In silico design of stable single-domain antibodies with high affinity

Zhongyao Zhang[†,1,2,3,4], Rob van der Kant[†,1,2,3], Iva Marković[5,6], David Vizarraga[7], Teresa Garcia[1,2,3], Katerina Maragkou[1,2,3], Javier Delgado Blanco[7], Damiano Cianferoni[7], Gabriele Orlando[1,2,3], Gabriel Cia[1,2,3], Nick Geukens[12], Carlo Carolis[7], Alexander N. Volkov[10,11], Savvas N. Savvides[5,6], Maarten Dewilde[4,12], Joost Schymkowitz[*,1,2,3], Luis Serrano Pubul[*,7,8,9], Frederic Rousseau[*,1,2,3]

† These authors contributed equally to this work.

* Corresponding author Email: joost.schymkowitz@kuleuven.be, luis.serrano@crg.es, frederic.rousseau@kuleuven.be

Affiliations:

[1] Switch Laboratory, VIB Center for Brain and Disease Research, VIB; Leuven, 3000, Belgium.

[2] Switch Laboratory, Department of Cellular and Molecular Medicine, KU Leuven; Leuven, 3000, Belgium.

[3] Switch Laboratory, VIB Center for AI & Computational Biology, VIB; Leuven, 3000, Belgium.

[4] Laboratory for Therapeutic and Diagnostic Antibodies, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven; Leuven, 3000, Belgium.

[5] Unit for Structural Biology, Department of Biochemistry and Microbiology, Ghent University; Ghent, 9052, Belgium.

[6] Unit for Structural Biology, VIB-UGent Center for Inflammation Research; Ghent, 9052, Belgium.

23 [7] Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology;
24 Barcelona, 08003, Spain.
25 [8] Universitat Pompeu Fabra (UPF); Barcelona, 08002, Spain.
26 [9] ICREA; Barcelona, 08010, Spain.
27 [10] VIB- VUB Center for Structural Biology; Brussels, 1050, Belgium.
28 [11] Jean Jeener NMR Centre, Vrije Universiteit Brussel (VUB); Brussels, 1050, Belgium.
29 [12] PharmAbs, the KU Leuven Antibody Center, KU Leuven; Leuven, 3000, Belgium.

Abstract

Antibody-based therapeutics have become indispensable in modern medicine, but traditional methods of antibody discovery often present with limitations in developability, cross-reactivity, and ethical concerns. While deep learning and generative approaches have shown promise in the design of high affinity protein binders, *de novo* antibody design remains challenging. Here, we present EvolveX, a structure-based computational pipeline for designing antibody fragments. EvolveX utilizes ModelX and empirical force field FoldX to optimize complementarity-determining regions (CDRs) and TANGO for aggregation analysis. We demonstrate the ability of EvolveX to redesign a single-domain VHH antibody fragment targeting mouse Vsig4 to address two challenges: enhancing stability and affinity for the original target and redesigning it for high affinity to the human ortholog. The redesigned variants of VHH fragments specific to mouse Vsig4 showed improved physicochemical properties, while retaining binding affinities comparable to the original version. Notably, EvolveX improved the binding affinity of VHHs to human Vsig4 by over 1000-fold, transforming low-affinity binders into nanomolar-affinity molecules. Structural analyses by X-ray crystallography and NMR confirmed the accuracy of the designs, which display optimized interactions with the antigen. NGS and re-modelling analysis further demonstrated the efficiency of FoldX-based design pipeline. Collectively, our study highlights EvolveX's potential to overcome current limitations in antibody design, offering a powerful tool for the development of next-generation therapeutics with enhanced specificity, stability, and efficacy.

Main Text:

Antibodies have become highly promising and dominating tools in both therapeutics and diagnostics. Over the past decades, most of biologics approved by the FDA are therapeutic antibodies [1]. Traditional antibody discovery methods primarily rely on animal immunization. While effective, these methods present challenges such as suboptimal developability, limited and unpredictable cross-reactivity, and difficulties in targeting conserved binding epitopes [2,3]. Additionally, ethical concerns regarding the use of animal models further complicate the immunization-based approach [4]. These limitations have led to a growing shift toward computational-based antibody design and discovery approaches, where various parameters of antibody development can be optimized before experimental testing.

