Synthetic single domain antibodies for the conformational trapping of membrane proteins

Iwan Zimmermann# [1], Pascal Egloff# [1], Cedric A. Hutter# [1], Peter Stohler [2], Nicolas Bocquet [2], Melanie Hug [2], Martin Siegrist [2], Lisa Svacha [2], Jennifer Gera [2], Samira Gmür [3], Peter Spies [3], Daniel Gygax [3], Eric R. Geertsma* [4], Roger J.P. Dawson* [2], Markus A. Seeger* [1]

1) Institute of Medical Microbiology, University of Zurich, Zurich, Switzerland

2) Roche Pharma Research and Early Development, Therapeutic Modalities, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland

3) University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland

4) Institute of Biochemistry, Goethe University Frankfurt, Frankfurt am Main, Germany

# These authors contributed equally

* Corresponding authors

ABSTRACT (150 words)

Single domain antibodies called nanobodies are excellent affinity reagents for membrane proteins. However, their generation relies on immunizations, which is only amenable to robust proteins and impedes selections in the presence of non-covalent or toxic ligands. To overcome these key limitations, we developed a novel *in vitro* selection platform, which builds on synthetic nanobodies called sybodies. Inspired by the shape diversity of natural nanobodies, three sybody libraries exhibiting different randomized surface shapes were engineered for high thermal stability. Using ribosome display, exceptionally large libraries were pre-enriched against membrane protein targets and subsequently funneled into a robust phage display process, thereby reducing selection bias. We successfully generated conformation-selective, high affinity sybodies against the human glycine transporter GlyT1, the human equilibrative nucleotide transporter ENT1 and a bacterial ABC transporter. Our platform builds exclusively on commercially available reagents and enables non-specialized labs to generate conformation-specific binders against previously intractable protein targets.

1

INTRODUCTION

Conformation-specific binders raised against membrane proteins have the ability to manipulate cells directly at the cell surface and are exquisite tools for basic science and drug discovery[1-3]. However, binder selection against this difficult class of target proteins is challenging[4-6]. A particularly successful method relies on the immunization of camelids for the pre-enrichment of B-cells that encode target-specific nanobodies, the variable domains of heavy-chain-only antibodies[6]. The extraordinary success of nanobodies is rooted in the simplicity and robustness of the single chain nanobody scaffold and in its characteristic variability at the complementarity determining region 3 (CDR3), which is frequently found to penetrate deeply into cavities of membrane protein targets[7,8]. CDR3 loops of variable length and orientation create diverse binder shapes, thus permitting an optimal surface-complementarity to antigens.

Despite of the outstanding track record of camelid nanobodies, there are two fundamental drawbacks linked to immunizations. First, the target space is limited to non-toxic, non-conserved and stable proteins. Delicate targets (e.g. human membrane transporters) readily unfold upon injection into the blood stream due to the applied adjuvants and the camelid's high body temperature. Second, it is very difficult to favor target conformations with non-covalent ligands because they dissociate from the protein shortly after injection.

Pure *in vitro* binder selection methods promise to solve the drawbacks of immunization, but often suffer from i) limited library sizes (e.g. <10^9 for phage or yeast display), resulting in weak binding affinities ii) poor framework design, giving rise to a large fraction of aggregated or poorly expressing library members and iii) selection bias, leading to poorly enriched pools[9]. The consequence of these shortcomings is the requirement for extensive single-clone screening efforts after selection.

In this work, we introduce a selection platform, tailored for membrane protein targets, which overcomes the current limitations of immunizations and *in vitro* selections[10,11]. At the core of our technology are synthetic and highly stable scaffolds called sybodies that were inspired by camelid nanobodies and their natural shape diversity to achieve an optimal surface complementarity for various epitope shapes. To compensate for the incremental antibody maturation taking place *in vivo*, we constructed very diverse sybody libraries and harnessed the capacity of ribosomes to display exceptionally large library sizes. Our approach permits the selection and preparative production of sybodies within three weeks and requires only standard laboratory materials. The platform generated conformation-selective sybodies against two disease-relevant human SLC transporters in their inhibitor-locked states and by trapping an ABC transporter in its transient ATP-bound conformation.

2

RESULTS

Design and construction of three synthetic sybody libraries

Three nanobody structures representing different binding modes served as templates for library design. First, a GFP-binder with a short CDR3 interacts via a **concave** surface with its target (PDB: 3K1K)[12]. Second, a β_2-adrenergic receptor binder inserts a medium length CDR3 as an extended **loop** into a receptor cavity (PDB: 3P0G)[7]. And third, a lysozyme binder displays a long CDR3, which is tethered down via an extended hydrophobic core, and interacts with its target lysozyme via a **convex** surface (PDB: 1ZVH)[13] (Fig. 1, Supplementary Fig. 1, Supplementary Table 1). The resulting sybody libraries were therefore dubbed **"concave"**, **"loop"** and **"convex"**, accordingly.

The concave and loop library share the same scaffolds, while the convex library has a different scaffold containing an extended hydrophobic core (Supplementary Fig. 2a and b). A single, conserved disulfide bond at the center of the immunoglobulin domain is common to all three scaffolds. To characterize the scaffolds, non-randomized sybodies representing the three libraries were generated by gene synthesis. The corresponding purified proteins eluted as a single species on size exclusion chromatography (Supplementary Fig. 3a). Scaffold engineering resulted in high melting temperature of 74 °C, 75 °C and 95 °C for the concave, loop and convex scaffold, respectively, which corresponds to a stability increase of 21 °C to 35 °C over their natural precursors (Supplementary Fig. 3b). Based on these scaffolds, the three sybody libraries were constructed by randomizing all three CDRs using defined mixtures of trinucleotides, thereby obtaining a balance between charged, polar, aromatic and apolar amino acids tailored for selections against membrane proteins[14] (Supplementary Note 1 and Supplementary Fig. 2c). The three sybody libraries were flanked with the required sequence elements for *in vitro* transcription and ribosome display, each exhibiting an experimental diversity of 9×10^{12} members (Fig. 1, Supplementary Table 2).

Sybody selections using ribosome and phage display

To enable display of $> 10^{12}$ different library members, sybody selections start with ribosome display whereby a stable ternary complex between an encoding mRNA, the ribosome and the folded nascent polypeptide chain is formed. The resulting genotype-phenotype linkage is established purely *in vitro*. This stands in contrast to technologies such as phage or yeast display, which require transformation steps into cells, thus limiting the effective library sizes significantly to approximately 10^9. To circumvent the tedious in-house preparation of ribosomal extracts for ribosome display[15], we assessed the capacity of the commercially available *in vitro* translation kit PUREfrex®SS

(GeneFrontier) to display functional nanobodies. We demonstrated that a very large fraction (84.6 ± 3.5 %) of the input mRNA encoding for a GFP binder could be displayed and recovered using this kit (Supplementary Fig. 4, Supplementary Note 2). Test selections with three consecutive rounds of ribosome display were carried out against maltose binding protein (MBP) by immobilizing biotinylated MBP on magnetic beads, followed by panning with the displayed sybodies (Supplementary Fig. 5). Sybody pools were found to be strongly enriched after the third selection round as monitored by qPCR. Using FX cloning[16], single sybodies were introduced into expression vector pNb_init, which directs the protein into the periplasm by virtue of a PelB leader sequence and adds a Myc- and a His-tag to the C-terminus of the sybody for detection by ELISA (Fig. 1). Of note, pNb_init contains BspQI restriction sites, which permits to release sybodies for subcloning into expression plasmids pBXNPH3 and pBXNPMH3 for the production of tag-less binders (Supplementary Fig. 6). ELISA analysis of the selections of the concave, loop and convex library revealed about 20 %, 50 % and 30 % of all wells as strong and specific hits against MBP. Crystal structures of three convex MBP binders (Sb_MBP#1-3) in complex with MBP were solved at resolutions ranging from 1.4 - 1.9 Å (Supplementary Table 3). All crystallized sybodies were highly similar to their natural precursor (e.g. RMSD of 1.02 Å comparing Sb_MBP#1 and 1ZVH), thereby validating our library design, which kept selected residues of the CDRs constant to assure folding of an extended hydrophobic core (Supplementary Fig. 7). Sb_MBP#1-3 have binding affinities ranging from 24-500 nM as determined by surface plasmon resonance (SPR). They bind into the cleft between the two lobes of MBP and thereby trap the target in its ligand-free conformation (Fig. 2a, Supplementary Figs. 8 and 9)[17]. In support of this notion, SPR measurements revealed decreasing sybody binding affinities at increasing maltose concentrations. A Schild plot analysis revealed 58 fold decreased sybody binding affinity at saturating maltose concentration and an affinity of 1.0 µM of MBP for maltose (Fig. 2b), which is in close agreement with the literature[18]. An analysis of the binding interface highlighted CDR3 residues W101, Q104, S105 and W110, which are identical among the three binders (Supplementary Fig. 8, Supplementary Note 3). Further contacts are mediated by variable randomized residues of CDR1, CDR2 and CDR3 as well as several invariant framework residues. Overall, the sybodies bind to MBP in an analogous fashion as natural nanobodies, namely via interactions predominantly mediated by CDR3 residues[13,19].

Overcoming selection bias

Three consecutive rounds of ribosome display using the same type of beads for immobilization – as successfully applied to raise binders against MBP – were insufficient to obtain sybodies against membrane proteins. In an empiric process monitoring numerous test selections by qPCR, we evolved

a strategy to overcome selection bias (Fig. 1). At the level of sybody display, the output from the initial ribosome display selection round, yielding 10^6 - 5×10^6 different sybodies as determined by qPCR, is used to generate a focused phage display library with a size range of 10^7- 5×10^7. To avoid background enrichment, solution panning was performed and surfaces for subsequent target immobilization were alternated. These simple measures introduce maximal changes in each selection round and proved highly effective in enriching sybodies against challenging membrane proteins.

Conformational trapping of a bacterial ABC transporter

ABC transporters harness the energy of ATP binding and hydrolysis to transport substrates against a concentration gradient. In the absence of ATP, the bacterial ABC transporter TM287/288 almost exclusively adopts an inward-facing (IF) state, and two crystal structures were solved in this conformation[20,21]. In contrast, a structure of outward-facing (OF) TM287/288 is still missing due to difficulties in stabilizing this alternate conformation. The transition to the OF state requires ATP binding (Fig. 3a)[22], but ATP hydrolysis constantly reverts the transporter back to its IF state. In order to populate the OF state, a glutamate to alanine substitution (E517A) in the ATP-binding cassette was introduced, which blocks ATP hydrolysis without impairing ATP binding[22]. Using the sybody selection platform (Fig. 1), binders were selected *in vitro* against TM287/288(E517A) in the presence of ATP (Fig. 1, Fig. 3). Using qPCR and AcrB as background control[14], we observed strong sybody enrichment of 170, 220 and 25 fold for the concave, loop and convex library, respectively, after the second round of phage display. For each library, 190 clones were analyzed for binding against TM287/288(E517A) in the presence of ATP by ELISA, of which 60 % were ELISA positive. Of 48 sequenced ELISA hits, 40 were unique, indicating high diversity after binder selection (Supplementary Figs. 10-13). The unique sybodies were named Sb_TM#1-40, and 37 thereof could be purified at sufficient yields and quality for further analyses. State specificity of these binders was assessed by SPR using wildtype and E517A mutant of TM287/288 as ligands and the sybodies as analytes in the presence and absence of ATP (Fig. 3b). SPR data of 31 binders could be quantified and revealed 20 sybodies specific for OF TM287/288(E517A) as defined by an affinity increase of at least three-fold upon addition of ATP (Supplementary Fig. 10). The sybodies Sb_TM#26 and Sb_TM#35 exhibited a particularly strong state-specificity and were analyzed for their capacity to inhibit ATP hydrolysis of TM287/288 (Fig. 3c). The IC_{50}s of ATP hydrolysis were found to be 62 nM and 65 nM for Sb_TM#26 and Sb_TM#35, respectively, while the corresponding inhibition curves saturated at residual ATP activities of 17.2 % and 3.2 %. In summary, sybody selections against OF TM287/288 resulted in a large number of high affinity binders trapping the transporter in its ATP-bound state and thereby strongly inhibiting

ATPase activities. Sb_TM#35 was successfully used to determine the first outward-facing crystal structure of TM287/288 at high resolution (publication in preparation).

Conformation-specific stabilization of the human SLC transporters GlyT1 and ENT1

There are only a small number of approved drugs or drugs in development, which therapeutically target human SLC transporters, indicating untapped potential[23]. A main reason behind these shortcomings is the intricate architecture and low thermal stability that makes human SLC transporters notoriously difficult to work with in early drug discovery stages[24]. Here we focus on two transporters with a high need for conformation-specific binders, namely the equilibrative nucleoside transporter 1 (ENT1) that is involved in ischemia and acts as a biomarker in pancreatic cancer[25] as well as on the glycine transporter 1 (GlyT1) that plays an important role in diseases of the central and peripheral nervous system[26] (Fig. 4a, Fig. 5a). Multiple attempts to raise mouse antibodies or nanobodies against these targets by immunizations failed in our hands, presumably due to low thermal target stability and the limited number of accessible epitopes.

In order to obtain conformation-specific binders against ENT1 and GlyT1, we performed selections at 4 °C in the presence of the inhibitors S-(4-Nitrobenzyl)-6-thioinosine (NBTI) and a Bitopertin-like molecule Cmpd1, respectively. For ENT1, the concave but not the convex and loop shape library was enriched 4-fold over background after binder selection. One concave sybody called Sb_ENT1#1 was identified by ELISA, purified as monodisperse protein. SPR measurements revealed an affinity of 40 nM (Fig. 4b, Table 1). To further characterize Sb_ENT1#1, a thermal shift scintillation proximity assay (SPA-TS) was established. ENT1 and sybody were incubated at varying temperatures in the absence of inhibitor, followed by measuring binding of tritiated NBTI. Sb_ENT1#1 binding led to a sharper transition trajectory and an increase in melting temperature by 6.1 °C (Fig. 4c), indicating that sybody binding increases the population of the inhibited conformation of ENT1. Supporting this notion, the absolute binding signal for NBTI was increased by more than seven-fold in the presence of Sb_ENT1#1 at temperatures well below T_m (Fig. 4d).

For GlyT1, seven sybodies from the concave (Sb_GlyT1#1-3) and loop (Sb_GlyT1#4-7) library were identified by ELISA and complex formation was confirmed by size-exclusion chromatography (SEC) (Fig. 5b). Binding kinetics were determined by SPR revealing a wide range of affinities from 494 pM to 2.52 μM (Fig. 5c, d, Table 1). Competition binding SPR analysis using GlyT1 pre-saturated with Sb_GlyT1#6 revealed binding of Sb_GlyT1#1-4 but not of Sb_GlyT1#5 and Sb_GlyT1#7 (Fig. 5e), demonstrating that at least two non-overlapping epitopes on GlyT1 are recognized. The large differences in affinities correlated well with SPA-TS analysis using the commercially available tritiated

inhibitor Org24598 that addresses the same binding site as Cmpd1 [27]. Of the seven sybodies, Sb_GlyT1#1-5 increased the T_m by 0.9-2.6 °C, whereas Sb_GlyT1#6 and Sb_GlyT1#7 stabilized the transporter by 8.8°C and 10 °C, respectively (Fig. 5f). Absolute SPA binding signals obtained at 23 °C increased up to 1.8-fold (Sb_GlyT1#7), suggesting that all sybodies stabilize the inhibited conformation of GlyT1 (Fig. 5g).

In conclusion, the identified sybodies trap inhibited conformations of ENT1 and GlyT1 thereby enhancing ligand binding. Hence, sybodies increase assay sensitivity for inhibitor screening and can serve as crystallization chaperones for structure-based drug design.

DISCUSSION

Pure *in vitro* binder selections operate independent of target toxicity and sequence conservation, allow for a wide range of selection conditions including low-affine ligands to trap targets in desired conformations and are also significantly faster than selections depending on a slow immune response. Despite of these advantages over *in vivo* procedures, successful *in vitro* selections against membrane proteins are limited to a few specialized labs[5,28-33].

Libraries based on synthetic scaffolds usually exhibit a single shape and are randomized at only one region of their surface. Since a large fraction of the membrane protein surface is buried beneath lipids or detergents, suboptimal shape-complementarity between its few accessible epitopes and the randomized binder surface is a key limiting factor that can impede successful selections. To overcome this common shortcoming, we engineered three synthetic single domain antibody libraries of different shapes. Thereby, we created a large paratope space, which is a key feature of the sybody platform to target membrane proteins with a limited number of suitable epitopes.

Crystal structures revealed that each nanobody contains a dedicated set of aromatic or aliphatic CDR residues, which point towards the hydrophobic core and thereby contribute to scaffold stability. Importantly, scaffolding CDR residues are harmonized among CDRs of the same nanobody (see for example 1ZVH and convex library), but vary among different natural nanobodies. Consequently, CDRs of natural nanobodies cannot be exchanged without the risk of destabilizing the scaffold. We took this into account by engineering the three sybody scaffolds based on individual nanobody structures and thereby achieved high thermal stability of the sybodies.

Besides the favorable biophysical library properties, large binder diversities are very critical for *in vitro* selections, in order to compensate for affinity maturation taking place in animals as a result of

somatic hypermutation. In this work, we show that nanobodies and sybodies can be efficiently displayed on ribosomes using a commercial kit. Thereby, 10^{12} binder candidates can be efficiently displayed, whereas phage or yeast display libraries are typically limited to below[34,35]. By monitoring output mRNA using qPCR, we learned that the initial ribosome display round generates a pool of around 5 x 10^6 sybodies, from which a focused phage display library can be constructed with minimal effort in analogy to nanobody cloning from B-cells of immunized camelids[6]. We further realized that binder generation against challenging membrane proteins profits from radical changes in display format and target immobilization between each selection round to overcome otherwise inevitable biases. The sybody platform is therefore built as a selection cascade, in which the library is pre-enriched by ribosome display and then funneled into a phage display library of optimal size.

The sybody platform was validated by determining crystal structures of three sybody-MBP complexes, unequivocally proving the integrity of the scaffold as well as the utility of the randomized binding surface. We generated sybodies that specifically recognize the ATP-bound state of an ABC transporter, which cannot be populated in the blood stream of an animal during immunization. A remarkable achievement of the method is the rapid generation of conformation-selective sybodies against disease-relevant human drug discovery transporter targets, which were previously intractable to antibody or nanobody immunization techniques. The libraries and protocols will be made fully available to academic labs in order to facilitate the spread and further development of the technology. In conclusion, the sybody platform is a remarkably fast and reliable technology enabling the next generation of challenging drug discovery targets including receptors, channels and transporters.

MATERIALS AND METHODS

Construction of vectors for phage display and sybody expression

An FX cloning vector for the periplasmic production of VHHs preceded by an N-terminal decaHis-tag and an HRV 3C protease cleavage side, designated pBXNPH3 (Supplementary Fig. 6), was constructed by polymerase chain reaction (PCR) using Phusion polymerase and pBXNH3[16] as template in combination with the 5'-phosphorylated primers pBXNPH3_#1 and pBXNPH3_#2 . The resulting 5848 bp product was DpnI-digested, column-purified, ligated and transformed to chemically-competent *ccdB*-resistant *E. coli* DB3.1. A variant designated pBXNPHM3 (Supplementary Fig. 6), which is compatible with the periplasmic production of VHHs as a fusion to an N-terminal maltose binding protein (MBP), decaHis-tag and HRV 3C protease site, was constructed from fragments of pBXNH3[16] amplified using primer pair pBXNPHM3_#1 (holding an NotI restriction site) and pBXNPHM3_#2 (5'-phosphorylated) and pET26FX[36] amplified using primer pair pBXNPHM3_#3 (5'-phosphorylated) and pBXNPHM3_#4 (holding an NotI restriction site). The resulting products of 4241 bp (vector backbone) and 2732 bp (insert holding *mbp* and *ccdB*) bp, respectively were DpnI-digested, gel-purified, cut with NotI, ligated and transformed into *E. coli* DB3.1. To obtain pNb_init, the ampicillin (*amp*) resistance gene of pBXNPH3 was replaced by a chloramphenicol (*cat*) marker. Because the kill cassette of pBXNPH3 contained a chloramphenicol marker as well, pBXNPH3 containing nanobody 1ZVH was used as template to amplify the vector without the *amp* gene with the primer pair pBXNPH3_blunt_for/ pBXNPH3_EcoRI_rev. The *cat* gene was amplified from pINIT_cat[16] (Addgene #46858) using primers Cm_EcoRI_for and Cm_blunt_rev (5'-phosphorylated). The PCR products were column purified, digested with EcoRI, purified by gel extraction, ligated and transformed into *E. coli* MC1061. The resulting plasmid was amplified by primers Nb_init_for (5' phosphorylated) and Nb_init_rev and circularized by ligation. The resulting vector was cut with SapI and the kill cassette excised from pINIT_cat using the same restriction enzyme was inserted, resulting in vector pNb_init. The FX cloning vector for phage display, named pDX_init, was constructed based on pMESy4 [6]. The internal SapI site preceding the lac promoter in pMESy4 was removed by generating a short 500 bp PCR fragment with a mutation in the SapI recognition site using the primer pair pDX_init_#1/ pDX_init_#2. The NcoI/SapI-digested PCR product was gel-purified and ligated into NcoI/SapI-digested and dephosphorylated pMESy4 and transformed into *E. coli* MC1061. The amber stop codon on the resulting vector was replaced by a glutamine residue using Quickchange mutagenesis and primer pair pDX_init_#3/ pDX_init_#4. The resulting vector backbone was amplified using primer pair pDX_init_#5/ pDX_init_#6, thereby introducing SapI sites as part of the open reading frame, digested with SapI and ligated with a SapI-digested PCR-fragment holding the counterselection marker *sacB* amplified from pINITIAL[16] using primer pair pDX_init_#7/ pDX_init_#8.

