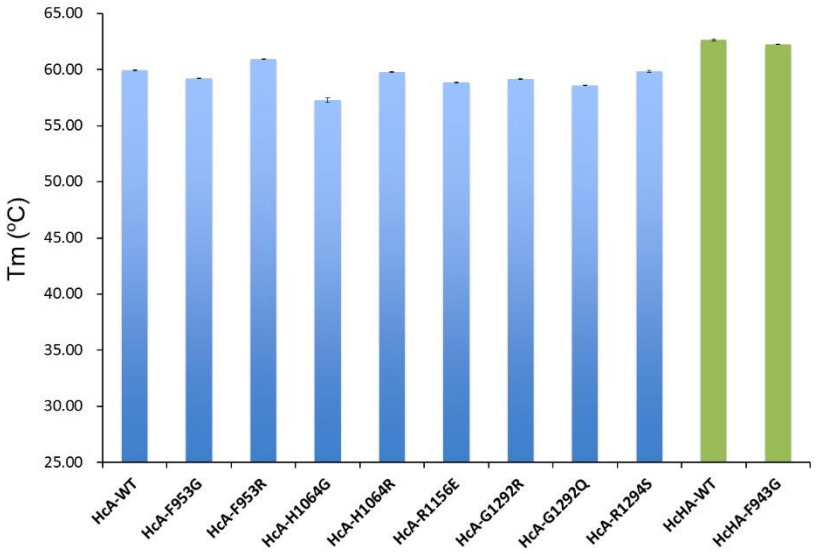
SPR binding curves of H_CA and gSV2C (red) overlaid with a 1:1 Langmuir binding model fit (black). Clear deviations were observed between the measured sensorgrams and the fit.



Supplementary Figure 5

Electron densities of the N559 glycan of SV2C in the gSV2C-H_CA complex.

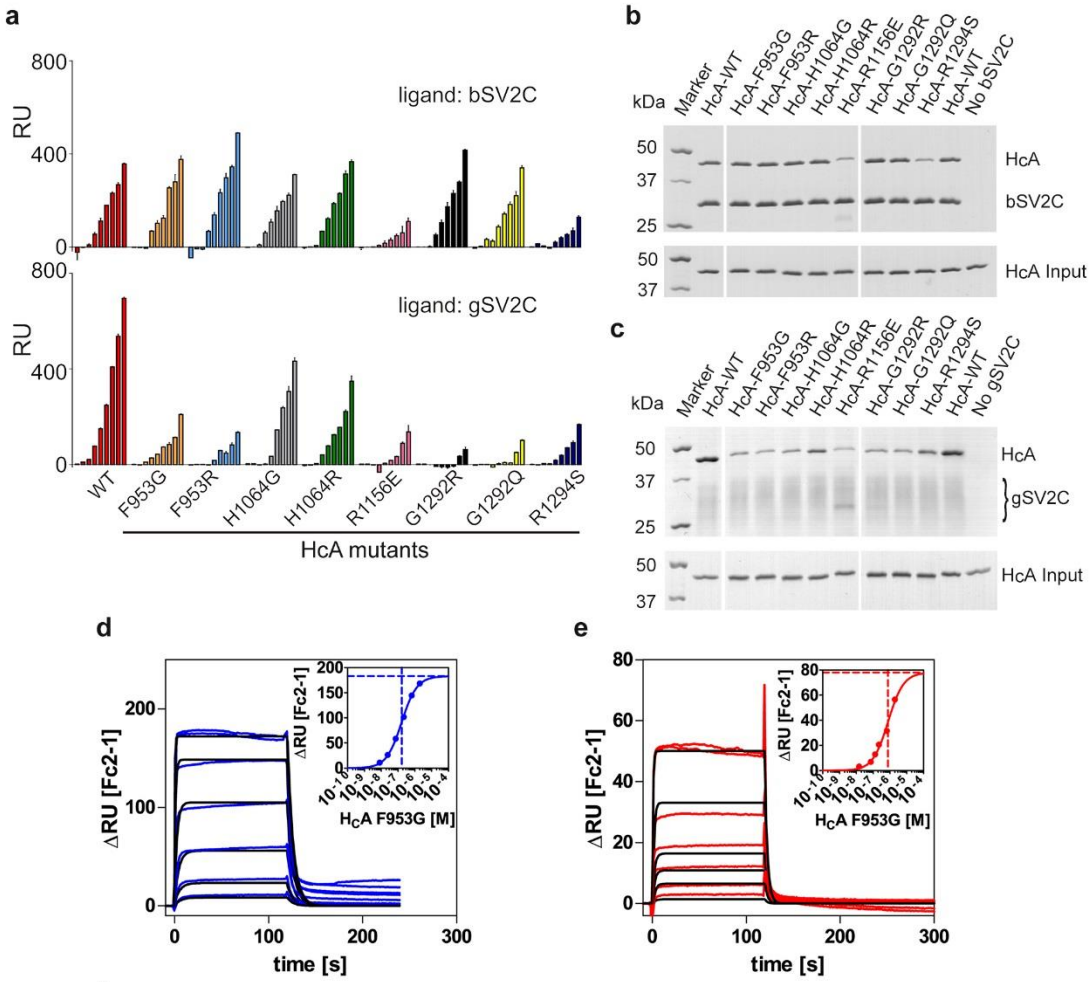
(a) Key glycan-binding residues of H_CA and the N559 glycan are shown as stick models. Water molecules facilitating the H_CA-glycan association are shown as green spheres. A simulated-annealing omit electron density map contoured at 1.5 σ was overlaid with the final refined model. **(b)** A different view with a rotation ~90° about a vertical axis.