Recent advancements in deep learning networks and diffusion-based generative approaches for designing antibody sequences have attracted increasing attention. However, *de novo* design of high affinity antibodies remains largely unexplored and challenging [5-8]. In this context, structure-based antibody design has emerged as a promising approach[3,9-13]. Utilizing the computational methods, three-dimensional structure of antibodies and their target antigens can be predicted and modeled, allowing for the further optimization of the full complex structures and their binding interfaces. [11,12]. Through a combination of computational molecular modeling, rational design of sequences, and experimental validation, structure-based *in silico* antibody design holds the potential to streamline the antibody development process and to facilitate the discovery of antibodies with multiple properties such as enhanced binding affinity and stability [12,13].

Here we introduce our innovative *in silico* antibody design pipeline, EvolveX, where the structure optimization is achieved using ModelX [14-16], interaction energy and antibody stability are evaluated by the empirical forcefield FoldX [16,17] and aggregation propensity is analyzed by

72 TANGO [18-20]. EvolveX enables the design of antibody sequences that effectively bind to the
73 predefined epitope while maintaining good thermodynamic stability and avoiding the introduction
74 of aggregation propensity. To validate our approach, we employ a process termed *re-paratoping*,
75 where sequence information from a known antibody-protein complex was removed, and new
76 sequences were designed to target the same epitope. The entire pipeline is evaluated through a
77 series of experimental assays, including phage display, biophysical characterization, NMR, X-ray
78 crystallography, and next-generation sequencing (NGS), demonstrating the successful design of
79 stable VHHs with high affinity starting from both high and low-affinity protein complexes [21].
80 Overall, our finding underscores the potential of structure-based *in silico* antibody design as a
81 powerful tool in biotechnology and drug discovery.

82 ***In silico* rebuilding of an antibody paratope starting from a high affinity complex**

83 Wen et al. [21] generated a VHH, Nb119 (in this work further referred to as VHH_WT), through
84 classical immunization against the extracellular IgV domain of the mouse Vsig4 protein. This
85 VHH showed high affinity towards mouse Vsig4 (mVsig4, K_D ~5 nM), but a thousand-fold lower
86 affinity towards human Vsig4 (hVsig4, K_D ~5 μM). The structure of the antibody-antigen complex
87 was determined by X-ray crystallography for both orthologs. We used the crystal structure of the
88 complex of the VHH to mVsig4 (VHH_WT-mVsig4, PDBID: 5IMM and 5IMO, [21]) to validate
89 the capacity of EvolveX to rebuild an antibody paratope starting from a high affinity complex (Fig.
90 1a). To this end, we removed the sequence information from CDR2 and CDR3 of the VHH_WT-
91 mVsig4 complexes since these regions are most involved in binding to mVsig4: Each amino acid
92 residue in CDR2 and CDR3 (except Proline and Glycine) that resided within 8 Å from the ligand
93 was mutated to alanine. From this all-alanine starting point, the compatibility of each of the
94 selected positions with all 20 amino acids was explored using the *PSSM* function in FoldX, which