9

Sybody expression and purification

Sybodies were expressed in *E. coli MC1061*, which were grown in terrific broth containing 25 µg/ml chloramphenicol (in case of pNb_init) or 100 µg/ml ampicillin (in case of pBXNPH3 as well as pBXNPHM3) to an OD_{600} of 0.7 at 37 °C. Then the temperature was switched to 25 °C and cells were induced with 0.002 % (w/v) L-arabinose for 15 h. Cells were disrupted using a microfluidizer processor (Microfluidics) at 25'000 lb/in^2 in TBS (20 mM Tris-HCl pH 7.5, 150 mM NaCl) supplement with 2 mM $MgCl_2$ and 25 µg/ml DNAse I. Cell debris was removed by centrifugation at 8'000 g for 20 min and 15 mM imidazole pH 7.5 was added prior to loading onto a gravity flow Ni-NTA superflow column of 2 ml bed volume (Qiagen). The column was washed with 25 ml TBS containing 30 mM imidazole pH 7.5 in case of pNb_init (50 mM in case of pBXNPH3 or pBXPHM3) and the sybody was eluted with 5 ml TBS containing 300 mM imidazole pH 7.5. If expressed in pBXNPH3 as well as pBXNPHM3, the Ni-NTA purified sybodies were dialyzed against TBS in the presence of 3C protease for 3 hours and loaded onto Ni-NTA columns to remove the His-tag or the His-tagged MBP as well as the 3C protease. Tag-free sybodies were eluted from the Ni-NTA column using TBS containing 30 mM imidazole. Sybodies were concentrated using centrifugal filters with a 3 kDa cut-off (Amicon Ultra-4) and separated by size exclusion chromatography (SEC) using Superdex 200 Increase 10/300 GL (GE Healthcare) or Sepax-SRT10C (Sepax Technologies) in TBS.

Assembly of sybody libraries

Synthetic genes encoding for three non-randomized scaffold sybodies (convex, loop and concave) were ordered at DNA2.0 and their sequences are provided in Supplementary Table 2. These scaffold sybodies contain serines and threonines in the positions to be randomized in the respective libraries and served as PCR templates for library assembly. Primers (Supplementary Table 4) were added to the PCR reaction at a concentration of 0.8 µM if not specified differently. Primers for the sybody assembly were ordered in PAGE-purified form (Microsynth). If not otherwise mentioned, Phusion High-Fidelity DNA Polymerase (NEB) was used for PCR amplification. Five double-stranded PCR products, which served as megaprimers in the assembly of the library were first amplified from the genes encoding for the frameworks of the concave/loop and convex library. The gene of the concave/loop sybody framework was amplified with primer pairs FW1_a_b_for/FW1_a_b_rev (megaprimer 1), FW3_a_b_for/Link2_a_rev (megaprimer 2) or FW3_a_b_for/Link2_b_rev (megaprimer 3). The gene of the convex sybody framework was amplified with primer pairs FW1_c_for/FW1_c_rev (megaprimer 4) or FW3_c_for/Link2_c_rev (megaprimer 5). Megaprimers were gel-purified. In a second step, the individual CDR regions of the libraries were assembled by

10

overlap extension PCR using Vent DNA Polymerase (NEB), applying 35 cycles and an annealing temperature of 60 °C. 100 µl of the PCR reactions contained 10 µl 10 x Vent buffer, 1 µl Vent DNA polymerase, 5 µl DMSO, 0.4 mM dNTPs, 1 µM outer primers, 50 nM randomized primer, 25 nM megaprimers (if applicable) and 25 nM internal assembly primer (if applicable). Supplementary Table 5 lists the primers used for the assembly of the CDRs of the respective libraries. The assembly PCR reactions yielded single DNA species of the expected size, which was purified by PCR purification kit (Qiagen). Fragments containing CDR1 were digested with BsaI and fragments containing CDR2 with BpiI. Digestion with these two Type IIS restriction enzymes resulted in complementary overhangs of 4 base pairs. Digested DNA was purified by PCR purification kit and CDR1-CDR2 pairs belonging to the respective library were ligated with T4 ligase. Separation of the ligation product by DNA gel revealed almost complete ligation of the fragments. The three ligation products consisting of the respective CDR1-CDR2 pairs of the three libraries were purified by gel extraction and yielded around 800 ng of DNA. The purified ligation product served as template to amplify the CDR1-CDR2 region using primer pairs FW1_a_b_for/Link2_a_rev, FW1_a_b_for/Link2_b_rev and FW1_c_for/Link2_c_rev for the concave, loop and convex library, respectively. The resulting PCR product was cleaned up by PCR purification kit and digested with BsaI. The purified CDR3 regions were digested with BpiI. The resulting compatible overhangs were ligated and the ligation product corresponding to the final assembled library was purified by gel extraction. The DNA yields were 6.8 µg, 8.6 µg and 9.4 µg for the concave, loop and convex library, respectively.

Attachment of the flanking region for ribosome display

3.42 µg, 3.6 µg and 3.77 µg of the ligated concave, loop and convex library, respectively, were used as template for PCR amplification using primer pairs Med_FX_for/ Med_FX_rev for the concave and the loop library and Long_FX_for/ Long_FX_for for the convex library. The amount of ligated libraries used as template for PCR amplification represented the bottleneck of the library construction with diversities corresponding to 9×10^{12} for each of the three sybody libraries. The resulting PCR product was cleaned up by PCR purification kit (Macherey-Nagel), digested by BspQI, and column purified again. The pRDV vector[11] was made compatible with FX cloning by amplifying the vector backbone using the primer pair RD_FX_pRDV_for/RD_FX_pRDV_rev1. The resulting PCR product was digested with BspQI, ligated with the *ccdB* kill cassette excised from the vector pBXCA3GH[37] (Addgene #47071) with the same enzyme and transformed into *E. coli* DB3.1. The DNA region between the stem loop and the ribosome binding site was shortened by amplifying the resulting vector with the 5'-phosphorylated primer pair pRDV_SL_for/pRDV_SL_rev and ligating the obtained PCR product. The resulting vector was called pRDV_FX5 and served as template to amplify the 5' and 3' nucleotide

sequences required for in *vitro* translation of mRNA for ribosome display. This was performed by PCR amplification with primer pairs 5'_flank_for/5'_flank_rev and 3'_flank_for/tolAk_rev, respectively. The resulting PCR products were column purified, digested with BspQI and column purified again. The digested flanking regions (60 µg and 80 µg of the 5' and 3' flank, respectively) were ligated with 75 µg of the respective sybody library in a volume of 4 ml and using 80 µl T4 ligase (5 U/µl, ThermoFisher). After ligation, the DNA fragments were gel-purified and yielded 13.8 µg, 22 µg and 20.5 µg ligation products corresponding to the flanked concave, loop and convex libraries, respectively. 10 µg of each ligation was amplified by PCR using primers 5'_flank_for and tolAk_2 in a 96 well plate and a total of 5 ml PCR reaction and the resulting PCR product was column purified, yielding 180 µg, 188 µg and 192 µg for the concave, loop and convex libraries, respectively. 10 µg of each amplified library was *in vitro* transcribed using RiboMAX Large Scale RNA Production System (Promega) and yielded 1.5 mg of mRNA for each library.

Purification and biotinylation of target proteins

The coding sequence of GFP was cloned using primer set GFP_FX_FV/GFP_FX_RV into FX vector pBXNH3. GFP was purified by Ni-NTA, followed by HRV 3C protease cleavage, rebinding on Ni-NTA and SEC in PBS. Chemical biotinylation of GFP was carried out in PBS using EZ-Link™ Sulfo-NHS-LC-Biotin (Thermo Fisher) and mass spectrometry analysis revealed that the reaction contained predominantly GFP having one biotin moiety attached. The coding sequence of MBP was cloned using primer sets MBP_FX_FW/MBP_FX_FW into FX vector pBXNH3CA[37], which results in a fusion protein consisting of an N-terminal His10-tag, a 3C protease cleavage site, MBP and a C-terminal Avi-tag. In order to produce Avi-tagged TM287/288, a GS-linker was first introduced into pBXNH3CA between the 3C cleavage site and the adjacent SapI site by amplifying the vector with the primer pair GS-Linker_FW (5' phosphorylated) and GS-Linker_RV, each containing half of the GS-linker as overhang, and blunt-end ligation of the resulting PCR product. The resulting expression vector was called pBXNH3LCA. TM287/288 was cloned into pBXNH3LCA by FX cloning, which attaches a cleavable His10-tag to the N-terminus of TM287 separated by a GS-linker and an Avi-tag to the C-terminus of TM288. GFP, MBP-Avi and TM287/288-Avi were expressed in and purified from *E. coli* by Ni-NTA chromatography as described previously[11,20]. The enzymatic site-specific biotinylation of the Avi-tag was carried out at 4°C over night using purified BirA in 20 mM imidazole pH 7.5, 200 nM NaCl, 10 % glycerol, 10 mM magnesium acetate, 0.03 % β-DDM in case of TM287/288 and a two-fold excess of biotin over the Avi-tag concentration. 3C protease was added as well to this reaction mixture to cleave off the His10-tag. Next day, the mixture was loaded onto Ni-NTA columns to remove the His-tag , BirA and the 3C protease. Biotinylated target proteins were eluted from the Ni-

NTA column using TBS containing and 30 mM imidazole and in case of TM287/288 0.03 % β-DDM. Finally, biotinylated target proteins were purified by SEC using a Superdex 200 Increase 10/300 GL (GE Healthcare) in TBS containing in case of TM287/288 0.03 % β-DDM.

Human ENT1(2-456) was expressed by transient transfection in HEK293 freestyle cells as wild-type, full length protein using a pCDNA3.1(+) base vector (Invitrogen) and synthesized, codon optimized genes cloned by Genewiz. The protein was designed as N-terminal fusion of His-GFP-3C-Avi with His being a 10-fold repeat of histidine, GFP is the enhanced green fluorescent protein, 3C is the 3C precision protease cleavage site (LEVLFQGP) and Avi the sequence corresponding to the Avi-tag (GLNDIFEAQKIEWHE). Biotinylation was performed *in vivo* during protein production by co-transfection of cells with two different pCDNA3.1(+) plasmids coding for the ENT1 construct and *E.coli* BirA ligase and by supplementing the medium with 50 µM biotin during fermentation. Cells were harvested in all cases 65 hours post transfection and were flash frozen at -80 °C. For purification, cell pellets were thawed and resuspended in solubilization buffer containing 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 1 Roche protein inhibitors complete tab per 50 ml of buffer in a 1 to 3 ratio (3 ml of buffer for 1 g of cells) under gentle agitation for 30 min until homogeneity. Subsequently, 1 % (w/v) of lauryl maltose neopentyl glycol (LMNG) and 10 µM S-(4-Nitrobenzyl)-6-thioinosine (NBTI) were added to the suspension and incubated for 1 hour at 4 °C. Supernatant was cleared by ultracentrifugation at 100'000xg using a Beckmann Ti45 rotor for 45 minutes and incubated overnight at 4 °C with 15 ml of preconditioned TALON affinity resin under gentle stirring. Resin was collected by low speed centrifugation and washed with a total of 15 column volumes 20 mM Tris-HCl pH 7.5, 0.003 % (w/v) LMNG, 20 mM imidazole, 300 mM NaCl and 10 µM NBTI several times. Using an empty pharmacia XK16 column, resin was collected and further washed with a buffer containing 20 mM Tris-HCl pH 7.5, 0.003 % (w/v) LMNG, 20 mM imidazole, 300 mM NaCl, 10 µM NBTI and 15 µM dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) lipid and washed a second time using the same buffer containing 40 mM imidazole. Protein was eluted at 300 mM imidazole and subjected to a desalt step on a 53 ml GE Hi Prep 26/10 desalting column using desalting buffer consisting of 20 mM Tris-HCl pH 7.5, 0.003 % (w/v) LMNG, 300 mM NaCl, 10 µM NBTI and 15 µM DOPS. To remove the GFP-His tag and reduce protein glycosylation, 3C-Prescission protease was added at a concentration of 1 unit per 50 µg of protein together with 10 µg/ml of PNgaseF and 10 µg/ml Endo-alpha-N-acetylgalactosaminidase. The mixture was incubated overnight at 4 °C and His-GFP tag removal was monitored by fluorescence-detection size-exclusion chromatography (FSEC) analysis. To completely remove GFP, protein was purified by an affinity purification step using a HiTrap TALON column collecting the flow-through. Subsequently, protein was concentrated using an Amicon filter unit at 30 kDa molecular weight cut-off to about 0.5-1 ml volume and further purified via size–exclusion chromatography using a Superdex 200 10/300 GL increase column equilibrated in

13

the SEC buffer containing 20 mM Tris-HCl pH 7.5, 0.003 % (w/v) LMNG, 300 mM NaCl, 10 µM NBTI and 15 µM DOPS. SEC fractions corresponding to the biotinylated ENT1 protein were concentrated on an Amicon filter unit with a cut-off of 30 kDa to a final concentration of 1.35 mg/ml and stored at -80 upon flash freezing in liquid nitrogen. Quality and biotin modification of the protein were analysed by LC-MS revealing close to complete biotinylation.

Human GlyT1 was cloned as a codon optimized gene into a modified pOET1 base vector (Oxford Expression) by Genewiz and expressed either as a C-terminal 3C-GFP-His or C-terminal Avi-3C-GFP-His fusion in *Spodoptera frugiperda* (Sf9) cells. Cells were grown to 2 million cells/ml in Sf9III medium in 50 liter wave bags (Sartorius, 25l maximal volume) and infected using 0.25-0.5 % (v/v) of virus. 72 hours post infection at viabilities higher than 85%, cells were harvested by centrifugation at 3000xg for 10 minutes and 4 °C, washed in PBS and re-centrifuged for 20 minutes. Cells were filled in plastic bags and frozen by putting the bag in a -80 °C freezer. Thawed biomass was further washed twice by resuspension and centrifugation for 20 minutes at 5000xg in a buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 Roche Complete Tablet per 50 ml volume. Washed cells were resuspended for solubilization for 30 minutes at 4 °C while stirring in a buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 30 mM imidazole pH 7.5, 1 % (w/v) LMNG, 0.1 % CHS, 15 µM DOPS and 100 µM Cmpd1, a more soluble analogon of the glycine transporter 1 reuptake inhibitor Bitopertin. The suspension was centrifuged at 100'000xg in a Ti45 rotor for 20 minutes at 4°C to collect supernatant. Protein was purified by batch purification using TALON affinity resin (GE Healthcare), incubated with resin for 16 hours under stirring at 300 rpm and then centrifuged at 500xg for 2 minutes in a 50 ml Falcon tube. Four wash steps with a buffer containing 50 mM Tris-HCl, 150 mM NaCl, 30 mM imidazole pH 7.5, 0.05 % (w/v) LMNG, 0.005 % (w/v) CHS, 15 µM DOPS, 50 µM Cmpd1 and 1 Roche Complete Tablet per 50 ml volume were followed by loading the resin into a XK26 column (GE Healthcare), washed again with four column volumes to finally elute the protein with the same buffer that contained 300 mM imidazole. The Avi-3C-GFP-His but not the 3C-GFP-His protein was treated with HRV-3C protease (Novagen) and in-house produced PNGase F (*F. meningosepticum*) to cleave the GFP-His tag and trim existing glycosylations. Subsequently, the Avi-tagged GlyT1 was desalted into a buffer optimal for enzymatic biotinylation consisting of 20 mM bicine pH 8.3, 150 mM potassium-glutamate pH 7.5, 0.05 % (w/v) LMNG, 0.005 % CHS, 15 µM DOPS and 50 µM Cmpd1 to remove imidazole and subjected to another round of TALON affinity purification to remove the cleaved GFP-His tag in the same buffer. The flow through containing the GlyT1 protein was concentrated to 1.1 mg/ml concentration with an Amicon Ultra 4 filter unit (Millipore) with a molecular weight limit of 30 kDa for complete biotinylation using the BirA-500 biotinylation kit (Avidity) according to the protocol. Biotinylation was monitored by liquid-chromatography-coupled mass spectroscopy. Both the Avi fusion as well as the 3C-GFP-His fusion protein were further purified

14

via size-exclusion chromatography using a Superdex 200 10/300 GL increase column and a buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % (w/v) LMNG, 0.005 % (w/v) CHS, 15 µM DOPS and 50 µM Cmpd1.

Sybody selections against MBP

To display sybody libraries on ribosomes, 10 µl of the PURE*frex®SS* translation mix was prepared. The kit components were mixed to a total volume of 9 µl and incubated at 37 °C for 5 min. 1 µl of 2 µM library RNA was added to the translation mix and incubated at 37 °C for 30 min. The ribosomal complexes were diluted in 100 µl ice cold WTB buffer (50 mM Tris-acetate pH7.5, 150 mM NaCl, 50 mM magnesium acetate) supplemented with 0.05 % Tween 20, 0.5 % BSA and 5 mg/ml Heparin. 10 µl Dynabeads MyOne Streptavidin T1 (Life Technologies) were washed 3 times with 500 µl WTB and blocked with 500 µl WTB 0.5% BSA for 1 hour followed by 3 washes 500 µl of WTB-T-BSA (0.05% Tween 20, 0.5% BSA). The magnetic beads were coated in 100 µl WTB-T-BSA containing 50 nM biotinylated MBP for 1 hour followed by 3 washes of 500 µl WTB-T-BSA. The ribosomal complexes were incubated with the beads for 20 min followed by 3 washes of 500 µl WTB-T. During the last wash step, the beads were placed in a fresh tube. The RNA was eluted by resuspending the beads in 100 µl TBS supplemented with 50 mM EDTA pH 8.0 and 100 µg/ml yeast RNA and incubated for 10 min at room temperature. The eluted RNA was purified using the RNeasy micro kit (Qiagen) and eluted in 14 µl RNase-free water. Reverse transcription was performed by mixing 14 µl of the eluted RNA with 2µl of RT_Primer at 100 µM and 4 µl of 10 mM dNTPs. The mixture was heated to 65 °C for 5 min, and then cooled on ice. Using this mixture, a 40 µl RT reaction was assembled according to the manual (Affinity Script, Agilent) and incubated 1 hour at 37 °C, followed by 5 min at 95°C. The cDNA was purified using the PCR purification kit (Macherey Nagel) and eluted in 30 µl elution buffer. 25 µl of the purified cDNA was amplified by PCR using the primers Medium_ORF_for and Medium_ORF_rev for the concave and loop library, and Long_ORF_for and Long_ORF_rev for the convex library, respectively. The PCR product was purified via gel and used as template in an assembly PCR to add the flanking regions for *in vitro* transcription using megaprimers. Megaprimers to flank the concave and loop sybodies were obtained by amplifying pRDV_FX5 containing the non-randomized loop sybody using primer pairs 5'_flank_for/ Medium_ORF_5'_rev and Medium_ORF_3'_for/ tolAk_rev. Megaprimers to flank the convex sybodies were obtained by amplifying pRDV_FX5 containing the non-randomized convex sybody using primer pairs 5'_flank_for/ Long_ORF_5'_rev and Long_ORF_3'_for/ tolAk_rev. Flanking was performed by assembly PCR using 200 ng sybody pool obtained from RT-PCR, 120 ng of 5'-flank, 360 ng of 3'-flank and 5 µM of outer primers 5'_flank_for and tolAk_2 in a volume of 100 µl. The resulting PCR product was separated on

gel, purified and used as input material for 10 µl reaction of the RiboMAX Large Scale RNA Production System (Promega). The resulting RNA was purified using the RNeasy kit (Qiagen) and used as input RNA of the next round. The second round was performed according to the first round. In the third round, the PCR product of the amplified cDNA was amplified using primers Med_FX_for/ Med_FX_rev (concave and loop library) or Long_FX_for/ Long_FX_for (convex library) to add FX overhangs to the DNA and subsequently cloned into the pNb_init vector by FX cloning for expression and ELISA.

Sybody selections against membrane proteins

One round of ribosome display was performed as in the MBP section with the following exceptions. Tween 20 was replaced by 0.1 % β-DDM for selections against TM287/288(E517A), 0.1 % β-DDM and 0.005 % CHS against ENT1, or by 0.05 % LMNG and 0.005 % CHS against GlyT1. Selections against TM287/288(E517A), ENT1 and GlyT1 were carried out in the presence of 1 mM ATP, 10 µM NBTI and 50 nM Cmpd1, respectively. Instead of immobilizing the target prior to panning, solution panning was performed by incubating ribosomal complexes and 50 nM biotinylated target protein for 30 min prior to the pulldown via streptavidin coated magnetic beads (Dynabeads MyOne Streptavidin T1). The PCR product of the amplified cDNA was used in a PCR using Med_FX_for/ Med_FX_rev for concave/loop sybodies or Long_FX_for/ Long_FX_for for convex sybodies to add FX overhangs to the DNA and subsequently cloned into the phagemid vector pDX_init. To this end, amplified sybody pools and pDX_init were digested with BspQI, gel-purified and 500 ng sybody insert and 1µg pDX_init backbone were ligated in 50 µl using 5 units of T4 DNA Ligase (Thermo Scientific). The ligation reaction was mixed with 350 µl electrocompetent *E.coli* SS320 on ice. After electroporation (Bio Rad Gene Pulser, 2.5 kV, 200 Ω, 25pF), the cells were immediately resuspended in 25 ml SOC and shaken at 37 °C for 30 min for recovery. Subsequently, the cells were diluted in 250 ml 2YT, 2 % glucose, 100 µg/ml ampicillin and grown overnight shaking at 37 °C. For phage production, this phagemid-containing *E.coli* SS320 overnight culture was inoculated 1:50 in 50 ml 2YT, 2 % glucose, 100 µg/ml ampicillin and grown to OD_{600} of 0.5. 10 ml of this culture was superinfected with 3 x 10^{11} plaque forming units M13KO7 Helper Phage at 37 °C without shaking for 30 min. Cells were collected by centrifugation and resuspended in 50 ml 2YT, 100 µg/ml ampicillin, 25 µg/ml kanamycin and incubated at 37 °C shaking overnight for phage production. Next day, cells were pelleted by centrifugation and 40 ml of the culture supernatant were mixed with 10 ml 20 % PEG6000 (v/v), 2.5 M NaCl and incubated on ice for 30 min to precipitate the phages, which were subsequently pelleted by centrifugation. The pellet was resuspended in phosphate-buffered saline (PBS) and cleared twice by centrifugation in a tabletop centrifuge at full speed. Phage concentration was determined by UV-

Vis spectroscopy. The first round of phage display was performed in a Maxi-Sorp plate (Nunc) which was coated overnight with 100 µl per well of 60 nM neutravidin in TBS. Next day, the plate was washed 3 times with TBS and blocked with TBS containing 0.5 % BSA for 1 hour. The phages were diluted to 10^{12} phages per ml in ice cold TBS-D-BSA (containing 0.1 % β-DDM for TM287/288(E517A), 0.1 % β-DDM and 0.005% CHS and ENT1 or 0.05 % LMNG and 0.005 % CHS for GlyT1, and 0.5 % BSA for all targets) and incubated in solution for 20 min with 50 nM biotinylated target protein (in the presence of 1 mM ATP + 5 mM $MgCl_2$ for TM287/288(E517A), 10 µM NBTI for ENT1 or 50 nM Cmpd1 for GlyT1). The plate was prepared by washing 3 times with ice cold TBS-D-BSA and 100 µl of the phage/target mix was added per well and incubated for 10 min at 4 °C. The plate was subsequently washed 3 times with 250 µl/well ice cold TBS-D (devoid of ligands and BSA). Phages were eluted by adding 100 µl TBS with 0.25 mg/ml trypsin and incubation at room temperature for 30 min. Trypsin was inhibited by adding 0.125 mg/ml 4-(2-Aminoethyl)benzenesulfonyl fluoride to the elution. For infection, a culture of *E.coli* SS320 was grown in 2YT to an OD_{600} of 0.5. 100 µl of eluted phages was added per milliliter of the culture and incubated at 37 °C without shaking for 30 min. The cells were then diluted 1:10 in 2YT, 2 % glucose, 100 µg/ml ampicillin and grown overnight shaking at 37 °C resulting in a preculture for phage production for the next round. The second round of phage display was performed according to the first one except that magnetic beads (Dynabeads MyOne Streptavidin C1) were used to pull down target-phage complexes. The resulting culture of infected cells was then used to purify the phagemids (MiniPrep, Qiagen) containing the selected sybodies. Sybody sequences were subcloned into pNb_init using FX cloning and transformed into *E. coli MC1061* for ELISA analysis and protein purification.