Supplementary Figure 6

Single-site mutations of H_CA and H_CHA adopt wild-type-like structures.

The thermal stability of proteins was measured using a fluorescence-based thermal shift assay on a StepOne real-time PCR system (ThermoFisher) [4]. Specifically, protein melting was monitored using a hydrophobic dye, SYPRO Orange (Sigma-Aldrich), as the temperature was increased in a linear ramp from 20ºC to 95ºC. The midpoint of the protein-melting curve (Tm) was determined using the software provided by the instrument manufacturer. The data are presented as mean ± S.D., n = 3. All H_CA and H_CHA mutants showed Tm values comparable to the wild-type protein, indicating correct protein folding.



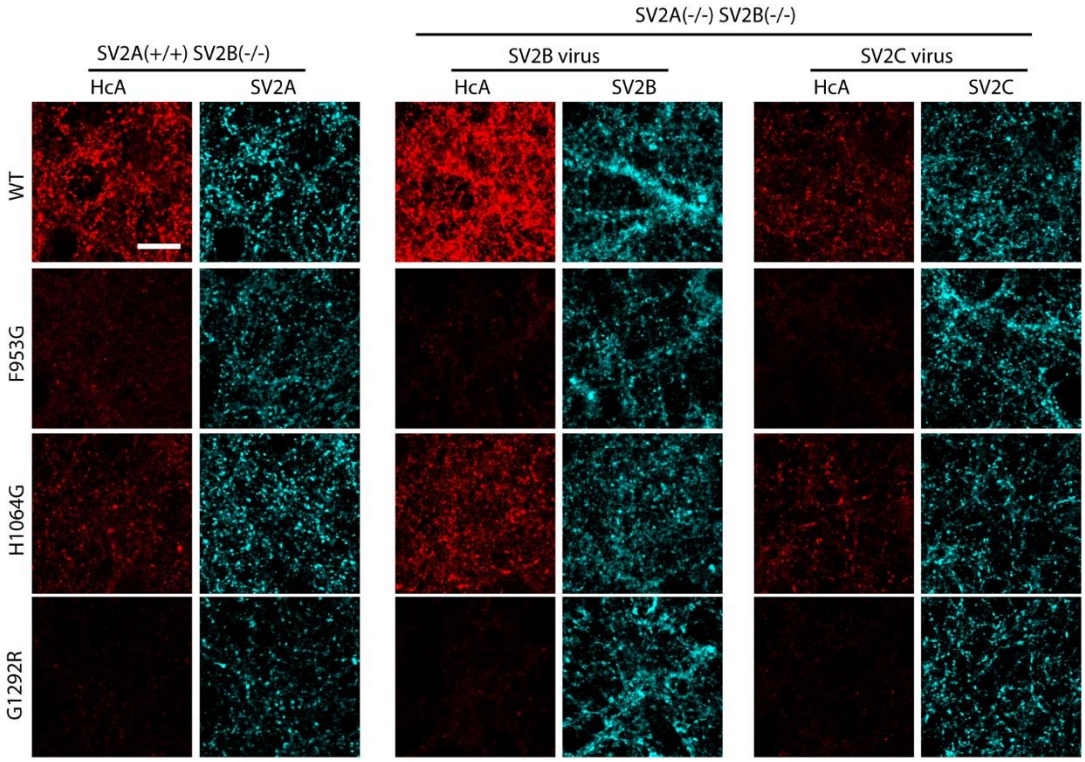
f

	k_a [$M^{-1}s^{-1}$]	k_d [s^{-1}]	K_D kinetics [M]	K_D steady state [M]
bSV2C	$9.8\pm0.4 \times 10^{5}$	$1.7\pm0.1 \times 10^{-1}$	$1.8\pm0.0 \times 10^{-7}$	$1.6\pm0.0 \times 10^{-7}$
gSV2C	$4.7\pm0.7 \times 10^{5}$	$3.3\pm0.4 \times 10^{-1}$	$7.0\pm0.2 \times 10^{-7}$	$7.6\pm0.4 \times 10^{-7}$

Supplementary Figure 7

Characterization of bindings between H_CA variants and human bSV2C and gSV2C, respectively.