95 yields the impact of each point mutation on both the thermodynamic stability of the VHH, as well
96 as interaction energy with the ligand (see methods for details). Based on these numbers, random
97 starting sequences were generated to be compatible with the complex structure. These starting
98 sequences were used as input for a genetic algorithm, where selection was based on the interaction
99 energy computed by FoldX, with preset ceilings for the thermodynamic stability of the VHH (see
100 Fig. 1a and methods for details).
101 The resulting sequences were filtered based on thresholds for the FoldX prediction of interaction
102 energy and VHH stability [17], as well as TANGO predicted aggregation propensity [18]. Additionally,
103 Z-scores were calculated based on densities derived from publicly available crystal structures of
104 VHHs bound to protein ligands, for cavities within the VHH, cavities at the interface of VHHs
105 with their ligand and VHH net charge (see methods for details). The top 6000 sequences with the
106 lowest summed Z-scores were obtained as oligonucleotides and cloned into the VHH_WT
107 backbone in a phagemid vector. The resulting library was enriched for mVsig4 binders through
108 several rounds of phage display using progressively lower concentrations of mVsig4 down to 5
109 nM (see Supplementary Fig.1 and methods for details). Periplasmic extracts were prepared for 188
110 randomly picked VHH clones from the enriched libraries, yielding 36 unique sequences
111 (Supplementary Table.1), and subsequently analyzed for mVsig4 binding using the Amplified
112 Luminescent Proximity Homogenous Assay (AlphaScreen) (Fig. 1b). This revealed clones with
113 AlphaScreen signal to noise ratios (S/N ratio) similar- or higher than VHH_WT, of which 12 were
114 selected. Of these VHH variants, we determined the dissociation velocity rate constant governing
115 dissociation from mVsig4 (k_d) using BioLayer Interferometry (BLI) (Fig. 1c, Supplementary
116 Table.2), which was quite similar in all cases.

117 The top four candidates (VHH_m1, VHH_m2, VHH_m3, and VHH_m4) with slightly more
118 favorable k_d values than the rest were subsequently recombinantly produced at larger scale (10
119 mg), chromatographically purified, and biophysically characterized. Binding affinities (K_D) for
120 mVsig4 were determined using Surface Plasmon Resonance (SPR) (Fig. 1d and e, Supplementary
121 Fig.3) and BLI (Fig. 1e, Supplementary Fig.2), revealing that the four VHH candidates have
122 affinities close to VHH_WT (Fig. 1e). The four VHHs were subjected to chemical denaturation
123 analysis using urea and guanidine hydrochloride (Supplementary Fig.4). Curve fitting indicated a
124 significant increase in stability by about -2 kcal/mol for all four designed sequences (Fig. 1f)
125 compared to VHH_WT. Thermal protein unfolding (T_m) and aggregation onset temperatures (T_{agg})
126 were determined simultaneously using Intrinsic Tryptophan Fluorescence (ITF) (Fig. 1g,
127 Supplementary Fig.5) and Right-Angle Light Scattering (RALS), respectively (Fig. 1h,
128 Supplementary Fig.6). The designed sequences showed increased thermal stability, about 5 °C in
129 both T_m and T_{agg}, and the relative order in stability coincides quite well between the chemical and
130 thermodynamic denaturation with VHH_m4 being the more stable and VHH_m3 the least.
131 Although EvolveX computationally explored 5 positions in CDR2 and 12 in CDR3 (~$1x10^{22}$
132 theoretical combinations), the four selected candidates were very similar to each other in sequence,
133 differing only at two positions where we found chemically similar amino acids: a Tyr or Phe at
134 position 107 and an Asp or a Glu at position 112, respectively (Fig. 1i, Supplementary Table.3).
135 Interestingly the two more stable VHH sequences VHH_m2 and VHH_m4 contain a Glu at
136 position 112, while having a Phe or a Tyr at position 107 does not seem to have an important role
137 in stability. Compared to wild type, the designed variants carried two mutations in the CDR2 loop
138 and 5 or 6 mutations in the CDR3 loop (Fig. 1i).

139 The successful identification of the multi-parametric optimized sequences with EvolveX
140 demonstrates the capability of our computational method in designing VHH CDR sequence
141 combinations when starting from a structure of a high affinity complex.