Sybody identification by ELISA

Single sybody clones were picked and expressed in 1 ml terrific broth containing 25 µg/ml chloramphenicol in a 2ml 96 deep well plate. After expression, the cells were pelleted by centrifugation and resuspended in 50 µl B-PER II for lysis. The lysate was diluted with 950 µl TBS and centrifuged to pellet cell debris. ELISAs were carried out in Maxi-Sorp plates (Nunc) coated overnight with 100 µl/well of 5 µg/ml Protein A in TBS. The plate was washed 3 times with 250 µl TBS and blocked with 250 µl TBS-BSA. All washing steps were performed using 3 times 250 µl TBS containing detergent (Tween-20 for MBP; 0.03 % β-DDM for ENT1 and TM287/288; 0.05 % LMNG for GlyT1) between all incubation steps, which were carried out in 100 µl TBS-D-BSA for 20 min. These steps were anti-myc antibody 1:2000 (Sigma Aldrich M4439) followed by 5 fold diluted sybody lysate, then 50 nM of the biotinylated target protein or biotinylated control protein and finally streptavidin-HRP 1:5000 (Sigma Aldrich, S2438) (Fig. 1). The ELISA was developed by adding 100 µl of 0.1mg/ml TMB

(Sigma Aldrich 860336) in 50 mM Na_2HPO_4, 25 mM citric acid and 0.006 % H_2O_2. ELISA signals were measured at an absorbance of 650 nm.

Monitoring of binder enrichment by qPCR

With the term "enrichment", we refer to the experimentally determined fold excess of polynucleotides eluted from a selection round against a target of choice versus an analogous selection round against another immobilized protein, which was not used for selections in preceding rounds. In order to determine enrichments, qPCR was performed in a 10 µl reaction containing SYBR select Master Mix (Thermo Fischer Scientific), 300 nM of each primer, 5 % DMSO and 2 µl cDNA (ribosome display) or phages (phage display) diluted 10 fold in H_2O. Standard curves for each primer pair were determined using a dilution series of the phagemid pDX_init or a PCR product corresponding to the sequence of the cDNA obtained after ribosome display. Their initial concentration was determined by UV-Vis spectroscopy. PCR efficiencies for all primer pairs were between 95 and 98 %. Cycling conditions were: 2 min 95 °C initially for polymerase activation, followed by 10 sec 95 °C and 30 sec 63 °C for 45 cycles. The runs were performed in a 7500 fast qPCR machine (Thermo Fischer Scientific). The following primer pairs were used: qPCR_RD_5'_for in combination with qPCR_ RD_S&M_5'_rev for the concave and loop library or qPCR_ RD_L_5'_rev for the convex library to determine the amount of full length cDNA, qPCR_ RD_tolA_3'_for/ qPCR_ RD_tolA_3'_rev to determine the total amount of cDNA, qPCR_PD_pDX_for/ qPCR_PD_pDX_rev to determine the amount of phages and qPCR_3K1K_for/ qPCR_3K1K_rev to determine the amount of 3K1K cDNA.

Protein stability measurements using Thermofluor

Thermofluor was performed in a 25 µl reaction of PBS containing 100x SYPRO Orange (Life Technologies) and 0.5 mg/ml sybody. The reaction was heated with a 1 % ramp from 25 to 99 °C in a 7500 fast qPCR machine (Thermo Fischer Scientific) while the fluorescence intensity was measured through a ROX filter. The raw data was extracted and fitted as described previously[38].

Surface plasmon resonance

Binding affinities were determined using surface plasmon resonance (SPR). MBP binders were analyzed using a Biacore X100 machine (GE healthcare). 570 response units (RU) of biotinylated MBP

were immobilized on a streptavidin coated SPR chip (Sensor Chip SA). Sybodies Sb_MBP#1-3 were purified in TBS, 0.05 % Tween-20 as described above. SPR measurements were carried out in the same buffer. Affinities of Sb_MBP#1 were in addition determined in the same buffer supplemented with increasing maltose (4-O-α-D-Glucopyranosyl-D-glucose) concentrations (0, 5, 10, 25, 50 and 100 µM). In these analyses, 860 response units (RU) of biotinylated MBP were immobilized. Data were analyzed by plotting sybody affinity ratios determined in the presence (K_D') and absence (K_D) of maltose against the maltose concentration and fitting the data with the following equation:

$$y = \frac{\frac{x}{B} + 1}{\frac{\alpha x}{B} + 1}$$

B corresponds to the binding affinity of maltose ($K_{D,maltose}$) and α corresponds to the allosteric constant.

Affinities of sybodies directed against TM287/288 were measured using a ProteOn™ XPR36 Protein Interaction Array System (Biorad). Biotinylated TM287/288 and TM287/288(E517A) mutant were immobilized on a ProteOn™ NLC Sensor Chip at a density of 1500 RU. Sybodies expressed in pNb_init were SEC-purified in TBS, and SPR analysis was carried out in the same buffer containing 0.03 % β-DDM and either 1 mM $MgCl_2$ or 1 mM $MgCl_2$ + 0.5 mM ATP, to measure binding affinities in the presence or absence of ATP.

All ENT1 and GlyT1 SPR experiments were performed on a Biacore® T200 (GE Healthcare, Uppsala, Sweden) instrument at 18 °C in running buffers containing either 20 mM HEPES pH 7.5, 150 mM NaCl, 0.001 % (w/v) LMNG, 1.25 µM DOPS (ENT1) or 20 mM Citrate pH 6.4, 150 mM NaCl, 0.004 % (w/v) LMNG (GlyT1) at 50 µl/min flow rate. Running buffers were freshly prepared, filtered with Express™Plus steritop filters with 0.22 µm cut off (Millipore, Billerica, MA, USA) and degassed prior the SPR analysis. Biotinylated ENT1 or GlyT1 were captured on streptavidin pre-coated SA sensors (GE Healthcare BR-1000-32). First, streptavidin sensors were conditioned with 3 consecutive 1 minute injections of high salt solution in sodium hydroxide (50 mM NaOH, 1 M NaCl). Next, biotinylated protein samples were applied to a streptavidin sensor surface for protein immobilization levels of about 1000 RU and 300 RU for ENT1 and GlyT1, respectively. Finally, free biotin solution (10 µM in running buffer) was injected once (1 x 1min) over the sensor surface to block remaining streptavidin binding sites. Dose-response experiments were performed at sybody concentrations up to 475 nM (ENT1) and up to 2 µM (GlyT1). All monitored resonance signals were single referenced, i. e. signals monitored on the binding active channel were subtracted with signals from a reference channel. The sensor surface was not modified with any protein but saturated with free biotin. Data

analysis and curve fitting were performed using Biacore T200 Evaluation (v2.0) and GraphPadPrism 6.07.

Biolayer Interferometrie

The Octet® RED96 System (FortéBio, Pall Inc.) uses disposable sensors with an optical coating layer immobilized with streptavidin at the tip of the sensor. Sensors were decorated with biotinylated MBP to reach a stable baseline, arbitrarily set to 0 nm (Fig. 9). Sensors were dipped in a well containing 500 nM Sb_MBP#1 which leads to the formation of the Sybody-MBP complex. Sensors containing the complex were sequentially dipped in a row of wells containing 500 nM Sb_MBP#1 and increasing concentrations of maltose (0.1, 1, 10, 100, 1000 µM). The reversibility of the competition was shown by decreasing maltose concentrations (1000, 100, 10, 1, 0.1, 0 µM) again in the presence of 500 nM Sybody Sb_MBP#1. The experiments were done in duplicates.

Thermal shift scintillation proximity assay (SPA-TS) using tritiated small molecule inhibitors

Ligand binding assays were performed in a buffer containing either 50 mM Tris-HCl pH 7.5, 150 nM NaCl and 0.004 % (w/v) LMNG or 20 mM Citrate pH 6.4, 150 mM NaCl and 0.004 % (w/v) LMNG, respectively. For each analysis, 140 µl protein solution (7nM for ENT1 and 10 nM for GlyT1) was added to each of 12 wells of a 96-well Eppendorf PCR plate at 4 °C and incubated subsequently for 10 minutes with a temperature gradient from 30-60°C (ENT1) or 23-53°C (GlyT1) across twelve wells in a Techne Prime Elite thermocycler. Subsequently, the plate was centrifuged at 2250xg for 3 minutes at 4 °C and 135 µl of protein solution transferred to a 96-well Optiplate (Perkin Elmer) preloaded with 15 µl Copper SPA beads (20 mg/ml) per well (PerkinElmer) to obtain a final bead concentration of 0.3 mg/well. After 15 minutes of incubation at 4 °C and 1000 rpm on a BioShake iQ, a final concentration of 6 nM tritiated [^{3}H]-NBTI (Perkin-Elmer) or [^{3}H]-Org24598 compound (1.2mCi/ml specific activity, 50 µl of a 24 nM stock solution) was added and incubated for 45 minutes at 4°C and 1000 rpm on a Bioshake iQ. Scintillation analysis was performed using a TopCount Microplate Scintillation Counter and apparent melting temperatures (T_ms) were determined in GraphPad Prism 6.07 using a non-linear fit to a Boltzmann sigmoidal function.

ATPase inhibition of TM287/288

ATPase activity was measured as described previously[21] in TBS containing 0.03 % β-DDM and 10 mM MgSO$_4$ at increasing concentrations of sybodies Sb_TM#26 or Sb_TM#35. ATP concentration was 50 µM, TM287/288 concentration was 25 nM, assay temperature was 25°C and incubation time was 20 min. The data were fitted to a hyperbolic decay curve with the following function (SigmaPlot):

$$y = y_0 + \frac{a\ IC_{50}}{IC_{50} + x}$$

in which *y* corresponds to the ATPase activity at the respective sybody concentration divided by the ATPase activity in the absence of inhibitor normalized to 100 %, *IC$_{50}$* corresponds to the sybody concentration for half-maximal inhibition, *y$_0$* corresponds to the residual activity at infinite sybody concentration, *a* corresponds to the maximal degree of inhibition, and *x* corresponds to the sybody concentration.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

MAS, ERG and RJPD conceived the project. ERG and MAS designed the sybody library. MAS and IZ assembled the library. IZ and PE established the sybody selection platform. CAH established the ELISA setup. ERG designed and constructed FX cloning expression vectors. IZ and PE selected sybodies against MBP and determined the sybody-MBP complex structures. SG and PS performed SPR and Biolayer Interferometry measurements for MBP sybodies under the supervision of DG. CAH and IZ selected and characterized by SPR sybodies against TM287/288, and CAH performed the ATPase activity assays. RJPD and NB designed all experiments and expression vectors for GlyT1 and ENT1. IZ selected sybodies against ENT1 and GlyT1. MS performed fermentation, PS, NB and LS performed small-scale expressions and protein purification and PS, NB and JG performed scintillation proximity assays (SPA-TS) for small molecules and sybodies for GlyT1 and ENT1. JG performed reverse phase chromatography experiments for GlyT1. MG performed SPR measurements for GlyT1 and ENT1. PE, IZ, ERG, RFPD and MAS wrote the manuscript.

22

FIGURE LEGENDS

Figure 1:

Selection of sybodies against membrane proteins within three weeks. (**a**) Three synthetic libraries exhibiting highly variable randomized surfaces (concave, loop and convex) each harboring a diversity of 9×10^{12} were designed based on thermostabilized nanobody frameworks. (**b**) The *in vitro* selection platform is built as a diversity funnel, starting with 10^{12} sybodies displayed on ribosomes for pre-enrichment, followed by a focused phage display library of 10^{7} clones and binder identification by ELISA (typically 96 clones). The platform builds on fragment exchange (FX) cloning using Type IIS restriction sites encoded on the phage display (pDX_init) and expression vector (pSb_init) backbones, which generate AGT and GCA overhangs for efficient and facile subcloning. Key elements for reliable selections against membrane proteins are the shape variability of the sybody libraries, exceptionally high experimental diversities using ribosome display and the change of display system during the selection process.

Figure 2:

Structural and biochemical characterization of convex sybody Sb_MBP#1 (**a**) Crystal structure of the Sb_MBP#1/MBP complex. MBP is shown as blue surface, the convex sybody Sb_MBP#1 is shown as grey cartoon with CDRs 1-3 colored in yellow, orange and red, respectively. Sybody residues mediating contacts to MBP are shown as sticks. (**b**) Maltose and sybody Sb_MBP#1 compete for binding to MBP. In the depicted Schild analysis, the sybody affinity ratios determined in the presence (K_D') and absence (K_D) of maltose is plotted against the maltose concentration. The binding affinity for maltose $K_{D,maltose}$ was determined as 1.0 µM. The allosteric constant α amounts to 0.017, i.e. the ratio K_D'/ K_D saturates at a value of 58.

Figure 3:

Conformational trapping of ABC transporter TM287/288 (**a**) In the absence of nucleotides, ABC transporter TM287/288 adopts its inward-facing (IF) state and captures substrates from the cytoplasm. ATP binding is required to achieve a partial population of the outward-facing (OF) state, which allows for substrate exit to the cell exterior. Sybodies were selected in the presence of ATP against the transporter mutant TM287/288(E517A), which is incapable of ATP hydrolysis and predominantly populates the OF state in this condition. (**b**) SPR analysis of convex sybody Sb_TM#35 in the presence and absence of ATP using wildtype TM287/288 and TM287/288(E517A) as ligands. Concentrations of Sb_TM#35: 0, 9, 27, 81, 243, 729 nM. (**c**) ATPase activities of wildtype TM287/288 at increasing concentrations of loop sybody Sb_TM#26 or convex sybody Sb_TM#35. IC_{50}

23

corresponds to the sybody concentration required for half-maximal inhibition and y_0 to the residual ATPase activity at infinite sybody concentrations.

Figure 4:

Conformation-specific binding of Sb_ENT1#1 to the inhibition state of human ENT1. (**a**) Snake plot of human ENT1. (**b**) SPR analysis of Sb_ENT1#1 binding to biotinylated ENT1 revealing a K_d of 40 nM. (**c**) Scintillation proximity assay thermal shift (SPA-TS) analysis of human ENT1 in the presence and absence of Sb_ENT1#1 using [^{3}H]-NBTI inhibitor. Sb_ENT1#1 stabilizes an inhibited conformation as evidenced by a shift of the apparent melting temperature (T_m) by 6.1 °C and **(d)** a 7-fold increase of the absolute SPA signal measured at 30.1 °C.

Figure 5:

Inhibition-state specific sybodies against human GlyT1. (**a**) Schematic of a GlyT1 homolog (PDP ID: 4M48) embedded in a lipid bilayer, illustrating the limited number of surface-accessible epitopes. (**b**) RP8-HPLC analysis of sybody-GlyT complexes previously separated by SEC. (**c, d**) SPR analysis of Sb_GlyT1#1 (K_d = 307 nM) and Sb_GlyT1#6 (K_d = 494 pM). Due to a slow off-rate, SPR analysis of Sb_GlyT1#6 was performed in a single cycle measurement. (**e**) SPR analysis reveals binding of Sb_GlyT1#1-4 to the GlyT1/Sb_GlyT1#6 complex, indicating the presence of two binding epitopes. Sb_GlyT1#5 and Sb_GlyT1#7 compete for binding with Sb_GlyT1#6. **(f)** SPA-TS analysis of Sb_GlyT1#1-7 using [^{3}H]-Org24598 reuptake inhibitor. Shifts of the melting temperature (T_m) are highest for Sb_GlyT1#6 and Sb_GlyT1#7 with values of 8.8 °C and 10 °C, respectively, and correlate well with (**g**) increased absolute SPA signals measured at 23 °C.

Supplementary Figure 1:

Variable sybody scaffolds based on three natural nanobodies. CDR1, CDR2 and CDR3 are colored in yellow, orange and red, respectively. In the left panel, crystal structures of natural nanobodies in complex with GFP (PDB: 3K1K) (**a**), a GPCR (PDB: 3P0G) (**b**) and Lysozyme (PDB: 1ZVH) (**c**) are shown, which served as starting point to delineate scaffolds for randomization. Nanobody residues contacting the target proteins are depicted as sticks. The target proteins are colored in blue. In the middle panel, homology models of three framework nanobodies are shown as cartoons and randomized residues (defined as serines and threonines in these examples) are highlighted as sticks. The three sybody libraries exhibit a concave (**a**), loop (**b**) or convex (**c**) binding surface, respectively. The right panel shows the randomized surface of the three libraries with the side chains of the

24

randomized positions highlighted in color. Note that the concave library contains randomized residues outside of the CDR regions, which are colored in purple.

Supplementary Figure 2:

Framework sequences and randomized positions. (**a** and **b**) Sequences of the framework sybodies are aligned with the sequences of their natural precursors. The frameworks of the concave and the loop library are identical (**a**) while the convex library has its own scaffold (**b**). Residues of the natural precursor nanobodies differing from the framework sequence are marked in blue. The three CDR regions are underlined. Invariant CDR residues contributing to the hydrophobic core of the respective scaffold are marked in green. Note that the differently shaped libraries exhibit alternative sets of invariant CDR residues that precisely match the corresponding scaffolds. This harmonization is a critical and unprecedented feature of our synthetic nanobdy libraries, as it allows for the first time to include variable CDR lengths without the risk of scaffold destabilization. Randomized residues are highlighted as red S (for which randomization mixture 1 was used), as red T (mix 2) and orange T (mix 3). (**c**) Amino acid composition of randomized positions obtained by three different trinucleotide randomization mixtures. Details are provided in Supplementary Note 1.

Supplementary Figure 3:

Biophysical characterization of sybodies. Three framework sybodies representing the concave, the loop and the convex library and containing serines and threonines in the randomized positions were generated by gene synthesis (sequences provided in Supplementary Fig. 2). (**a**) SEC analysis of periplasmatically expressed concave, loop and convex framework sybodies using a Superdex 75 300/10 GL column. (**b**) Determination of melting temperature (T_m) of framework sybodies and their natural precursors 3K1K and 1ZVH using dye SYPRO Orange (ThermoFluor).

Supplementary Figure 4:

Ribosome display of single domain antibodies. (**a**) The non-randomized convex sybody was either purified containing a C-terminal 3x-FLAG tag or displayed on ribosomes containing the same tag using the commercial kit PURE*frex®SS* (GeneFrontier). 3C protease cleavage was used to liberate the displayed sybody from the ribosomal complex. Western blotting analysis using anti-3x-FLAG antibody and purified sybody as standard revealed a display efficiency of 82 % of input mRNA for ribosome display. (**b**) 10^6 mRNA molecules encoding the GFP-specific 3K1K nanobody were displayed on ribosomes using PURE*frex®SS* together with 10^{12} mRNA molecules encoding the non-randomized convex sybody. The ribosomal complexes were pulled down using either biotinylated GFP or MBP immobilized on magnetic beads. The mRNA of isolated ribosomal complexes was isolated, reverse

25

transcribed and the resulting cDNA was analyzed by qPCR. This analysis revealed that 84.6 ± 3.5 % of the input 3K1K mRNA was retrieved on GFP-coated beads, while virtually no background binding of the non-randomized convex sybody nor 3K1K binding to MBP was observed.

Supplementary Figure 5:

Sybody selections against MBP. (**a**) Sybodies were selected against MBP using three rounds of ribosome display and MBP immobilized on magnetic beads. Sybodies were expressed in pNb_init and analyzed by ELISA. (**b**) Binder enrichment was monitored using qPCR by comparing the cDNA output after panning against the target MBP versus the control protein GFP. (**c**) ELISA analysis of convex pool after selection round 3. MBP-specific DARPin off7 was used as positive control.[11] (**d**) SEC analysis of sybody Sb_MBP#3 alone and in complex with MBP using a Superdex 200 300/10 GL column. (**e**) SDS-PAGE analysis of Sb_MBP#3/MBP complex after SEC. (**f**) Crystals formed by sybody-MBP complexes.