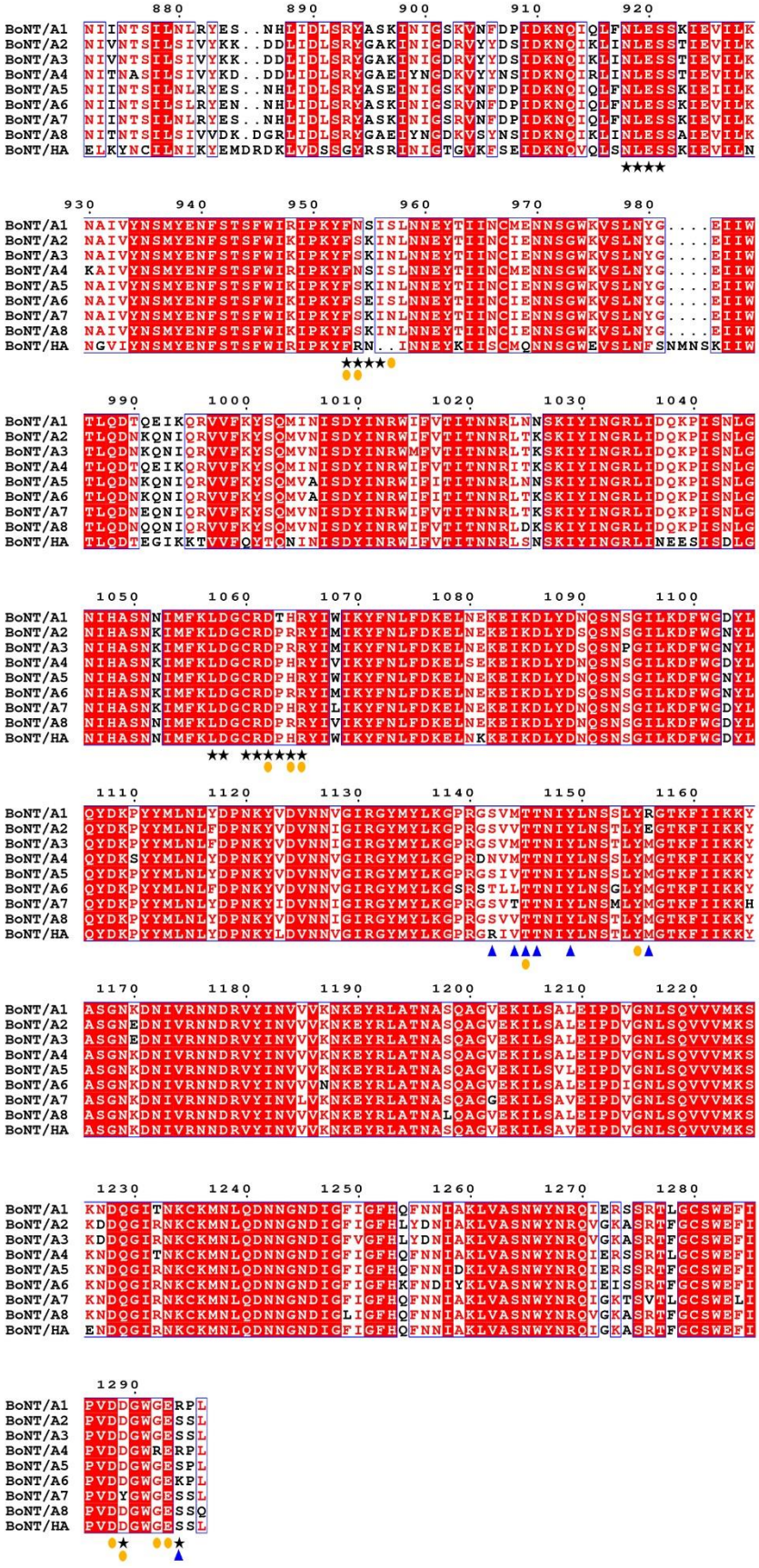
(a) Surface plasmon resonance was used to examine the changes of binding affinity between H_CA variants and SUMO-bSV2C or gSV2C, respectively. SV2C was covalently immobilized to a CM5 chip as a ligand whereas H_CA variants were analytes. Bars from left to right represent the responses when H_CA was applied at 10 pM, 1 nM, 10 nM, 25 nM, 50 nM, 75 nM, 100 nM, and 200 nM, respectively. RU stands for arbitrary response unit. **(b-c)** Interactions between H_CA variants (preys) and SUMO-bSV2C or gSV2C (baits) were examined by a pull down assay. **(d-f)** Binding kinetics and affinity between H_CA-F953G and immobilized bSV2C (107 RU; panel **d**) or gSV2C (74 RU; panel **e**) were determined by injecting 1:3 dilution series ranging from 2,000 nM to 8.23 nM. Values shown represent the mean ±S.D. (n = 2).