142 ***In silico* design of optimal CDR2/3 sequences targeting human Vsig4**

143 Subsequently, we set out to identify high-affinity binders starting from a low-affinity complex,
144 starting from structural templates obtained from the crystal structures of the low micromolar
145 affinity complex between VHH_WT and human Vsig4 (PDB ID 5IMK and 5IML, [21]) (Fig. 2a).
146 The backbone conformation of CDR3 was adapted using the *Bridging* command in ModelX [16]
147 using 5IMK as a starting point, adding 14 starting complex structures (see methods for details).
148 Sequence design as described above yielded 6000 candidate sequences (Fig. 2a). The mouse
149 VHH_m and human VHH_h sequences were combined into one phage library, which was selected
150 against hVsig4 using similar procedures as before (Supplementary Fig.1). Based on the
151 enrichment, the phages from the third round of selection with 50nM hVsig4 (R3) and round 4
152 selection with 5nM hVsig4 (R4) were chosen for screening. We randomly picked 188 colonies
153 from both libraries for expression and sequencing, yielding 112 unique sequences, which we
154 ranked for binding to hVsig4 using classical ELISA. The average ELISA signal from the
155 R4_hVsig4_5nM group is higher than that of the R3_hVsig4_50nM group and most of the clones
156 have a higher ELISA signal than VHH_WT (Fig. 2b, Supplementary Table.4). 56 sequences were
157 tested for off-rate using BLI (Fig. 2c, Supplementary Table.5). The 3 clones with the best off-rates
158 (VHH_h1, VHH_h2 and VHH_h3), the clone with the highest ELISA signal (VHH_h4), and the
159 most occurring clone (VHH_h5) were selected for purification and subsequent biophysical
160 characterization.

161 Remarkably, the measured affinity of the designed sequences situated in the low nanomolar
162 affinities (Fig. 2d and e), meaning an improvement in affinity to human Vsig4 of up to 1000-fold
163 over VHH_WT (Fig. 2d and e, Supplementary Fig.7, Supplementary Fig.8). Affinities were cross
164 validated with the VHH_m hits for binding to both mVsig4 and hVsig4. This revealed all designed
165 VHHs show specificity for their intended target (Supplementary Fig.9). All VHHs were subjected
166 to chemical and thermal denaturation. Chemical denaturation by urea (Supplementary Fig.10a)
167 and guanidine hydrochloride (Supplementary Fig.10b) indicated the five designed sequences had
168 an increase in stability from -1.5 up to -3.5 kcal/mol compared with the original human VHH_WT
169 (Fig. 2f and Supplementary Table.6). Both T_m (Fig. 2g, Supplementary Fig.11) and T_{agg} (Fig. 2h,
170 Supplementary Fig.12) were significantly improved for all designed sequences compared to
171 human VHH_WT (Supplementary Table.6).
172 VHH_h2, VHH_h4 and VHH_h5 exhibited the strongest binding to hVsig4 (K_D_BLI and K_D_SPR
173 ~ 5 nM), which is 1000-fold better than VHH_WT (K_D_BLI ~ 5 μM) (Fig. 2e, Supplementary
174 Table.6). VHH_h1 and VHH_h3 exhibited worse K_D values (10~40nM) to hVsig4 than the first
175 three candidates, while still being substantially better than VHH_WT. In all cases we found
176 common mutations among the five VHHs (WN to PF in CDR2, DK to NR, F to D and E to N in
177 CDR3) (Fig. 2i). Interestingly, when comparing the sequences analyzed for hVsig4 and mVsig4,
178 we found some positions that were mutated (although to different residues) in both cases at CDR2
179 (GS) and CDR3 (D and E). In the case of the hVsig4 we found more positions that were mutated
180 indicating the need of larger sequence changes to ensure high affinity binding.
181 To confirm the epitope and evaluate the predictions done by EvolveX, the designed VHHs were
182 crystallized in complex with hVsig4 (Fig. 3, Supplementary Fig.13, Supplementary Table.7). All
183 five high-resolution crystal structures revealed that the binding epitope is almost identical among