Supplementary Figure 6:

FX cloning vector series for phage display and purification of sybodies and nanobodies. Sybody pools from ribosome display (or nanobodies from immunized camelids) are amplified with primers containing restriction sites of Type IIS enzyme BspQI (isoschizomer of SapI) to generate AGT and GCA overhangs. BspQI restriction sites generating the same overhangs were introduced into the backbones of vector pDX_init for phage display and pNb_init for periplasmatic expression and attachment of Myc- and His-tag. Note that in pDX_init and pNb_init the BspQI restriction sites are part of the sybody open reading frame. Finally, sybodies/nanobodies are sub-cloned from pNb_init to the destiny vectors pBXNPH3 or pBXNPHM3 for periplasmic expression. Tag-less sybodies/nanobodies for structural biology purposes can be obtained by 3C protease cleavage. Importantly, the vector series permits for PCR-free subcloning once the sybodies have been inserted into phage display vector pDX_init.

Supplementary Figure 7:

Validation of sybody library design. Comparison of homology model of non-randomized convex sybody based on the coordinates of 1ZVH with the structure of selected convex sybody Sb_MBP#1 (determined in complex with MBP). CDR residues contributing to the hydrophobic core are highlighted as green sticks, randomized residues as sticks colored in yellow, orange and red for CDR1, CDR2 and CDR3, respectively.

Supplementary Figure 8:

Detailed analysis of sybody-MBP complex structures. (**a**) Structures of MBP (blue) in complex with Sb_MBP#1-3 (grey with CDR1, CDR2 and CDR3 in yellow, orange and red, respectively). The coordinates of MBP were used to perform superimposition. (**b**) Interaction of Sb_MBP#1 with MBP, shown along the MBP cleft from both sides. Sybody residues contacting MBP (distance ≤ 4 Å) are shown as sticks. (**c**) Detailed view of interacting residues of sybodies Sb_MBP#1-3. In the left panel, four randomized residues of CDR3 which are invariant among the three binders are labeled. (**d**) Sequence alignment of Sb_MBP#1-3. The CDR regions are underlined.

Supplementary Figure 9:

Biophysical analysis of sybody-MBP interactions. (**a**) SPR analysis of interaction between sybody Sb_MBP#1 (analyte) and biotinylated MBP (ligand), determined in triplicate for each analyte concentration. Concentrations of Sb_MBP#1: 0, 4.7, 14.1, 42.2, 126.7, 380 nM. Data were fitted with a 1:1 binding model to obtain k_{on}, k_{off} and $K_{D, kinetics}$. Inset shows binding equilibrium data to determine $K_{D, equilibrium}$. Sybodies Sb_MBP#2 and Sb_MBP#3 were analyzed accordingly. (**b**) Data table summarizing the values obtained from SPR analysis shown in (**a**). (**c**) Displacement of 500 nM Sb_MBP#1 bound to immobilized MBP by addition of increasing maltose concentrations was monitored using the Octet® RED96 System.

Supplementary Figure 10:

Analysis of sybodies raised against ABC transporter TM287/288. (**a**) Binding affinities of 31 sybodies belonging to the concave, loop and convex library were determined by kinetic SPR measurements using the ProteOn™ XPR36 Protein Interaction Array System in the presence and absence of ATP and using wildtype TM287/288 and the ATPase-deficient mutant TM287/288(E517A) as ligands. Binders which exhibit an affinity increase of at least three fold against TM287/288(E517A) in the presence of ATP were defined as state-specific and are marked in blue. (**b**) Phylogenetic trees of sybodies specific against TM287/288 as determined by ELISA. Note that some of the sybodies were not analyzed by SPR either due to low yields during purification or poor SPR data.

Supplementary Figure 11:

Sequence alignment of concave sybodies raised against TM287/288.

Supplementary Figure 12:

Sequence alignment of loop sybodies raised against TM287/288.

Supplementary Figure 13:

Sequence alignment of convex sybodies raised against TM287/288.

Supplementary Figure 14:

Fitted SPR raw data of sybodies raised against ENT1 and GlyT1.

Table 1 – Characterization of sybodies raised against ENT1 and GlyT1

Sybody	k_{on} [M^{-1}s^{-1}]	k_{off} [s^{-1}]	KD [M] (kinetics)	KD [M] (equilibrium)	ΔT_m(SPA-TS) [°C]	SPA signal (fold increase)
Sb_ENT1#1	1.86E+05	7.44E-03	4.00E-08		6.1	7
Sb_GlyT1#1	1.88E+05	5.77E-02	3.07E-07		0.9	1.2
Sb_GlyT1#2*	3.68E+04	9.28E-02	2.52E-06		1.5	1.1
Sb_GlyT1#3**				1.54E-07	1.7	1.1
Sb_GlyT1#4**				4.761E-07	2.1	1.5
Sb_GlyT1#5	4.54E+05	3.72E-02	8.19E-08		2.6	1.5
Sb_GlyT1#6***	1.00E+05	4.99E-05	4.94E-10		8.8	1.6
Sb_GlyT1#7***	2.01E+04	1.85E-04	9.18E-09		10	1.8

*K_d higher than highest concentration injected

**k_{on} and k_{off} outside of measurable range

*** single cycle measurement due to slow k_{off}

Supplementary Table 1 – Features of the three sybody libraries

Library	Template PDB entry/target	Binding interface in template	Length of CDR3	Number of randomized residues in library	Theoretical diversity of library
concave	3K1K/GFP	672 Å^2	6 aa	15	8.3 x 10^{17}
loop	3P0G/GPCR	901 Å^2	12 aa	16	4.3 x 10^{19}
convex	1ZVH/Lysozyme	533 Å^2	16 aa	18	2.8 x 10^{22}

Supplementary Table 2 – DNA sequences of framework sybodies and flanking regions for ribosome display

Framework sequence concave	CAGGTTCAGCTGGTTGAGAGCGGTGGTGGCCTGGTCCAAGCTGGCGGTTCGCTGCGTCTGAGCTGCGCCGCAAGCGGTTTCCCGGTGAGCAGCAGCACGATGACCTGGTATCGTCAGGCACCGGGCAAAGAACGTGAGTGGGTCGCGGCGATTTCCAGCTCTGGTAGCACCACGACCTACGCAGATTCTGTTAAGGGCCGCTTTACCATCAGCCGCGACAACGCGAAGAATACGGTCTATTTGCAGATGAATAGCCTGAAACCGGAAGATACCGCGGTTTACTACTGTACCGTGACCGTGGGTAGCACGTACACGGGCCAAGGTACCCAAGTGACTGTGAGC
Framework sequence loop	CAGGTTCAGCTGGTTGAGAGCGGTGGTGGCCTGGTCCAAGCTGGCGGTTCGCTGCGTCTGAGCTGCGCCGCAAGCGGTTTCCCGGTGAGCAGCAGCACGATGACCTGGTATCGTCAGGCACCGGGCAAAGAACGTGAGTGGGTCGCGGCGATTTCCAGCTCTGGTAGCACCACGACCTACGCAGATTCTGTTAAGGGCCGCTTTACCATCAGCCGCGACAACGCGAAGAATACGGTCTATTTGCAGATGAATAGCCTGAAACCGGAAGATACCGCGGTTTACTACTGTAACGTGAAAGACAGCGGTAGCTCCAGCAGCTCCTACGACTATTGGGGCCAAGGTACCCAAGTGACTGTGAGC
Framework sequence convex	CAAGTCCAGCTGGTGGAATCGGGTGGTGGTGGTAGCGTCCAGGCGGGTGGTAGCCTGCGTCTGAGCTGTGCGGCTAGCGGCTCTATTTCCAGCATCACGTACCTGGGCTGGTTTCGCCAGGCACCGGGCAAAGAGCGTGAGGGCGTCGCAGCGCTGAGCACCAGCTCCGGTACCACCTACTACGCGGACAGCGTTAAGGGTCGTTTCACGGTGAGCCTGGACAACGCCAAGAATACCGTGTATCTGCAAATGAACAGCTTGAAACCGGAAGATACTGCTTTGTATTACTGCGCGGCAGCCAGCAGCGGCTCCAGCAGCCCGCTGTCTAGCAGCAGCTATACGTACTGGGGTCAGGGCACCCAAGTTACCGTTTCT
5' flank ribosome display	TAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATATCCATGGGTAGT
3' flank ribosome display	GCAAAGCTTTATATGGCCTCGGGGGCCGAATTCGGATCTGGTGGCCAGAAGCAAGCTGAAGAGGCGGCAGCGAAAGCGGCGGCAGATGCTAAAGCGAAGGCCGAAGCAGATGCTAAAGCTGCGGAAGAAGCAGCGAAGAAAGCGGCTGCAGACGCAAAGAAAAAAGCAGAAGCAGAAGCCGCCAAAGCCGCAGCCGAAGCGCAGAAAAAGCCGAGGCAGCCGCTGCGGCACTGAAGAAGAAAGCGGAAGCGGCAGAAGCAGCTGCAGCTGAAGCAAGAAAGAAAGCGGCAACTGAAGCTGCTGAAAAAGCCAAAGCAGA

```
AGCTGAGAAGAAAGCGGCTGCTGAAAAGGCTGCAGCTGATAAGAAAGCGGCAGCAGAGAAAGCTGCAGCCGACAAAAAAG
CAGCAGAAAAAGCGGCTGCTGAAAAGGCAGCAGCTGATAAGAAAGCAGCGGCAGAAAAAGCCGCCGCAGACAAAAAAGCG
GCAGCGGCAAAAGCTGCAGCTGAAAAAGCCGCTGCAGCAAAAGCGGCCGCAGAGGCAGATGATATTTTCGGTGAGCTAAG
CTCTGGTAAGAATGCACCGAAAACGGGGGGAGGGGCGAAAGGGAACAATGCTTCGCCTGCCGGGAGTGGTAATACTAAAA
ACAATGGCGCATCAGGGGCCGATATCAATAACTATGCCGGGCAGATTAAATCTGCTATCGAAAGTAAGTTCTATGACGCA
TCGTCCTATGCAGGCAAAACCTGTACGCTGCGCATAAAACTGGCACCCGATGGTATGTTACTGGATATCAAACCTGAAGG
TGGCGATCCCGCACTTTGTCAGGCTGCGTTGGCAGCAGCTAAACTTGCGAAGATCCCGAAACCACCAAGCCAGGCAGTAT
ATGAAGTGTTCAAAAACGCGCCATTGGACTTCAAACCGTAG
```

Supplementary Table 3 - Data collection and refinement statistics

	Sb_MBP#1 (PDB: 5M13)	Sb_MBP#2 (PDB: 5M14)	Sb_MBP#3 (PDB: 5M15)
Data Collection			
Space group	$P2_12_12_1$ (19)	$P2_12_12_1$ (19)	$P2_12_12_1$ (19)
Cell dimensions			
a, b, c (Å)	58.298	57.890	57.030
	82.789	57.950	57.780
	102.583	281.540	286.530
α, β, γ (º)	90.00	90.00	90.00
	90.00	90.00	90.00
	90.00	90.00	90.00
Resolution (Å)	50 - 1.37	50 - 1.6	50 - 1.9
R_{meas} (%) [1]	6.5 (60.9)	5.9 (124)	7.8 (146.6)
I/σ_I	15.26 (3.47)	16.98 (1.82)	21.44 (2.17)
$CC_{1/2}$ (%)	99.9 (86.3)	99.9 (68.5)	100 (60.5)
Completeness (%)	99.4 (97.7)	100 (100)	100 (100)
Redundancy	6.1	6.5	12.9
Refinement			
Resolution (Å)	50 - 1.37	50 - 1.6	50 - 1.9
No. reflections (work/test)	102618/5131	126118/6307	75931/3797
R_{work}/ R_{free} (%)	16.82/18.60	19.04/21.56	20.92/ 25.70
No. atoms			
Protein	3873	7640	7619
Water	694	1040	422
B-factor ($Å^2$)			
Total	20.1	34.4	50.1
R.m.s deviations			
Bond lengths (Å)	0.005	0.003	0.003
Bond angles (°)	0.750	0.591	0.623

[1] Values in parentheses are for the last resolution shell

Supplementary Table 4: List of Primers

Primers for library assembly	
(triplets designated 111, 222 and 333 correspond to the trinucleotide mixtures 1-3 for randomization; all primers in 5' to 3' orientation))	
CDR1_a_b	GCA AGC GGT TTC CCG GTG 111 111 111 222 ATG 333 TGG TAT CGT CAG GCA CCG G
CDR1_c	C TGT GCG GCT AGC GGC 111 ATT 111 111 ATC 222 TAC CTG GGC TGG TTT CGC C
CDR2_a_b	GA AGA CCT GTC GCG GCG ATT 111 AGC 111 GGT 111 222 ACG 333 TAC GCA GAT TCT GTT AAG GGC CG
CDR2_c	CGA AGA CCT GCA GCG CTG 111 ACC 111 111 GGT 222 ACC TAC TAC GCG GAC AGC G
CDR3_a	GA AGA CCT GCG GTT TAC TAC TGT 333 GTG 222 GTG GGT 111 222 TAC 333 GGC CAA GGT ACC CAA GTG AC
CDR3_b	CGC GAA GAC CTC GTG AAA GAC 111 GGT 111 111 111 111 111 TAC GAC TAT TGG GGC CAA GGT ACC CAA GTG AC
CDR3_c	GAA GAC CTC TGC GCG GCA GCC 111 111 GGC 111 111 111 CCG CTG 111 111 111 111 TAT 222 TAC TGG GGT CAG GGC ACC CAA GTT ACC GTT TCT
FW1_a_b_for	CAG GTT CAG CTG GTT GAG AGC
FW1_a_b_rev	CAC CGG GAA ACC GCT TGC
FW1_c_for	CAA GTC CAG CTG GTG GAA TCG
FW1_c_rev	GCC GCT AGC CGC ACA G
FW2_a_b_rev	ATG CAT GGT CTC ACG ACC CAC TCA CGT TCT TTG CCC GGT GCC TGA CGA TAC CA
FW2_c_rev	ATG CAT GGT CTC ACT GCG ACG CCC TCA CGC TCT TTG CCC GGT GCC TGG CGA AAC CAG CCC AGG
FW3_a_b_for	CGC AGA TTC TGT TAA GGG CCG
FW3_c_for	ACC TAC TAC GCG GAC AGC G
FW4_a_b_rev	GCT CAC AGT CAC TTG GGT ACC TTG GCC
FW4_c_rev	AGA AAC GGT AAC TTG GGT GCC CTG
Link1_a_b_for	ATG CAT GAA GAC CTG TCG CGG CG
Link1_a_b_rev	ATG CAT GGT CTC ACG ACC CAC
Link1_c_for	TAT ATC GAA GAC CTG CAG CGC TG
Link1_c_rev	ATG CAT GGT CTC ACT GCG ACG
Link2_a_for	TAT ATC GAA GAC CTG CGG TTT ACT ACT G
Link2_a_rev	ATG CAT GGT CTC ACC GCG GTA TCT TCC GGT TTC
Link2_b_for	ATG CAT GGT CTC ACC GCG GTA TCT TCC GGT TTC
Link2_b_rev	ATG CAT GGT CTC ACA CGT TAC AGT AGT AAA CCG CGG
Link2_c_for	ATA TAT GAA GAC CTC TGC GCG GC
Link2_c_rev	ATG CAT GGT CTC AGC AGT AAT ACA AAG CAG TAT CTT CCG G
Primers for vector construction	
pBXNPH3_#1	CAG CAG TCC GGC AGC AGC GGT CGG CAG CAG GTA TTT CAT GGT TAA TTC CTC CTG TTA GCC
pBXNPH3_#2	CTC CTC GCT GCC CAG CCT GCA ATG GCC GCA GAT CAC CAT CAT CAT CAC CAT CAT CAT CAT CAT TTA
pBXNPHM3_#1	ATA TAT GCG GCC GCC ATA GTG ACT GGA TAT GTT G
pBXNPHM3_#2	CAT GGT TAA TTC CTC CTG TTA GCC CAA AAA
pBXNPHM3_#3	AAA TAC CTG CTG CCG ACC GCT GCT GCT GGT
pBXNPHM3_#4	ATA TAT GCG GCC GCA TTA GGC ACC CCA GGC TTT A

31

pBXNPH3_blunt_for	CTC ATG ACC AAA ATC CCT TAA CGT GAG
pBXNPH3_EcoRI_rev	ATA TAT GAA TTC ATG GGG AGA CCC CAC ACT AC
pDX_init_#1	ATA TAT GCT CTT CAA GCG GAA GAG AGC CCA ATA CGC AAA CCG
pDX_init_#2	CGT TAG TAA ATG AAT TTT CTG TAT GAG GTT TTG
pDX_init_#3	GAA CCT GAA GCC CAG TAC CCG TAC
pDX_init_#4	CGT ACG GGT ACT GGG CTT CAG GTT
pDX_init_#5	TAT AAC TTG AAG AGC CGG CTG CCA TGG CCG GCT GGG CC
pDX_init_#6	TAT AGC AGG AAG AGC TCA CCA CCA TCA CCA TCA CGA ACC TG
pDX_init_#7	TAT AGC TCT TCA AGT CTG CCC ACA TAT ACC TGC CGT TC
pDX_init_#8	TAT AGC TCT TCC TGC AGA CAC GTG TCA CGT GAG GCC
Cm_EcoRI_for	GCT CAT GAA TTC CCC GCG CG
Cm_blunt_rev	GTG CAA TGT AAC ATC AGA GAT TTT GAG ACA C
Nb_init_for	ATG CAG GAA GAG CTG GCG AAC AAA AAC TCA TCT CAG AAG AGG ATC TG
Nb_init_rev	ATA CTT GAA GAG CCG GCC ATT GCA GGC TGG GCA G
RD_FX_pRDV_for	ATA TAT GCT CTT CTG CAA AGC TTT ATA TGG CCT CGG GGG C
RD_FX_pRDV_rev1	TAT ATA GCT CTT CAA CTA CCC ATG GAT ATA TCT CCT TCT TAA AGT TAA AC
pRDV_SL_for	AGA CCA CAA CGG TTT CCC TCT AGA AAT AAT TTT GTT TAA CTT TAA G
pRDV_SL_rev	CCC TAT AGT GAG TCG TAT TAA TTT CGA TGG
GS-Linker_FW	GGC GGT GGC GGT AGT AGA AGA GCG AGC TGC AGA CTG
GS-Linker_RV	GCC GGA ACC ACT TGG ACC TTG AAA CAA AAC TTC TAA ATG ATG

Primers for target amplification

GFP_FX_FW	TAT AGC TCT TCT AGT CAA TTC AGC AAA GGA GAA GAA CTT TTC
GFP_FX_RV	TAT AGC TCT TCT TGC TGC ACT AGT TTT GTA GAG CTC ATC C
MBP_FX_FW	ATA TAT GCT CTT CTA GTA AAA TCG AAG AAG GTA AAC TGG TAA TCT GG
MBP_FX_RV	TAT ATA GCT CTT CAT GCG CTA CCC GGA GTC TGC GC

Primers for ribosome and phage display

Med_FX_for	ATA TGC TCT TCT AGT CAG GTT CAG CTG GTT GAG AGC G
Med_FX_rev	TAT AGC TCT TCA TGC GCT CAC AGT CAC TTG GGT ACC
Long_FX_for	ATA TGC TCT TCT AGT CAA GTC CAG CTG GTG GAA TCG
Long_FX_rev	TAT AGC TCT TCA TGC AGA AAC GGT AAC TTG GGT GCC C
RT_Primer	CTT CAG TTG CCG CTT TCT TTC TTG
Medium_ORF_for	AGT CAG GTT CAG CTG GTT GAG AGC G
Medium_ORF_rev	TGC GCT CAC AGT CAC TTG GGT ACC
Long_ORF_for	AGT CAA GTC CAG CTG GTG GAA TCG
Long_ORF_rev	TGC AGA AAC GGT AAC TTG GGT GCC C
5'_flank _for	CGA AAT TAA TAC GAC TCA CTA TAG GGA GAC
tolAk_rev	CCG CAC ACC AGT AAG GTG TGC GGT TTC AGT TGC CGC TTT CTT TCT
tolAk_2	CCG CAC ACC AGT AAG GTG TGC
5'_flank _rev	TAT AGC TCT TCA ACT ACC CAT GGA TAT ATC CC
3'_flank_for	TAT AGC TCT TCT GCA AAG CTT TAT ATG GCC TC
Medium_ORF_5'_rev	CGC TCT CAA CCA GCT GAA CCT GAC T
Long_ORF_5'_rev	CGA TTC CAC CAG CTG GAC TTG ACT
Medium_ORF_3'_for	GGT ACC CAA aGTG ACT GTG AGC GCA
Long_ORF_3'_for	GGG CAC CCA AGT TAC CGT TTC TGC A

Primers for qPCR

qPCR_RD_5'_for	GGG AGA CCA CAA CGG TTT CCC

32

qPCR_ RD_S&M_5'_rev	CAC CGG GAA ACC GCT TGC GGC
qPCR_ RD_L_5'_rev	GCC GCT AGC CGC ACA GCT C
qPCR_ RD_tolA_3'_for	GCC GAA TTC GGA TCT GGT GGC
qPCR_ RD_tolA_3'_rev	CTG CTT CTT CCG CAG CTT TAG C
qPCR_PD_pDX_for	GAC GTT CCG GAC TAC GGT TCC
qPCR_PD_pDX_rev	CAC AGA CAG CCC TCA TAG TTA GC
qPCR_3K1K_for	AGT GCC GGT GAT CGT AGC AG
qPCR_3K1K_rev	CCC AAT ATT CAA AGC CCA CGT T

Supplementary Table 5: Primers and megaprimers used to assembly the sybody libraries

Library	CDR	Randomized primer	Megaprimer	Assembly primer	Outer primers
concave	CDR1	CDR1_a_b	Megaprimer 1	FW2_a_b_rev	FW1_a_b_for/ Link1_a_b_rev
	CDR2	CDR2_a_b	Megaprimer 2	-	Link1_a_b_for/ Link2_a_rev
	CDR3	CDR3_a	-	-	Link2_a_for/ FW4_a_b_rev
loop	CDR1	CDR1_a_b	Megaprimer 1	FW2_a_b_rev	FW1_a_b_for/ Link1_a_b_rev
	CDR2	CDR2_a_b	Megaprimer 3	-	Link1_a_b_for/ Link2_b_rev
	CDR3	CDR3_b	-	-	Link2_b_for/ FW4_a_b_rev
convex	CDR1	CDR1_c	Megaprimer 4	FW2_c_rev	FW1_c_for/ Link1_c_rev
	CDR2	CDR2_c	Megaprimer 5	-	Link1_c_for/ Link2_c_rev
	CDR3	CDR3_c	-	-	Link2_c_for/ FW4_c_rev

Supplementary Table 6 – Biophysical characterization of the three sybody-MBP complexes analyzed by the PISA server (http://www.ebi.ac.uk/pdbe/pisa/)

Sybody	Chains	Interface area	Number of hydrogen bonds	Number of salt bridges
Sb_MBP#1	A – B	1147 $Å^2$	18	9
Sb_MBP#2	A – C	1095 $Å^2$	8	6
	B – D	1111 $Å^2$	10	6
Sb_MBP#3	A – C	1128 $Å^2$	7	3
	B – D	1196 $Å^2$	11	3

Supplementary Note 1 - Establishment of the sybody framework and randomization strategy

The natural nanobodies 3K1K, 3P0G and 1ZVH, which all contain a single disulfide bond and have CDR3 lengths of 6, 12 and 16 amino acids (Supplementary Table 1) served as templates to create the concave, loop and convex sybody library, respectively. Apart from the CDR regions, 3K1K and 3P0G share high sequence identities and therefore the concave and the loop library share the same framework. In contrast, 1ZVH contains an extended hydrophobic core and the convex library was therefore built on a different scaffold. Supplementary Figure 2a shows an alignment of 3K1K and 3P0G along with the framework sequences of the concave and the loop library. In the alignment, residues of 3K1K and 3P0G differing from the framework sequence are depicted in blue, whereas randomized residues of the libraries are highlighted in red and orange. The CDRs contain aliphatic or aromatic framework residues marked in green, which according to the template structures point into the hydrophobic core and contribute to scaffold stability. Supplementary Figure 2b shows an alignment of 1ZVH along with the framework sequence of the convex library, which is labelled accordingly. Three non-randomized scaffold sybodies representing the concave, the loop and the convex library were generated by gene synthesis. Their DNA sequences are given in Supplementary Table 2 and they contain serines and threonines at positions to be randomized in the libraries (Supplementary Fig. 2). The non-randomized scaffold sybodies were expressed in vector pBXNPH3 in the periplasm of *E. coli*, purified by Ni^{2+}-NTA chromatography and analyzed by size exclusion chromatography (SEC) (Supplementary Fig. 3a). The three framework sybodies eluted as monodisperse monomers from SEC. Thermal unfolding using the fluorescent dye SYPRO Orange was performed and revealed melting temperatures of 74 °C, 75 °C and 95 °C for the non-randomized concave, loop and convex sybody, respectively. Hence, sybodies are considerably more stable than their natural counterparts 3K1K (T_m = 39 °C, precursor of concave library) or 1ZVH (T_m = 74 °C, precursor of convex library). Thermal stability of the convex sybody is particularly high. We attributed this increased stability to the extended hydrophobic core to tether down CDR3 and the V51L substitution introduced into the framework prior to the CDR2 region.