Supplementary Figure 8

Binding of glycan-binding deficient H_CA mutants to neurons that express individual SV2 isoforms.

Hippocampal/cortical neurons cultured from SV2A(+/+)SV2B(-/-) mice served as neurons that only express SV2A. Neurons that only express SV2B or SV2C were created by infecting neurons cultured from SV2A(-/-)SV2B(-/-) mice with lentiviruses that express SV2B or SV2C, respectively. Neurons were then exposed to WT or indicated H_CA mutant (100 nM, 5 min), washed, fixed, and subjected to immunostaining analysis. H_CA was detected with a monoclonal human anti-BoNT/A antibody (RAZ-1) and SV2 was detected with a mouse monoclonal pan-SV2 antibody. Scale bar, 20 µm.


BoNT/A1
BoNT/A2
BoNT/A3
BoNT/A4
BoNT/A5
BoNT/A6
BoNT/A7
BoNT/A8
BoNT/HA

Supplementary Figure 9

Sequence alignment among eight BoNT/A subtypes and BoNT/HA.

The amino acid sequence of BoNT/A1-A8 and HA are taken from GenBank: AAQ06331.1 (A1), ACO83782.1 (A2), ABA29017.1 (A3), ACQ51417.1 (A4), ACG50065.1 (A5), ACW83608.1 (A6), AFV13854.1 (A7), AJA05787.1 (A8), and KGO15617.1 (HA). Sequence alignments were made using Clustal Omega [1] and ESPript [2]. Identical residues are indicated with white letters on a red background, similar conserved residues are in red letters, varied residues are in black letters. Key H_CA residues that are recognized by antibody CR1 (PDB code: 2NYY) [5] are indicated by black stars. H_CA residues that directly interact with SV2C peptide or the N559-glycan are labeled by blue triangles or yellow ovals, respectively.

Supplementary Table 1

Water-mediated interactions in the SV2C glycan-H_CA interface

Water #	Distance (Å)	Water #	SV2C N559-glycan	H_CA (or SV2C)
W1	3.15	W2		
	3.04			F953 [N]
	3.04			N954 [N]
	2.79			S957 [OG]
	2.97			D1062 [OD2]
W2	2.90		NAG603 [O7]	
	2.67			D1062 [OD2]
	2.76			R1065 [NH2]
W3	2.62		NAG603 [O3]	
	3.15		BMA604 [O6]	
	2.77			H1064 [O]
W4	2.84		FUC601 [O2]	
	3.39		NAG602 [O6]	
	2.70		NAG603 [N2]	
	2.66			D1289 [O]
W5	2.61		FUC601 [O2]	
	3.40		FUC601 [O3]	
	2.69			D1289 [OD1]
W6	2.79	W7		
	2.66			Y1155 [OH]
	2.84			D1288 [O]
	3.02			G1292 [N]
W7	2.84		FUC601 [O4]	
	3.38		FUC601 [O5]	
	2.60			E1293 [O]
W8	3.32		NAG602 [O7]	
	2.80			D539 [OD2] – SV2C
W9	2.95	W10		
	2.96			F953 [O]
	2.67			T1145 [OG1]
	3.12			N559 [OD1] – SV2C
W10	2.36		NAG602 [N2]	

REFERENCES

1. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
2. Gouet, P., Courcelle, E., Stuart, D.I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8 (1999).
3. Laskowski, R.A. & Swindells, M.B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-86 (2011).
4. Lo, M.C. et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **332**, 153-9 (2004).
5. Garcia-Rodriguez, C. et al. Molecular evolution of antibody cross-reactivity for two subtypes of type A botulinum neurotoxin. *Nat Biotechnol* **25**, 107-16 (2007).