184 all hit VHHs compared to the crystal structure of VHH_WT (Fig. 3b and c). When superimposing
185 on hVsig4 we see that the orientation for each VHH as well as for VHH_WT is slightly different
186 (Fig. 3b and Supplementary Fig.14). This might be due to the interaction between the designed
187 negative charges in CDR3 (Asp107 and Asp109) and the N-terminus of hVsig4 (Supplementary
188 Fig.14d). Overall, more residues are interacting with hVsig4 in the designed VHHs compared to
189 VHH_WT (Supplementary Table.8). Additionally, the conformation of CDR3 is very similar to
190 the VHH_WT crystal structures (mean pairwise backbone RMSD = 0.29 Å) (Fig. 3c).
191 To elucidate the biophysical parameters responsible for the increased stability of the designed
192 sequences over the original VHH, we studied VHH_WT and VHH_h4 (the most stable binder to
193 hVsig4) by nuclear magnetic resonance (NMR) spectroscopy. VHH_WT features a well-dispersed
194 heteronuclear single-quantum correlation (HSQC) spectrum, where nearly all peaks could be
195 unambiguously assigned (Supplementary Fig.15). Interestingly, most residues in the three CDRs
196 showed no backbone amide resonances (Supplementary Fig.16). The paucity of the CDR signals
197 was reported before for several other VHHs and attributed to a pronounced flexibility and
198 conformational heterogeneity of the CDR loops [22-24]. The HSQC spectrum of VHH_h4 is similar
199 to that of VHH_WT, but contains more peaks (Supplementary Fig.17), some of which could be
200 assigned to the residues in the CDR3 loop (Supplementary Fig.18). In addition, several VHH_h4
201 resonances not observed in the NMR spectra of the VHH_WT (and showing very weak signals in
202 triple-resonance NMR experiments, precluding their unambiguous assignment) could belong to
203 other CDRs. Overall, it appears that VHH_h4 undergoes a much less extensive conformational
204 exchange upon target binding than the WT. Predicted from the backbone chemical shifts, the
205 secondary structures of VHH_WT and VHH_h4 are very similar and agree well with that seen in
206 the X-ray structure (Supplementary Fig.16 and 18), which testifies to the high structural

207 conservation of the VHH fold outside the CDRs. To delve further into differences between
208 VHH_WT and VHH_h4, we performed NMR relaxation analysis (Supplementary Fig.19-21),
209 which revealed that free VHH_h4 in solution maintains the bound-like protein conformation, with
210 closely packed CDR loops undergoing limited motion. At the same time, the CDRs of VHH_WT
211 are largely disordered and involved in extensive conformational exchange. These differences could
212 explain the observed disparity in the thermal stability of the two proteins, where motional disorder
213 of CDRs renders VHH_WT less stable.

214 **FoldX modeling of VHH hit sequences reveals energy differences between VHH_WT and**
215 **VHH candidates**

216 After obtaining the sequences and the crystal structures of those Vsig4 candidates, we re-built the
217 structure models of all VHH candidates based on nine available structures (PDBID: 5IMM and
218 5IMO for VHH_mVsig4, and 5IMK, 5IML, 9EZU, 9EZV, 9EVW, 9EZH, and 9EZI for
219 VHH_hVsig4) using the FoldX *BuildModel* function. The interaction energy ($\Delta G_{interaction}$) and
220 VHH individual energy (ΔG_{VHH}) were calculated using FoldX *AnalyseComplex* function and
221 visualized in corresponding plots (Supplementary Fig.22). Triplicate running was performed for
222 each model. For mVsig4 hits, VHH_m1 and VHH_m3 exhibited slightly lower $\Delta G_{interaction}$ than
223 VHH_m2 and VHH_m4, consistent with the affinity data from SPR (Supplementary Fig.22a, Fig.
224 1e). All four hits demonstrated lower $\Delta G_{interaction}$ than VHH_WT, although the difference was not
225 statistically significant (Supplementary Fig.22a). Additionally, all four hits showed significantly
226 lower ΔG_{VHH} compared to VHH_WT, aligning with the ΔG values measured in the chemical
227 denaturation experiment and T_m/T_{agg} values measured in the thermal stability experiment
228 (Supplementary Fig.22b, Fig. 1f-h).