The CDRs of the three libraries concave, loop and convex library were randomized at positions visualized in Supplementary Figure 1 and specified in Supplementary Figure 2. The libraries were constructed using trinucleotide phosphoramidites for the synthesis of oligonucleotides, which permitted a defined amino acid composition in each randomized position. Three trinucleotide mixes were used for randomized residues placed 1) in loops, 2) at the transitions from loop to β-sheet and 3) in the middle of β-sheets. In Supplementary Figure 2a, the randomized positons are depicted as red colored S (for mix 1), red colored T (for mix 2) and orange colored T (for mix 3). The amino acid compositions of the three mixes are provided in Supplementary Figure 2b. Cysteines and prolines

35

were generally excluded in any of the three mixes. Mix 1 is enriched by the residues A, S, T, N, Y (10.6 %, each), contains D, E, Q, R, K, H, W at 5 % frequency each and harbors only few of the apolar amino acids F, M, V, I, L, G (2 %, each). Mix 2 lacks amino acids D and A, because these two residues are underrepresented at the end of β-sheets[39]. Mix 3 is devoid of D, N, Q, G, S and M, because these amino acids are found less frequently in the middle of β-sheets[39]. The theoretical diversity of the libraries amounts to 8.3×10^{17}, 4.3×10^{19}, and 2.8×10^{22} for the concave, loop and convex library, respectively. Three DNA fragments each containing one CDR of each of the three libraries were generated by assembly PCR. The resulting fragments were ligated in two subsequent steps using Type IIS restriction enzymes analogous to the assembly of designed ankyrin repeat proteins (details see materials and methods)[14]. In a next step, the three sybody libraries were flanked with the required sequence elements for *in vitro* transcription and ribosome display. Instead of cloning the library via BamHI/ HindIII into pRDV followed by PCR amplification of the flanked library, as it was reported previously for DARPins[15], we amplified the flanking regions by PCR and ligated them to the sybody libraries using the Type IIS restriction enzyme BspQ1. An experimental library diversity of 9×10^{12} for each of the three libraries was determined.

In recent years, the generation of two synthetic nanobody libraries have been described[34,35]. Unfortunately, the corresponding studies do not contain biophysical analyses, which would permit to compare thermal stability and structural integrity of these synthetic nanobodies with our in-depth characterized sybodies. Importantly, the CDR3s of both reported libraries have a length up to 18 amino acids and are therefore comparable to our convex library containing 16 residues. However, the authors of the respective studies did not design an extended hydrophobic core to tether down the long CDR3 as we did for the convex library nor did they stabilize it with a second disulfide bond as it is typically observed for natural nanobodies containing long CDR3 loops (see for example PDBs 1KXT, 1RJC and 3SN6, whose CDR3 lengths are 19). It must therefore be assumed that these synthetic nanobodies contain highly flexible CDR3 loops which are likely unsuited for structural biology applications. Further, the reported diversities of these synthetic nanobody libraries are in the range of 10^9-10^{10} due the generation of phage libraries. Hence, the diversity of our sybody libraries is 100-1000 times larger than any previously reported single domain antibody library.

36

Supplementary Note 2 - Ribosome display of sybodies and nanobodies

Ribosome display offers the advantage of displaying 10^{12} different library members with minor experimental effort. Since the preparation of home-made *in vitro* translation reagents is laborious and associated with variable levels of unfavorable Rnase acitivity[15], we used the commercial *in vitro* translation kit PURE*frex®SS* (GeneFrontier) and tested its applicability for ribosome display. The kit is devoid of reducing agents and contains oxidized glutathione (GSSG) and the disulfide bond isomerase DsbC and is thus suited to support the folding of disulfide-containing proteins such as nanobodies and sybodies. For transcription and ribosome display, the sybody library was flanked by sequence elements which attach a T7 promoter and a ribosome binding site to the 5' end and a *tolA* spacer containing a stem-loop to the 3' end. We experimentally tested display efficiency in two independent assays. In a first assay, we fused a 3xFLAG tag followed by 3C protease cleavage site to the C-terminus of the non-randomized loop sybody in a construct containing the flanking regions for transcription and ribosome display (Supplementary Fig. 4a). The mRNA of this construct was displayed on ribosomes using the PURE*frex®SS* kit, sybody-3xFLAG was cleaved from the nascent polypeptide chain by 3C protease and the entire protein mixtures was analyzed by Western blotting using an anti-FLAG antibody. A purified non-randomized convex sybody containing exactly the same 3xFLAG sequence and 3C protease cleavage site at the C-terminus served as standard for protein quantification by Western blotting. The analysis revealed that more than 70 % of the input mRNA was translated. To assure that ribosome display produces correctly folded binders, we displayed 10^6 mRNA encoding the 3K1K nanobody spiked into 10^{12} mRNA encoding the non-randomized convex sybody using PURE*frex®SS* kit and assessed binding to immobilized GFP (target of 3K1K) and MBP (dummy target) (Supplementary Fig. 4b). Quantitative PCR on reverse transcribed cDNA was used to determine the amount of mRNA which could be retrieved after binding and washing in comparison to the input mRNA added to the display reaction. For 3K1K panned against GFP, mRNA recovery was 80 %, while mRNA retrieved from 3K1K panning against MBP was not detectable. Background binding of the non-randomized convex sybody towards GFP and MBP was minimal (recovery fractions below 0.001 %). These two independent experiments clearly demonstrated that ribosome display of nanobodies works efficiently using the commercial PURE*frex®SS* kit.

Supplementary Note 3 – In depth crystallographic analysis of MBP sybodies

Three convex MBP-binders called Sb_MBP#1, Sb_MBP#2 and Sb_MBP#3 were co-crystallized with MBP. The sybodies were expressed in the pBXPNH3 construct and co-migrated with purified MBP on SEC (Supplementary Fig. 5). The complexes readily formed crystals, which were directly picked from the screening plate for X-ray data acquisition. Crystal structures were solved by molecular replacement and refined to resolutions ranging from 1.37 - 1.9 Å (Supplementary Table 3). The three nanobodies bind to the same epitope on MBP, which is located at the cleft between the two lobes of the protein (Fig. 2a and Supplementary Fig. 8) where maltose becomes trapped [17]. The buried sybody-MBP interfaces cover areas in the range of 1095 – 1196 $Å^2$ and contain variable numbers of hydrogen bonds and salt bridges (Supplementary Table 6). The interaction is mainly mediated by CDR3. Of particular importance are the CDR3 residues W101, Q104, S105 and W110 placed at randomized positions which are identical among the three binders (Supplementary Fig. 8). Further contacts are mediated by randomized residues of CDR1 and CDR2. However, these are highly variable. Of note, there are also several invariant framework residues participating in the protein-protein interaction. Overall, the synthetic nanobodies bind to MBP in an analogous fashion as their natural precursors, namely via interactions predominantly mediated by CDR3 residues[13,19]. Natural nanobodies containing CDR3 sequences of identical lengths and sequence identities greater than 80% are considered to stem from the same B-cell lineage[6]. Such nanobodies are considered to belong to the same family and to recognize the same epitope. In our synthetic convex library, the CDR3 is of identical length by design. While the CDR3 sequences of Sb_MBP#1 and Sb_MBP#2 are identical, they share only 63 % sequence identity with Sb_MBP#3. According to the criteria of natural nanobodies, Sb_MBP#3 would belong to a different binder family than Sb_MBP#1 and Sb_MBP#2. However, our structural analysis clearly shows that they recognize exactly the same epitope.

An important property of nanobodies is their ability to trap target proteins in specific conformational states[7,40]. Structures of MBP determined in the presence or absence of maltose revealed that maltose binding is associated with a large conformational rotation around a hinge between the two lobes of MBP resulting in a closure of the cleft between the lobes and thereby surrounding the sugar from both sides[17]. The structure of MBP crystallized in complex with sybodies corresponds to the ligand-free conformation; it deviates by a RMSD of 0.507 from the ligand-free structure (PDB: 1OMP) versus a RMSD of 3.786 from the maltose-bound structure (PDB: 1ANF). This finding agrees with the experimental conditions of the sybody selections; since the buffers did not contain maltose, MBP was primarily presented in its open conformation. Despite of the open conformation of MBP, Sb_MBP#1 is in contact with 14 out of the 20 side chains, which otherwise interact with maltose in the sugar-

bound structure. Therefore, Sb_MBP#1 stabilizes the unliganded form of MBP and thereby prevents cleft-closure and maltose binding.

The sybody structures also permitted to validate the design of the convex library. Supplementary Figure 7 shows a homology model of the convex non-randomized scaffold sybody carrying serines and threonines in the randomized positions along with the structure of the MBP-specific convex sybody Sb_MBP#1 determined in complex with MBP. A set of CDR residues was designed as invariable framework residues contributing to the hydrophobic core of the scaffold (marked in green in Supplementary Fig. 7). Homology model and structure are highly similar, in particular with regard to the scaffolding residues contributed by the CDRs. The largest differences were observed in the randomized residues of CDR3, whose structure likely varies among different convex sybodies due to loop flexibility.

REFERENCES

1   Lee, Y. C. *et al.* Inhibition of cell adhesion by a cadherin-11 antibody thwarts bone metastasis. *Molecular cancer research : MCR* **11**, 1401-1411, (2013).

2   Ahuja, S. *et al.* Structural analysis of bacterial ABC transporter inhibition by an antibody fragment. *Structure* **23**, 713-723, (2015).

3   Blanpain, C. *et al.* Multiple active states and oligomerization of CCR5 revealed by functional properties of monoclonal antibodies. *Mol Biol Cell* **13**, 723-737, (2002).

4   Doshi, R. *et al.* In vitro nanobody discovery for integral membrane protein targets. *Scientific reports* **4**, 6760, (2014).

5   Dominik, P. K. *et al.* Conformational Chaperones for Structural Studies of Membrane Proteins Using Antibody Phage Display with Nanodiscs. *Structure* **24**, 300-309, (2016).

6   Pardon, E. *et al.* A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc* **9**, 674-693, (2014).

7   Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175-180, (2011).

8   Ehrnstorfer, I. A., Geertsma, E. R., Pardon, E., Steyaert, J. & Dutzler, R. Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport. *Nat Struct Mol Biol* **21**, 990-996, (2014).

9   Bradbury, A. R. M., Sidhu, S., Dubel, S. & McCafferty, J. Beyond natural antibodies: the power of *in vitro* display technologies. *Nat Biotechnol* **29**, 245-254, (2011).

10   Hanes, J. & Plückthun, A. *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **94**, 4937-4942, (1997).

11   Binz, H. K. *et al.* High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* **22**, 575-582, (2004).

12   Kirchhofer, A. *et al.* Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* **17**, 133-138, (2010).

13   De Genst, E. *et al.* Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc Natl Acad Sci U S A* **103**, 4586-4591, (2006).

14   Seeger, M. A. *et al.* Design, construction, and characterization of a second-generation DARPin library with reduced hydrophobicity. *Protein science : a publication of the Protein Society* **22**, 1239-1257, (2013).

15   Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display: selecting and evolving proteins *in vitro* that specifically bind to a target. *Nat Methods* **4**, 269-279, (2007).

16   Geertsma, E. R. & Dutzler, R. A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* **50**, 3272-3278, (2011).

17   Spurlino, J. C., Lu, G. Y. & Quiocho, F. A. The 2.3-Å Resolution Structure of the Maltose-Binding or Maltodextrin-Binding Protein, a Primary Receptor of Bacterial Active-Transport and Chemotaxis. *J Biol Chem* **266**, 5202-5219, (1991).

18   Telmer, P. G. & Shilton, B. H. Insights into the conformational equilibria of maltose-binding protein by analysis of high affinity mutants. *J Biol Chem* **278**, 34555-34567, (2003).

19   Desmyter, A., Decanniere, K., Muyldermans, S. & Wyns, L. Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. *J Biol Chem* **276**, 26285-26290, (2001).

20   Hohl, M., Briand, C., Grütter, M. G. & Seeger, M. A. Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat Struct Mol Biol* **19**, 395-402, (2012).

21   Hohl, M. *et al.* Structural basis for allosteric cross-talk between the asymmetric nucleotide binding sites of a heterodimeric ABC exporter. *Proc Natl Acad Sci U S A* **111**, 11025-11030, (2014).

22   Timachi, M. H. *et al.* Exploring conformational equilibria of a heterodimeric ABC transporter. *eLife* **6**, (2017).

23   Lin, L., Yee, S. W., Kim, R. B. & Giacomini, K. M. SLC transporters as therapeutic targets: emerging opportunities. *Nat Rev Drug Discov* **14**, 543-560, (2015).

24   Cesar-Razquin, A. *et al.* A Call for Systematic Research on Solute Carriers. *Cell* **162**, 478-487, (2015).

25   Yang, C. & Leung, G. P. H. Equilibrative Nucleoside Transporters 1 and 4: Which One Is a Better Target for Cardioprotection Against Ischemia-Reperfusion Injury? *J Cardiovasc Pharm* **65**, 517-521, (2015).

26   Harvey, R. J. & Yee, B. K. Glycine transporters as novel therapeutic targets in schizophrenia, alcohol dependence and pain. *Nature Reviews Drug Discovery* **12**, 866-885, (2013).

27   Alberati, D. *et al.* Glycine reuptake inhibitor RG1678: a pharmacologic characterization of an investigational agent for the treatment of schizophrenia. *Neuropharmacology* **62**, 1152-1161, (2012).

28   Uysal, S. *et al.* Crystal structure of full-length KcsA in its closed conformation. *Proc Natl Acad Sci U S A* **106**, 6644-6649, (2009).

29   Kim, J., Stroud, R. M. & Craik, C. S. Rapid identification of recombinant Fabs that bind to membrane proteins. *Methods*, (2011).

30   Stockbridge, R. B., Koide, A., Miller, C. & Koide, S. Proof of dual-topology architecture of Fluc F- channels with monobody blockers. *Nature communications* **5**, (2014).

31   Seeger, M. A. *et al.* Tuning the drug efflux activity of an ABC transporter *in vivo* by *in vitro* selected DARPin binders. *PLoS One* **7** e37845, (2012).

32   Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O. & Grütter, M. G. Drug Export Pathway of Multidrug Exporter AcrB Revealed by DARPin Inhibitors. *PLoS Biol* **5**, e7, (2007).

33   Kodan, A. *et al.* Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc Natl Acad Sci U S A* **111**, 4049-4054, (2014).

34   Moutel, S. *et al.* NaLi-H1: A universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *eLife* **5**, (2016).

35   Yan, J. R., Li, G. H., Hu, Y. H., Ou, W. J. & Wan, Y. K. Construction of a synthetic phage-displayed Nanobody library with CDR3 regions randomized by trinucleotide cassettes for diagnostic applications. *J Transl Med* **12**, (2014).

36   Bertozzi, C., Zimmermann, I., Engeler, S., Hilf, R. J. C. & Dutzler, R. Signal Transduction at the Domain Interface of Prokaryotic Pentameric Ligand-Gated Ion Channels. *Plos Biol* **14**, (2016).

37   Bukowska, M. A. *et al.* A Transporter Motor Taken Apart: Flexibility in the Nucleotide Binding Domains of a Heterodimeric ABC Exporter. *Biochemistry* **54**, 3086-3099, (2015).

38   Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289-298, (2006).

39   Bhattacharjee, N. & Biswas, P. Position-specific propensies of amino acids in the β-strand. *Bmc Struct Biol* **10**, (2010).

40   Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535**, 448-+, (2016).
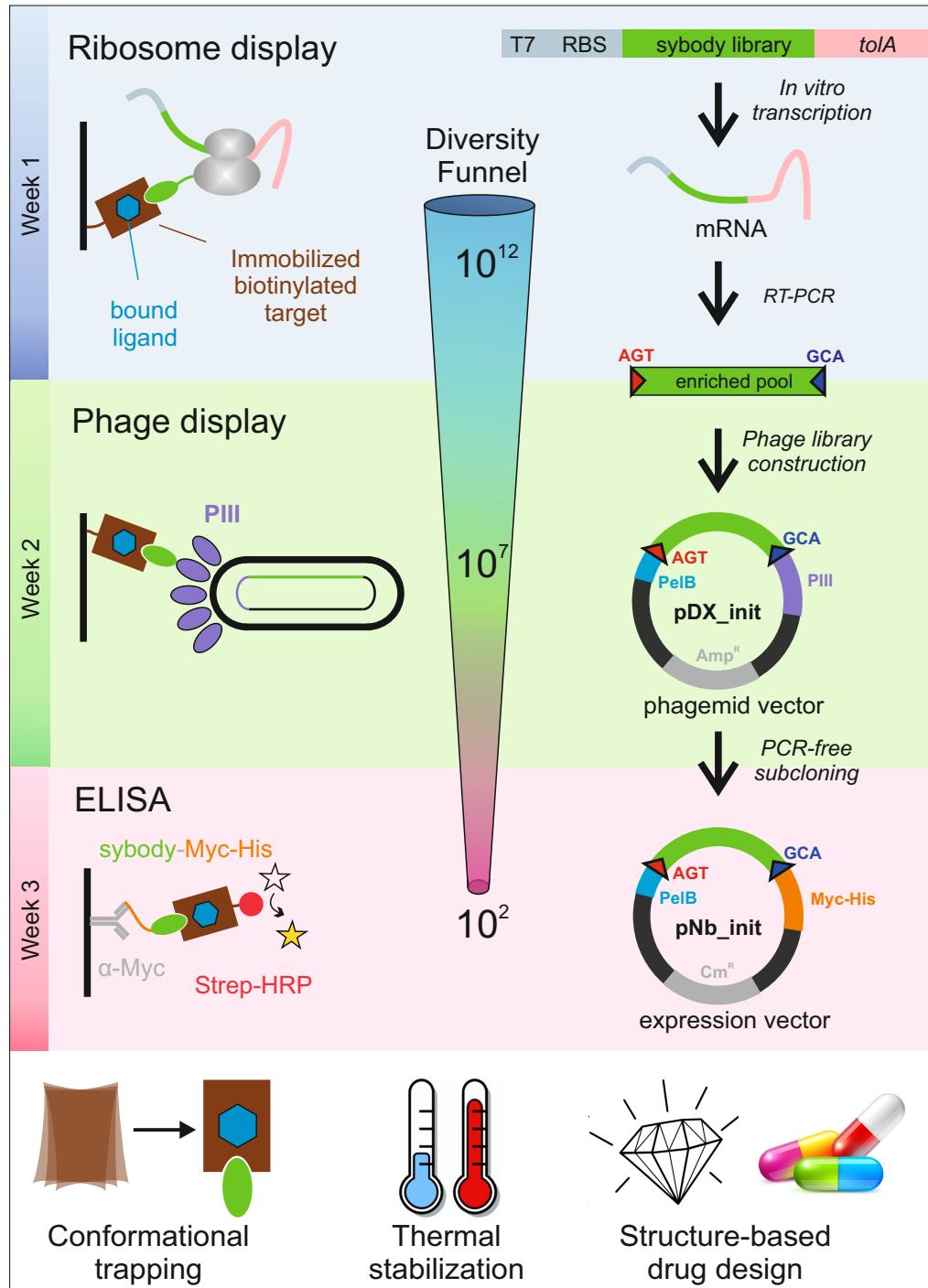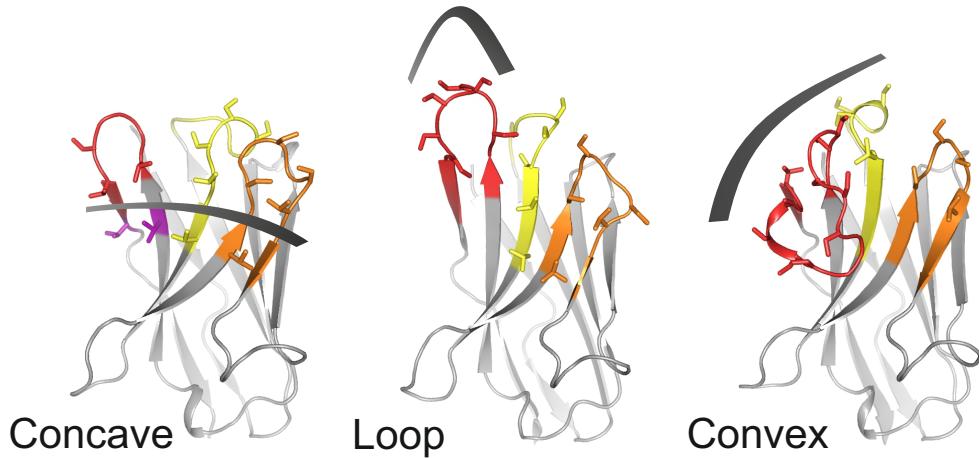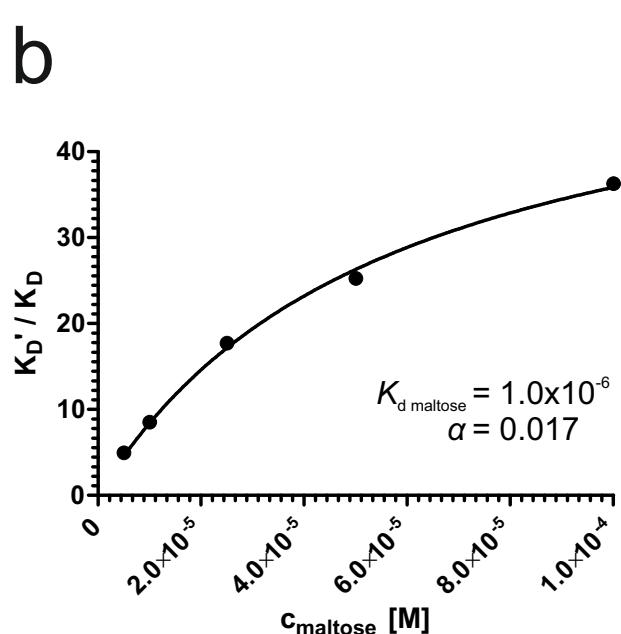
Figure 1
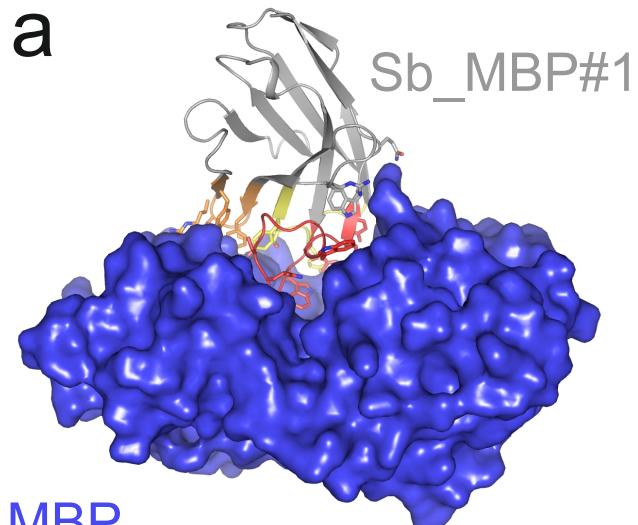
a

Sb_MBP#1

MBP

b

$K_{d\ maltose} = 1.0 \times 10^{-6}$

$\alpha = 0.017$

Figure 2

Figure 3

Figure 4

a

Extracellular space

Cytosol

b

c

d

e

f

g

Figure 5

A — GFP — 3K1K — Concave

B — GPCR — 3P0G — Loop

C — Lyso-zyme — 1ZVH — Convex
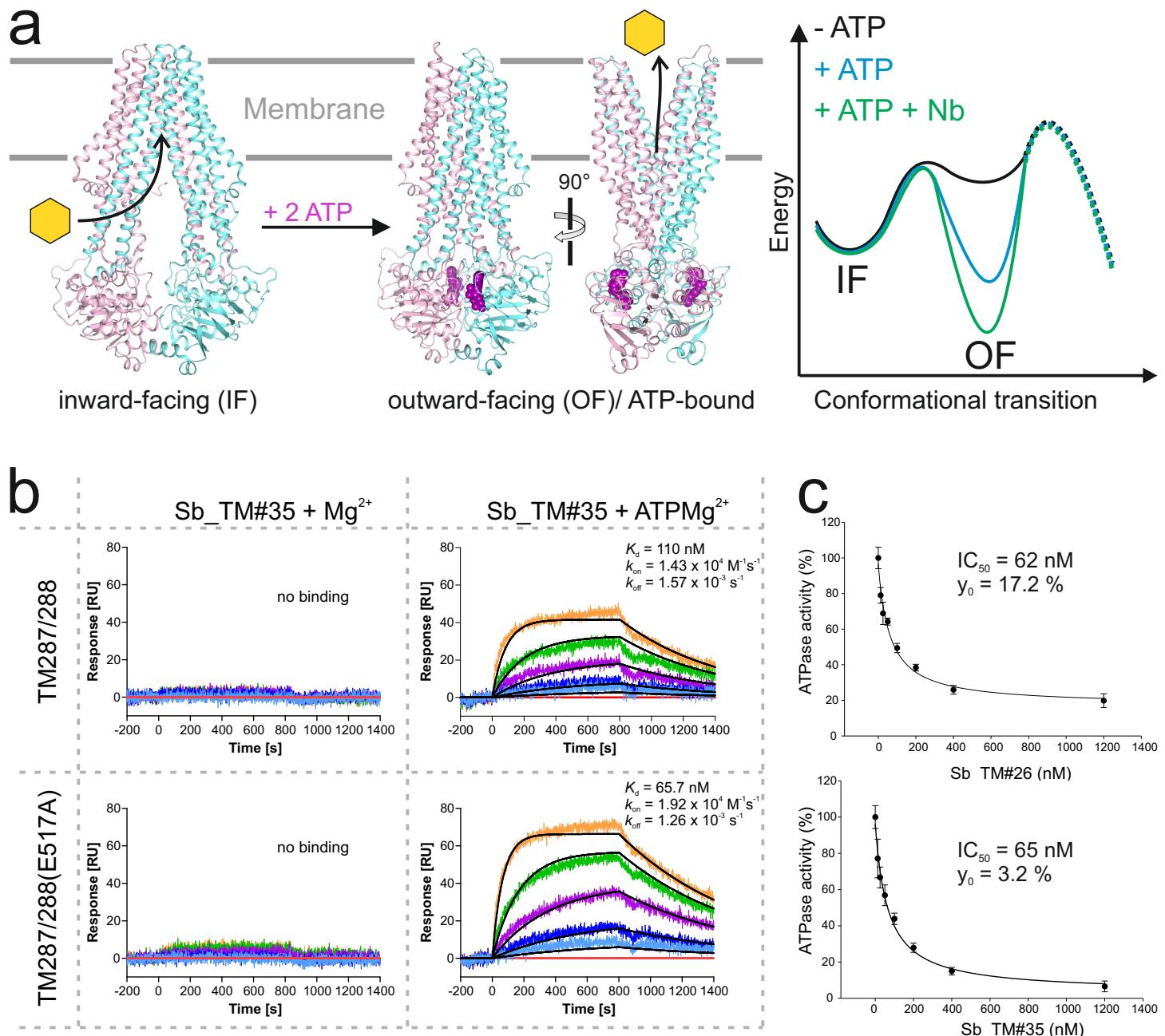
Supplementary Figure 1

a

```
3K1K      QVQLVESGGALVQPGGSLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGMSSAGDRSSY      60
Concave   QVQLVESGGGLVQAGGSLRLSCAASGFPVSSSTMTWYRQAPGKEREWVAAISSSGSTTY      60
3P0G      QVQLQESGGGLVQAGGSLRLSCAASGSIFSINTMGWYRQAPGKQRELVAAIHSGG-STNY      59
Loop      QVQLVESGGGLVQAGGSLRLSCAASGFPVSSSTMTWYRQAPGKEREWVAAISSSGSTTY      60
          ****  **** **  ************  ..    :* *********:** **.: *.*  :.*

3K1K      EDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYCNVNVG------FEYWGQGTQVTVS     114
Concave   ADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCTVTVG------STYTGQGTQVTVS     114
3P0G      ANSVKGRFTISRDNAANTVYLQMNSLKPEDTAVYYCNVKDYGAVLYEYDYWGQGTQVTVS     119
Loop      ADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCNVKDSGSSSSYDYWGQGTQVTVS     120
          :*********:*  ********************.*.          * *********
```

b

```
1ZVH      DVQLVESGGGSVQAGGSLRLSCAASGYIASINYLGWFRQAPGKEREGVAAVSPAGGTPYY      60
Convex    QVQLVESGGGSVQAGGSLRLSCAASGSISSITYLGWFRQAPGKEREGVAALSTSSGTTYY      60
          :*************************  *.:**. *******************:*  :.** **

1ZVH      ADSVKGRFTVSLDNAENTVYLQMNSLKPEDTALYYCAAARQGWYIPLNSYGYNYWGQGTQ     120
Convex    ADSVKGRFTVSLDNAKNTVYLQMNSLKPEDTALYYCAAASSGSSSPLSSSSYTYWGQGTQ     120
          **************:**********************.*   **.* .*.*******

1ZVH      VTVS     124
Convex    VTVS     124
          ****
```

c

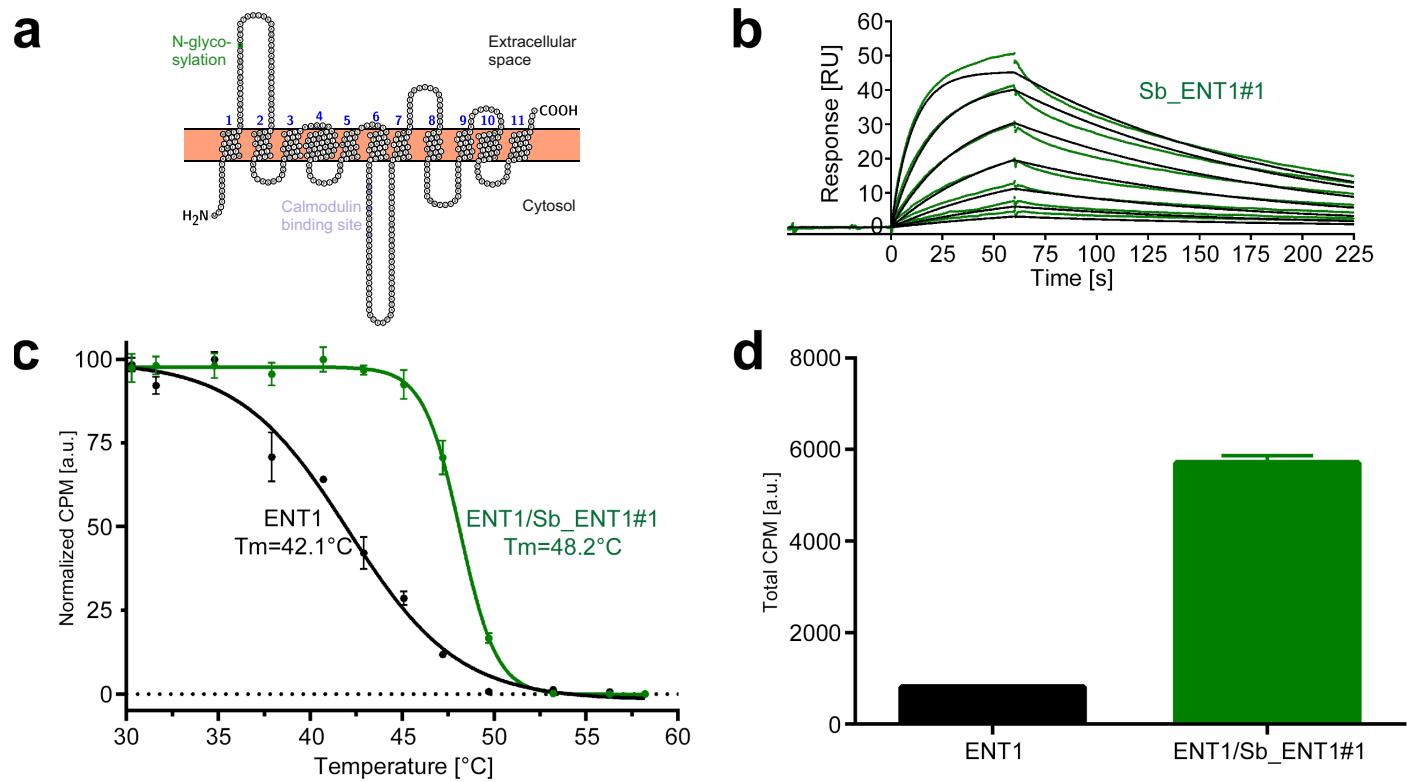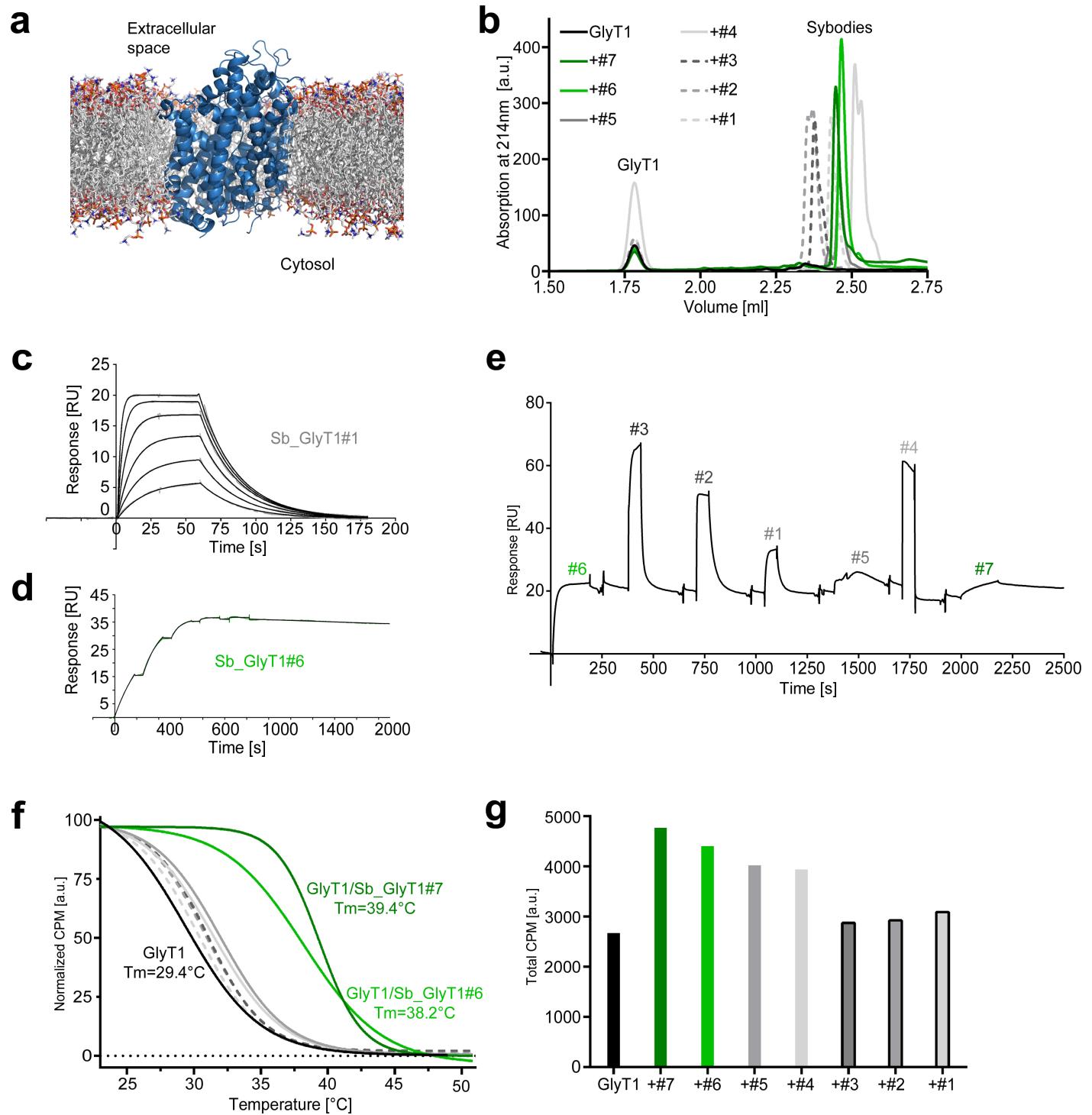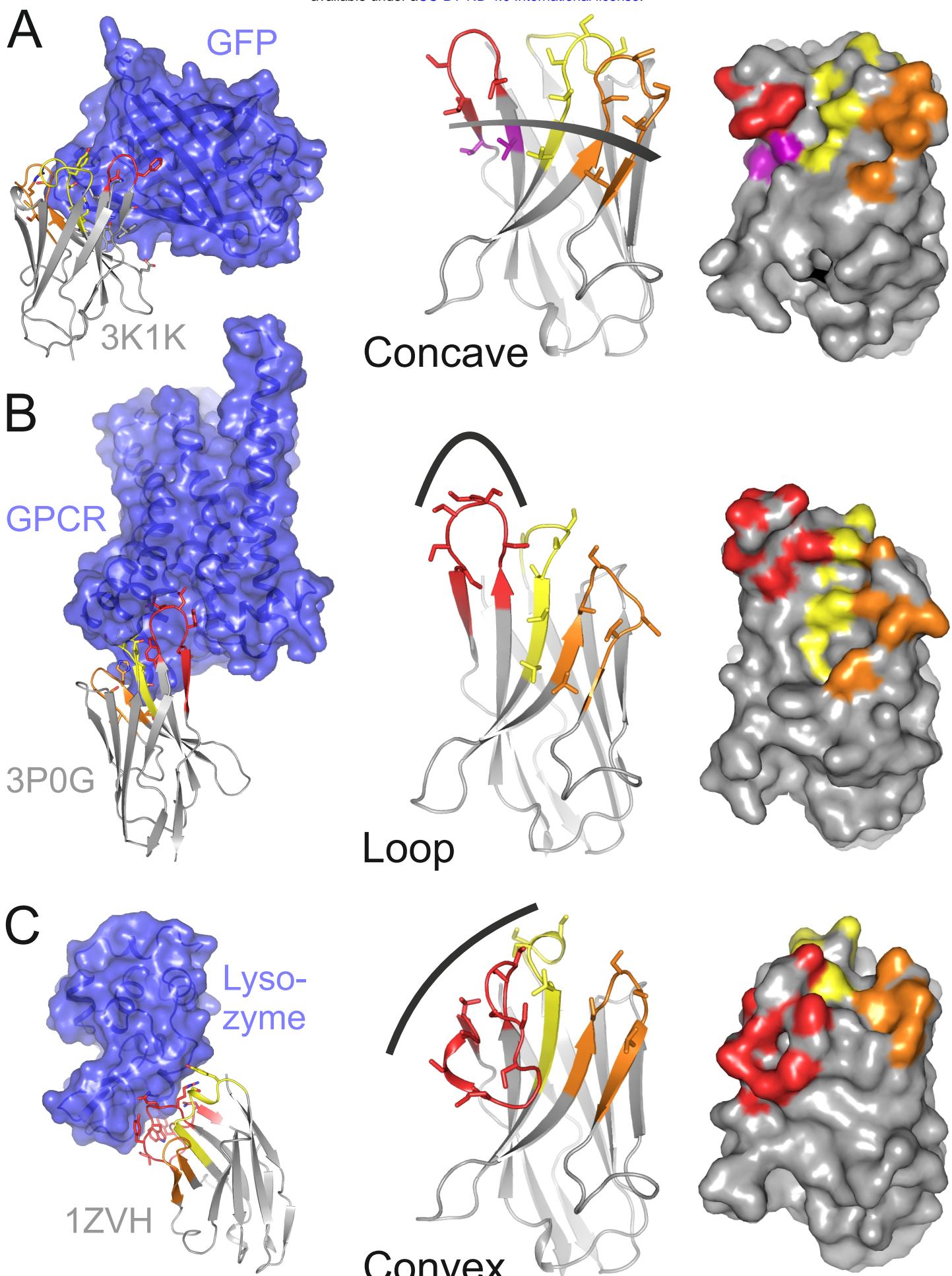| Amino acid	Amino acid composition in randomized positions (%)		
	Positions marked with red colored S	Positions marked with red colored T	Positions marked with orange colored T
K	5.00	7.50	11.00
R	5.00	7.50	11.00
H	5.00	7.50	11.00
D	5.00		
E	5.00	7.50	11.00
N	10.60	7.50	
Q	5.00	7.50	
S	10.60	10.00	
T	10.60	10.00	11.00
A	10.60		11.00
V	2.00	2.00	3.00
L	2.00	2.00	3.00
I	2.00	2.00	3.00
M	2.00	2.00	
F	2.00	2.00	3.00
Y	10.60	10.00	11.00
W	5.00	7.50	11.00
G	2.00	7.50	

Supplementary Figure 2
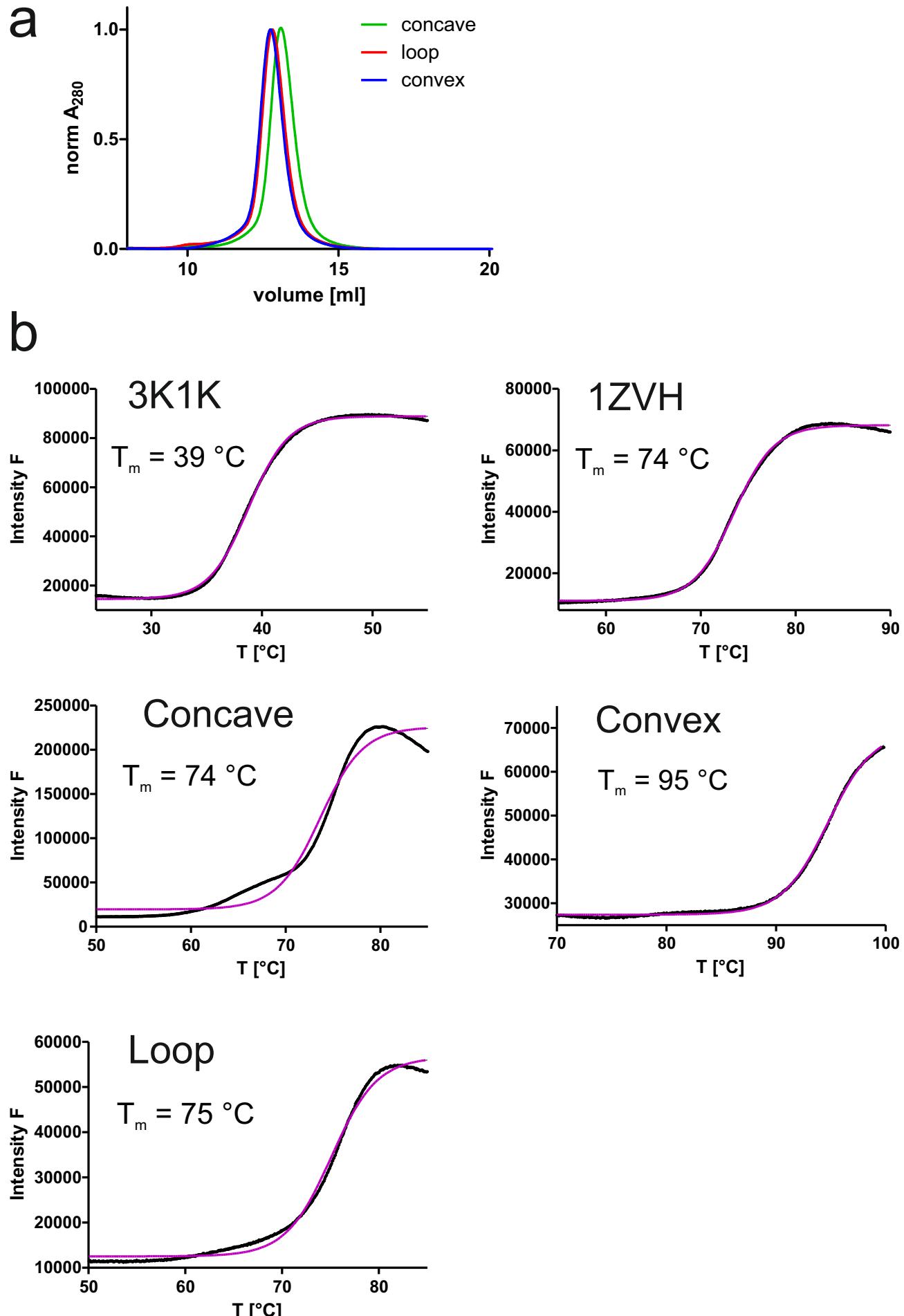
Supplementary Figure 3
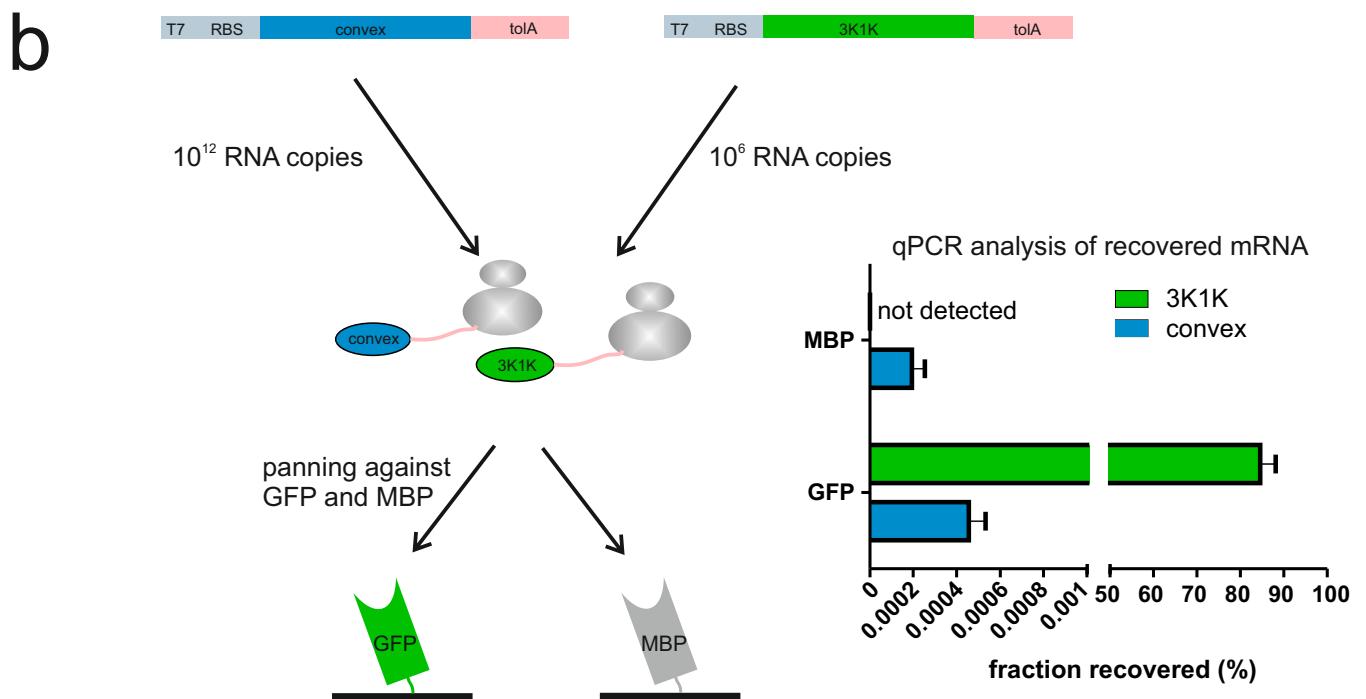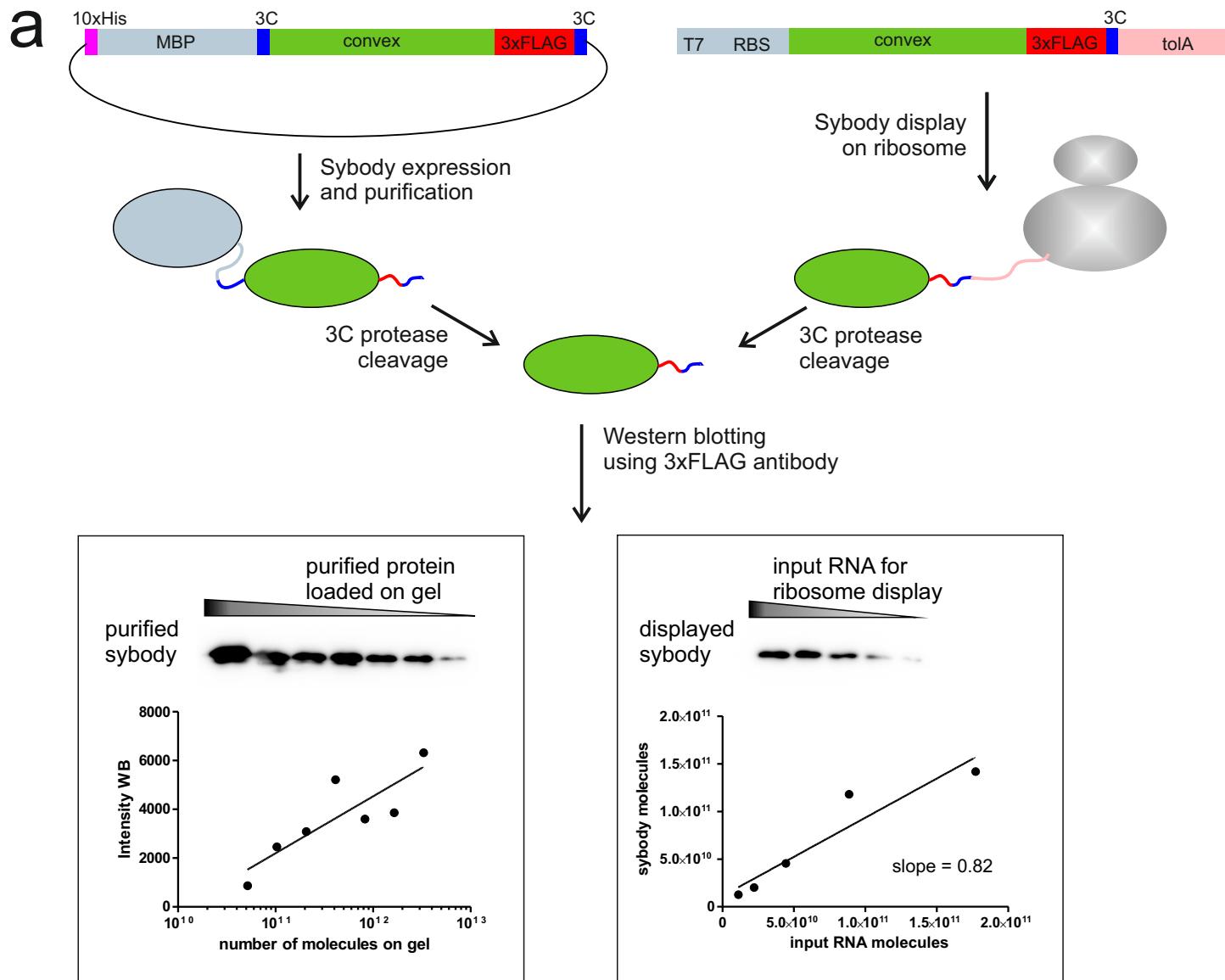
Supplementary Figure 4
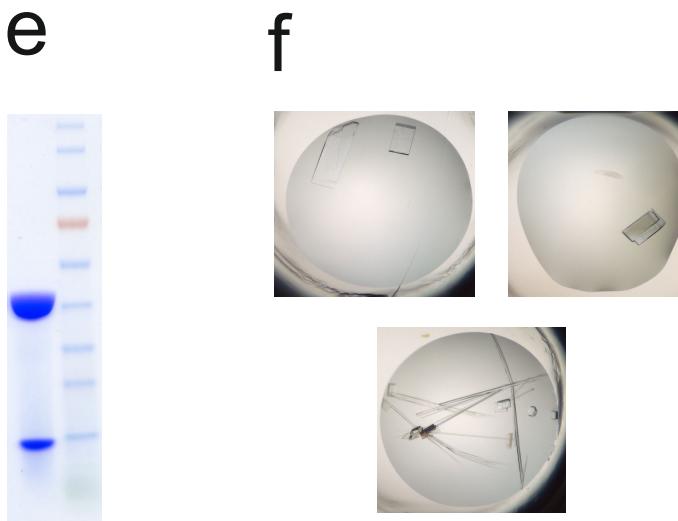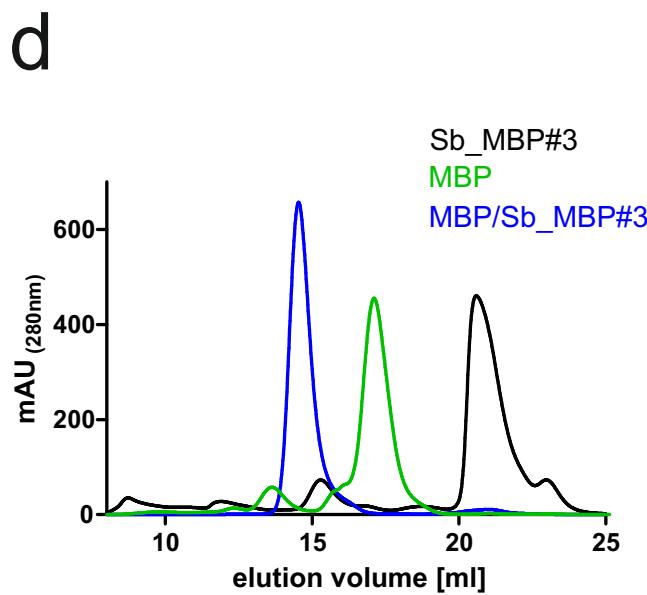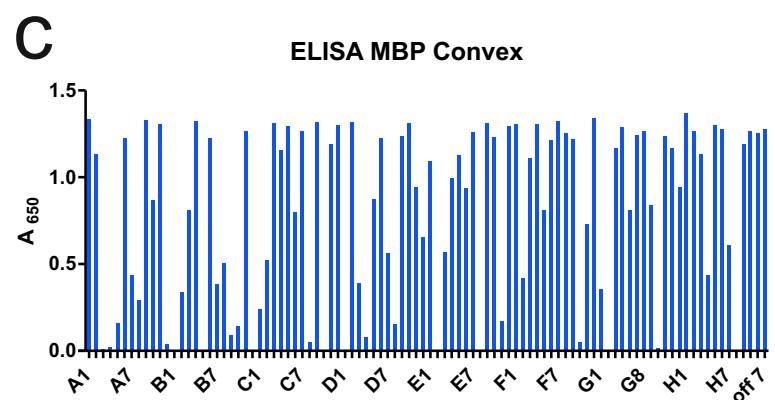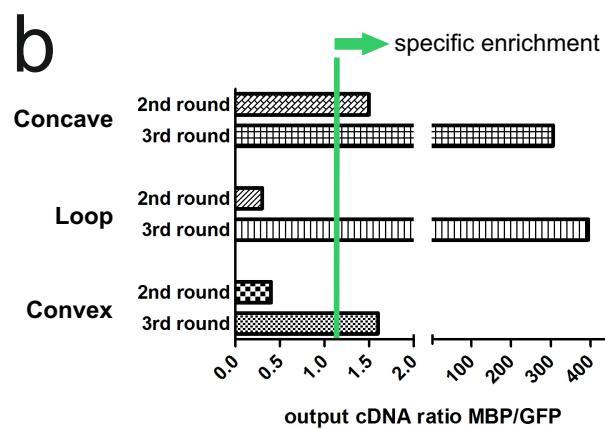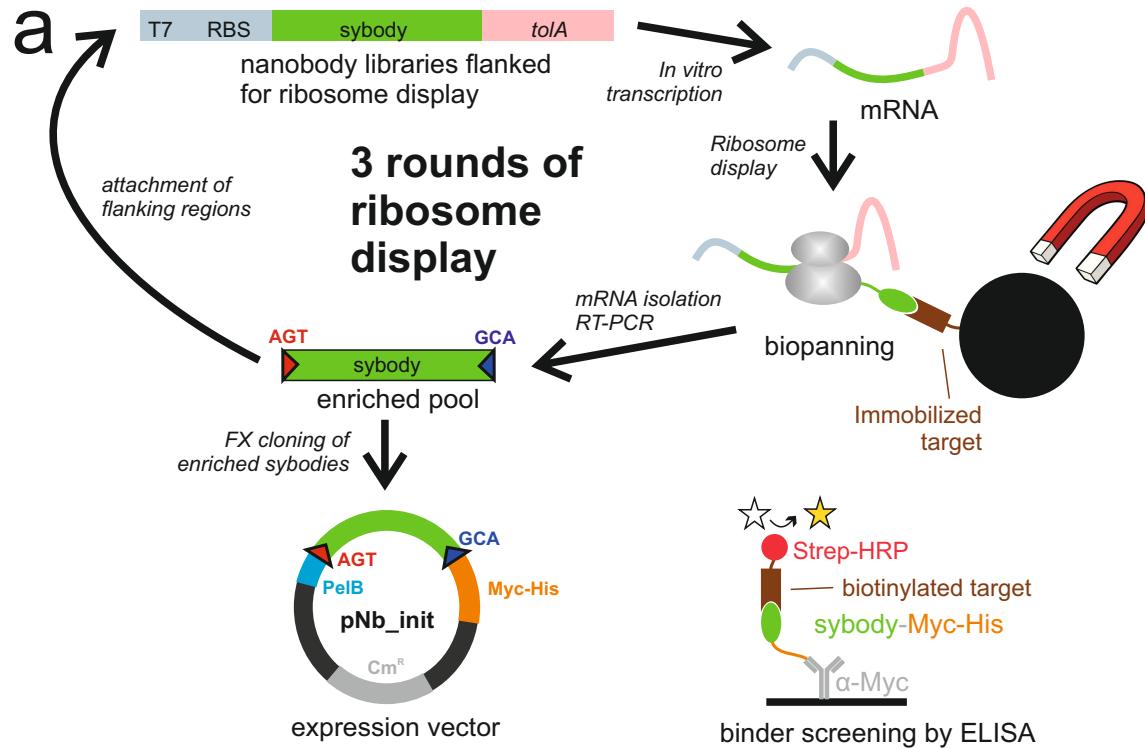
Supplementary Figure 5
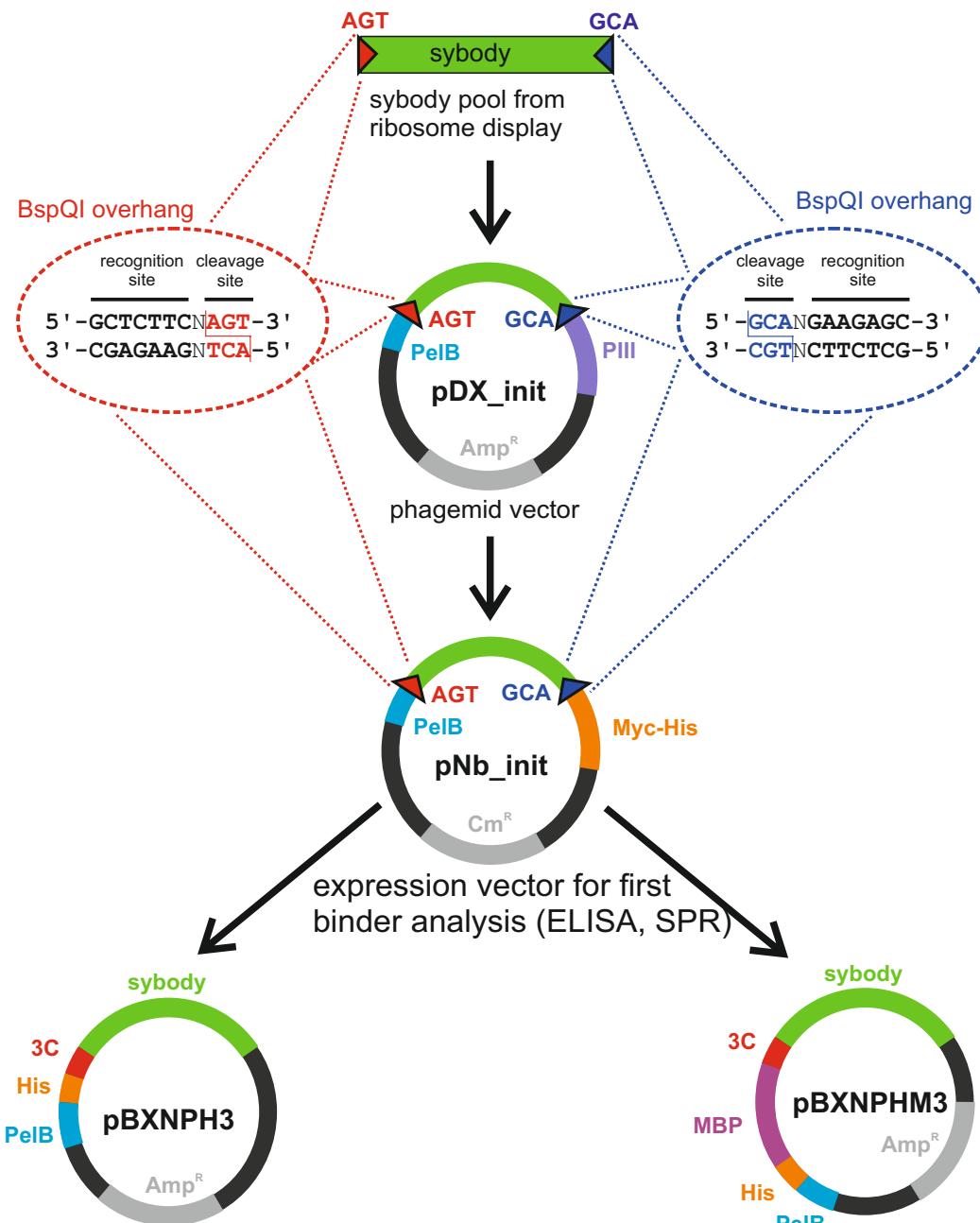
sybody expression vectors with cleavable tags for structural biology

MKYLLPTAAAGLLLLAAQPAMAADHHHHHHHHHHLEVLFQGPSDVQL
VESGGGSVQAGGSLRLSCAASGYIASINYLGWFRQAPGKEREGVAAV
SPAGGTPYYADSVKGRFTVSLDNAENTVYLQMNSLKPEDTALYYCAA
ARQGWYIPLNSYGYNYWGQGTQVTVSSA

MKYLLPTAAAGLLLLAAQPAMAMDIGINSDPNSSSHHHHHHHHHHPM
KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKF
PQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTW
DAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKA
KGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGA
KAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNI
DTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLEN
YLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTPGSLEVLFQG
PSDVQLVESGGGSVQAGGSLRLSCAASGYIASINYLGWFRQAPGKER
EGVAAVSPAGGTPYYADSVKGRFTVSLDNAENTVYLQMNSLKPEDTA
LYYCAAARQGWYIPLNSYGYNYWGQGTQVTVSSA

Supplementary Figure 6

Yellow/Orange/Red sticks: randomized residues CDR1/CDR2/CDR3

Green sticks: framework residues within CDRs contributing to the hydrophobic core
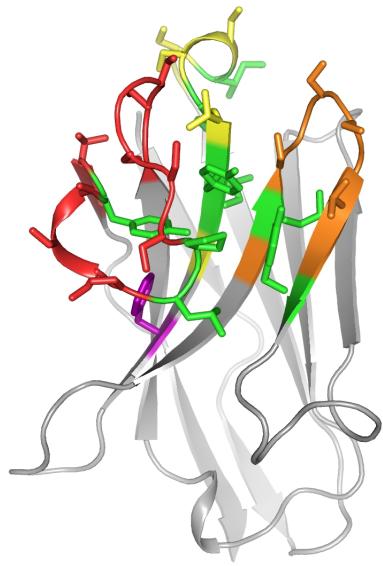
Homology model
Design Convex
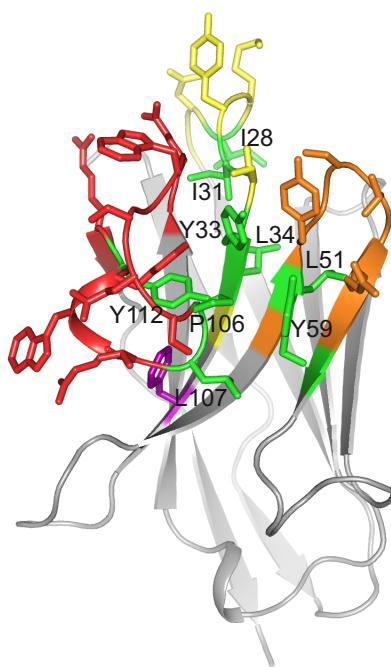
Crystal structure
Sb_MBP#1
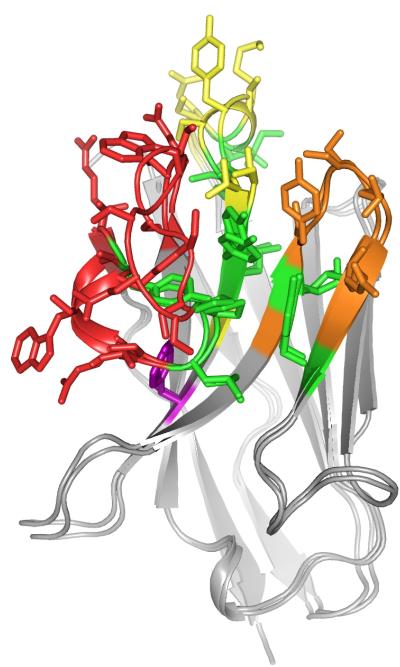
Superimposition
Model & structure

Supplementary Figure 7

Orange: randomized residue without interaction
Red: randomized residue with interaction, not conserved among three binders
Green: randomized residue with interaction, conserved among three binders
Blue: framework residue with interaction
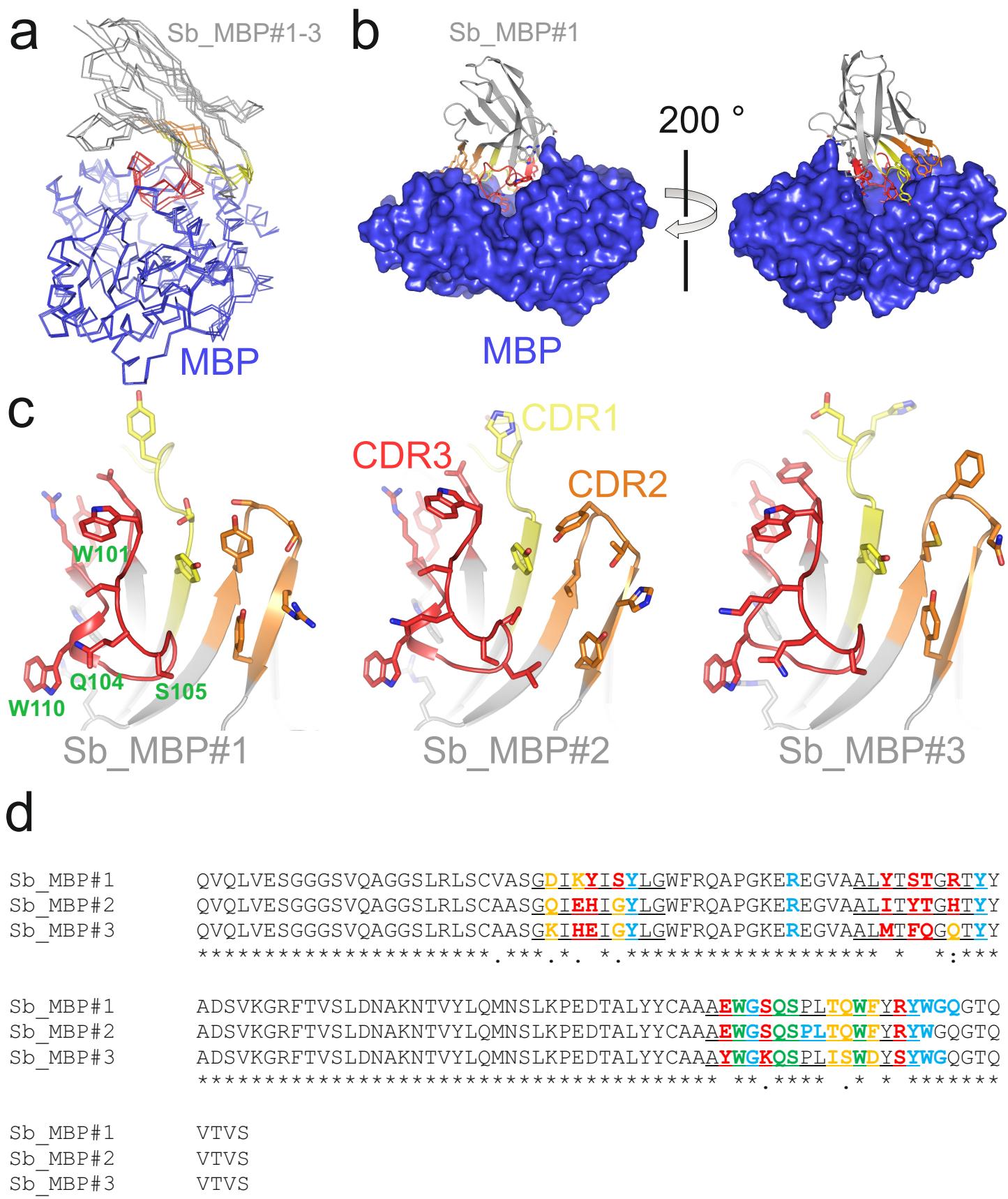
Supplementary Figure 8

a

Sb_MBP#1
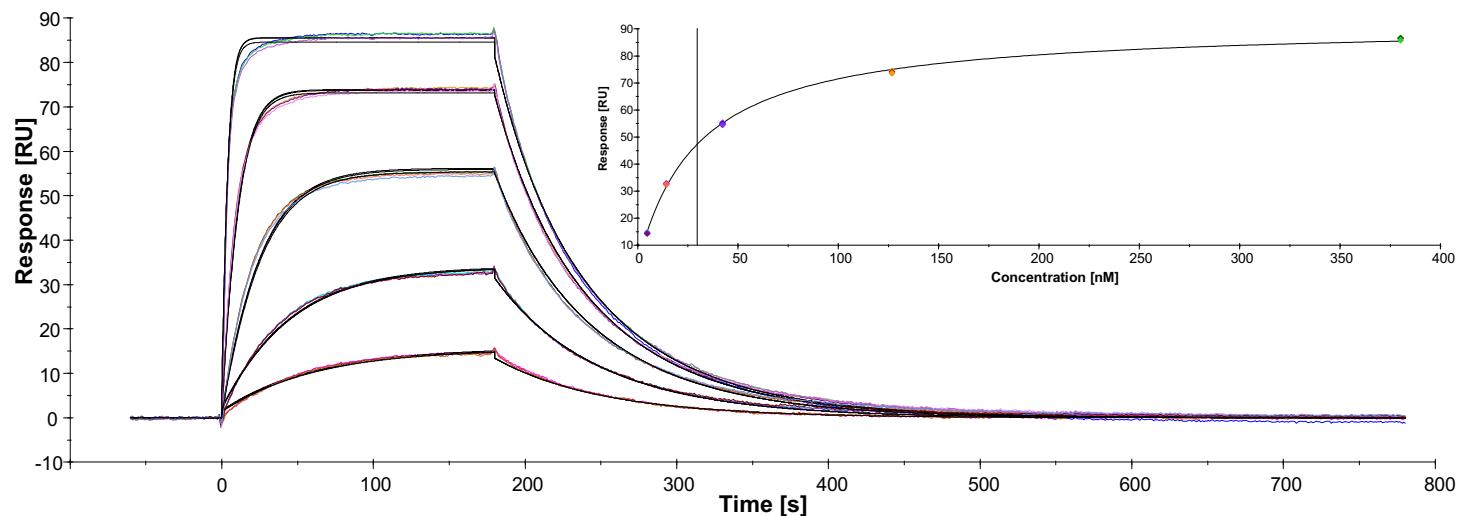
b

Sb_MBP#1		k_a (1/Ms)	k_d (1/s)	K_D (M)	R_{max} (RU)	Chi² (RU²)
Biacore X100	kinetics	7.82E+05	1.89E-02	2.41E-08	86.35	0.52
Biacore X100	equilibrium			2.97E-08	88.91	0.91
Sb_MBP#2						
Biacore X100	kinetics	3.88E+05	1.57E-02	4.04E-08	81.14	0.39
Biacore X100	equilibrium			4.83E-08	86.92	1.09
Sb_MBP#3						
Biacore X100	kinetics	5.24E+04	2.63E-02	5.02E-07	51.62	0.59
Biacore X100	equilibrium			7.03E-07	71.67	0.02

c
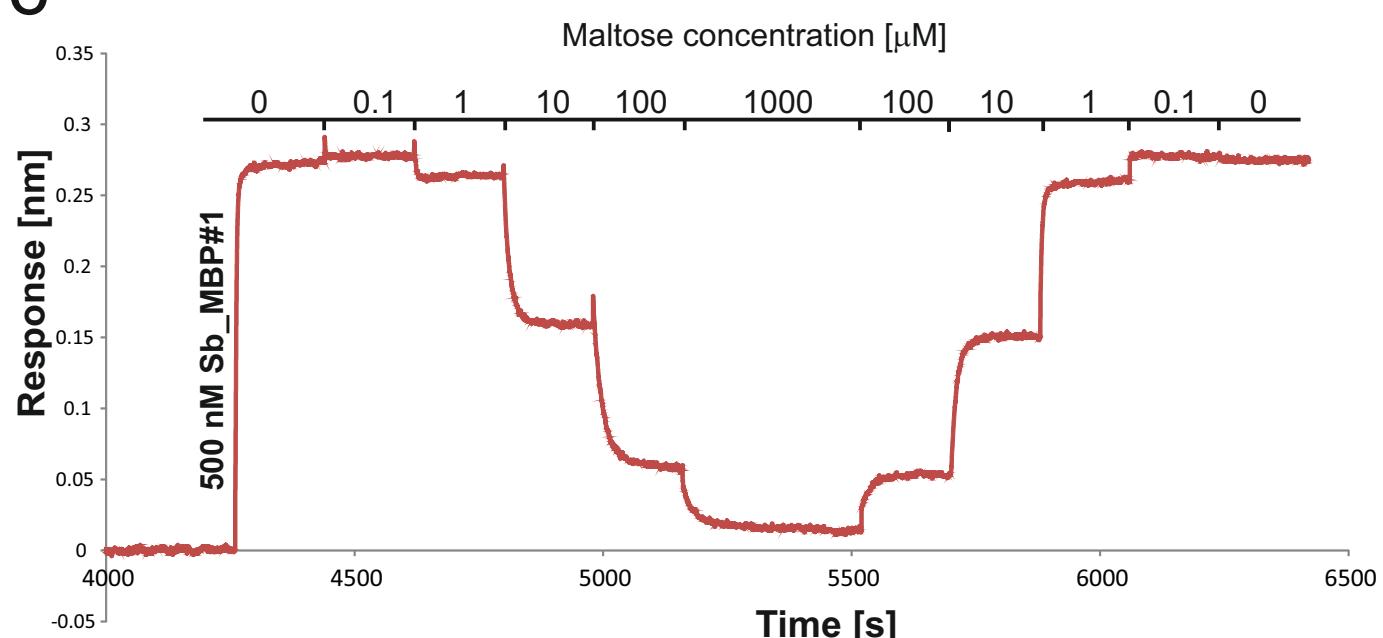
Supplementary Figure 9

a

	Binder	Binder affinities K_D [nM]			
		wildtype TM287/288		**TM287/288_E517A**	
	Binder	**Mg^{2+}**	**MgATP**	**Mg^{2+}**	**MgATP**
concave	Sb_TM#1	136	81	143	9.44
	Sb_TM#2	n.d.	5.26	n.d.	1.79
	Sb_TM#3	307	52.3	326	56.6
	Sb_TM#4	n.d.	5.41	n.d.	1.85
loop	Sb_TM#7	171	204	168	208
	Sb_TM#8	677	229	841	149
	Sb_TM#10	321	157	319	161
	Sb_TM#11	248	112	241	125
	Sb_TM#13	n.d.	1380	n.d.	1090
	Sb_TM#14	326	271	317	384
	Sb_TM#15	618	327	624	139
	Sb_TM#16	124	51.4	127	37.9
	Sb_TM#18	1520	1230	807	680
	Sb_TM#19	1780	476	3240	599
	Sb_TM#20	n.d.	807	1050	422
	Sb_TM#21	1100	347	1120	350
	Sb_TM#22	1810	868	2490	780
	Sb_TM#23	596	309	432	317
	Sb_TM#24	61.9	33.1	60.9	17.2
	Sb_TM#25	154	23.3	131	7.85
	Sb_TM#26	no signal	82.4	no signal	48.6
	Sb_TM#27	1970	269	1610	117
	Sb_TM#28	416	102	461	28.3
	Sb_TM#29	no signal	996	no signal	532
	Sb_TM#30	no signal	264	no signal	169
	Sb_TM#31	4620	1280	1770	558
convex	Sb_TM#33	4880	629	2430	228
	Sb_TM#35	no signal	110	no signal	65.7
	Sb_TM#37	no signal	173	no signal	66.1
	Sb_TM#38	242	71.3	231	20.5
	Sb_TM#39	738	140	981	422

🟦	state-specific
⬜ (tan)	state-unspecific
⬜	not determined

b
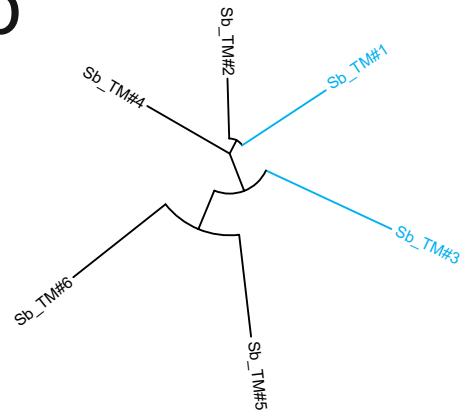
concave

loop

convex

Supplementary Figure 10

```
              10        20        30        40        50        60
Concave   QVQLVESGGGLVQAGGSLRLSCAASGFPVXXXXMXWYRQAPGKEREWVAAIXSXGXXTXY        60
α-TM#1    .....................L...........SETY.E.................N.W.WY.Y.     60
α-TM#2    ..............................WTAY.E.................E.W.WW.Y.        60
α-TM#3    ..............................ASYW.E.................W.Y.GN.W.        60
α-TM#4    ..............................NAQY.D.................E.F.WW.Y.        60
α-TM#5    ..............................ESRT.Y.................R.H.KM.Y.        60
α-TM#6    ..............................AHKT.Y.........D.....R.T.KY.Y.          60
                                      ‾‾‾‾‾‾‾‾‾                 ‾‾‾‾‾‾‾‾
                                        CDR1                      CDR2


              70        80        90        100       110
Concave   ADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCXVXVGX-XYXGQGTQVTVS        114
α-TM#1    ...................................A.E..E-W.Y.........      114
α-TM#2    ...................................A.E..D-H.Y.........      114
α-TM#3    ...................................A.H..A-G.T.........      114
α-TM#4    ...................................A.Q..L-E.Y.........      114
α-TM#5    ...................................W.W..SNH.K.........      115
α-TM#6    ...................................Y.T..Q-E.I.........      114
                                             ‾‾‾‾‾‾‾‾‾
                                               CDR3
```

Supplementary Figure 11

```
                                  10        20        30        40        50        60
Loop    QVQLVESGGGLVQAGGSLRLSCAASGFPVXXXXMXWYRQAPGKEREWVAAIXSXGXXTXYA  61
α-TM#7  .............................SSST.T...............W.Y.WS.K..  61
α-TM#8  .............................TNSY.A...............S.W.RH.Y..  61
α-TM#9  .............................REMN.Y...............A.N.QY.H..  61
α-TM#10 .............................DYWE.H...............Y.K.EE.K..  61
α-TM#11 .............................NQWY.A...............R.T.VK.W..  61
α-TM#12 .............................NAYN.Y...............A.S.SY.H..  61
α-TM#13 .............................TSTS.F...............D.K.AH.H..  61
α-TM#14 .............................WKEW.H...............D.T.HK.A..  61
α-TM#15 .............................QSRS.Y...............F.I.TT.E..  61
α-TM#16 .............................NFQE.H...............Y.N.QH.K..  61
α-TM#17 .............................YWNN.Y...............N.S.KY.H..  61
α-TM#18 .............................YEDS.F...............Y.Y.NS.A..  61
α-TM#19 .............................EAFH.R...............T.K.QH.F..  61
α-TM#20 .............................TQAN.Y...............D.D.MY.H..  61
α-TM#21 ..................S.....WESY.A...............D.H.EK.Y..  61
α-TM#22 ....A.............HHQY.A...............A.W.GG.Y..  61
α-TM#23 .............................EHTG.H...............Y.S.FK.Y..  61
α-TM#24 .............................THSY.A...............T.W.EK.Y..  61
α-TM#25 .............................HTEY.E...............E.Y.WW.Y..  61
α-TM#26 .............................YTRW.R...............E.S.-N.Y..  60
α-TM#27 .............................SNAY.E...............E.N.WW.Y..  61
α-TM#28 .............................YQKY.E...............E.N.WW.Y..  61
α-TM#29 .............................YATW.R...............E.M.QE.W..  61
α-TM#30 .............................YSTW.R...............E.T.YT.W..  61
α-TM#31 ....L.............YNTN.Y...............N.N.EY.H..  61
                                          ‾‾‾‾‾‾‾‾              ‾‾‾‾‾‾‾‾
                                            CDR1                   CDR2


                 70        80        90        100       110       120
Loop    DSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCNVKDX-GXXXXX-YDYWGQGTQVTVS  120
α-TM#7  .......................................Y-.TFWEI-.............  120
α-TM#8  .......................................S-.WEARW-.............  120
α-TM#9  .......................................Y-.SFFWG-.............  120
α-TM#10 ......................................WS.SELWW-.............  121
α-TM#11 .......................................D-.WQWDN-.............  120
α-TM#12 .......................................Y-.EWWKE-.............  120
α-TM#13 .......................................Y-.WWYGV-.............  120
α-TM#14 .......................................Y-.YVWAA-.............  120
α-TM#15 .......................................Y-.MWWWA-.............  120
α-TM#16 .......................................Y-.DWFEW-.............  120
α-TM#17 .......................................N-.AWWYY-.............  120
α-TM#18 .......................................H-.AFWWN-.............  120
α-TM#19 .......................................Y-.MWFMY-.............  120
α-TM#20 .......................................F-.WFFWD-.............  120
α-TM#21 .......................................N-.NVWWAW.............  121
α-TM#22 .......................................T-.SWWEW-.............  120
α-TM#23 .......................................Y-.DWLYT-.............  120
α-TM#24 .......................................S-.YQWQW-.............  120
α-TM#25 .......................................F-.AHAAE-.............  120
α-TM#26 .......................................S-.FNWFW-.............  119
α-TM#27 .......................................V-.NRQQT-.............  120
α-TM#28 .......................................Y-.NNWQY-.............  120
α-TM#29 .......................................Y-.DWFNY-.............  120
α-TM#30 .......................................Y-.WSWLA-.............  120
α-TM#31 .......................................D-.WTFWQ-.............  120
                                                 ‾‾‾‾‾‾‾‾‾
                                                   CDR3
```

Supplementary Figure 12

```
                  10          20          30          40          50          60
Convex    QVQLVESGGGSVQAGGSLRLSCAASGXIXXIXYLGWFRQAPGKEREGVAALXTXXGXTYYADS    63
α-TM#32   ..........................S.SS.T....................T.HA.T......    63
α-TM#33   ..........................S.SS.T....................T.NH.G......    63
α-TM#34   ..........................S.SS.T....................V.YG.Q......    63
α-TM#35   ..........................N.HH.S....................W.KD.N......    63
α-TM#36   ..........................S.SS.T....................S.SD.T......    63
α-TM#37   ..........................S.SS.T....................N.NS.Y......    63
α-TM#38   ..........................S.SS.T....................N.SS.Q......    63
α-TM#39   ..........................K.YW.T....................N.KQ.K......    63
α-TM#40   ..........................D.SD.S....................W.WI.W......    63
                                   ‾‾‾‾‾‾‾‾‾                  ‾‾‾‾‾‾‾‾
                                      CDR1                       CDR2

                  70          80          90         100         110         120
Convex    VKGRFTVSLDNAKNTVYLQMNSLKPEDTALYYCAAAXX-GXXX-PLXXXXYXYWGQGTQVTVS   124
α-TM#32   ...................................SW-.TWA-..IWYW.G..........   124
α-TM#33   ...................................SW-.VAH-..YWVW.G..........   124
α-TM#34   ...................................YW-.FWSA..WNYD.E..........   125
α-TM#35   ...............G...................DT-.SDT-..WDWV.W..........   124
α-TM#36   ...................................TY-.IWY-..TWWA.G..........   124
α-TM#37   ...................................FH-.DYM-..WWIL.G..........   124
α-TM#38   ...................................TY-.VSW-..VWWL.G..........   124
α-TM#39   ...................................YKV.NIS-..WALH.G..........   125
α-TM#40   ...................................YE-.YHR-..WDWV.T..........   124
                                             ‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾
                                                     CDR3
```

Supplementary Figure 13

Supplementary Figure 14