229 For the hVsig4 hits, all five hits showed significantly lower $\Delta G_{interaction}$ and ΔG_{VHH} relative to
230 VHH_WT across all five newly obtained crystal structures (Supplementary Fig.22 c-d). The hit
231 models generated on two original structures (5IMK and 5IML) did not show clear trend relative to
232 VHH_WT (Supplementary Fig.22 c-d). This shows the importance of performing backbone
233 moving to generate the optimized starting structure model when designing from a low-affinity
234 complex.

235 **FoldX modeling of NGS-identified CDR designs reveals energy change differences between**
236 **enriched and negative CDR designs**

237 To further investigate the relationship between enriched EvolveX-designed sequences and FoldX-
238 based energy, next-generation sequencing (NGS) was performed on libraries from all rounds of
239 selection (Supplementary Fig.1), including the empty control conducted in parallel with each
240 round (Fig. 4a). As CDR recombination can occur during the PCR steps of the sample preparation,
241 the sequences of CDR2 and CDR3 were extracted and analyzed separately. Enriched CDR designs
242 were filtered from R1 selection for both mVsig4 and hVsig4 based on specific criteria (count >
243 10, enrichment factor > 10). Notably, the counts for all enriched CDR designs exceeded 1000 (red
244 dots in Fig. 4b, Supplementary Fig.23a, c, i), significantly surpassing the counts for enriched non-
245 designs (black dots in Fig. 4b, Supplementary Fig.23a, c, i), which arose from mutations
246 introduced during cloning or phage selection. An equivalent number of negative CDR designs
247 were identified, defined as sequences present in the input but absent or minimally represented in
248 the R1 selection (Fig. 4c, Supplementary Fig. 23b, d, j). Models incorporating enriched or negative
249 CDR sequences were generated using the FoldX *BuildModel* function based on nine available
250 structures: 5IMM and 5IMO for VHH_mVsig4, and 5IMK, 5IML, 9EZU, 9EZV, 9EZW, 9EZH,
251 and 9EZI for VHH_hVsig4. The interaction energy change ($\Delta\Delta G_{interaction}$) and VHH individual

252 energy change ($\Delta\Delta G_{VHH}$) were calculated using FoldX *AnalyseComplex* function and visualized
253 in corresponding plots (Fig. 4d-e, Supplementary Fig.23e-h, k-l).
254 For mVsig4, 31 enriched CDR2 designs and 34 enriched CDR3 designs were identified
255 (Supplementary Fig.23a-d). Compared to negative CDR sequences, all models with enriched
256 CDR2 designs showed significantly lower $\Delta\Delta G_{interaction}$ and all models with enriched CDR3 designs
257 showed significantly lower $\Delta\Delta G_{VHH}$ on both VHH-mVsig4 structures (Supplementary Fig.23e-h).
258 For hVsig4, 38 enriched CDR2 and 115 enriched CDR3 designed sequences were identified and
259 analyzed across all seven VHH_hVsig4 structures (Fig. 4b-e, Supplementary Fig.23i-l). While no
260 significant differences were observed between enriched and negative CDR2 models in $\Delta\Delta G_{interaction}$
261 and $\Delta\Delta G_{VHH}$, the average $\Delta\Delta G_{interaction}$ across all enriched CDR2 designs was consistently lower
262 than that of the negative models (Supplementary Fig.23k-l). In five newly obtained structures
263 (9EZU, 9EZV, 9EZW, 9EZH, and 9EZI), enriched CDR3 models exhibited significantly lower
264 $\Delta\Delta G_{interaction}$ (Fig. 4d). In addition, all enriched CDR3 models demonstrated significantly lower
265 $\Delta\Delta G_{VHH}$ compared to the negative models across all analyzed structures (Fig. 4e). Overall, the
266 results demonstrated that FoldX parameters were highly effective in distinguishing optimal designs
267 from negative sequences on the proper structures. Beyond the top hits we identified, 30–100
268 optimal designs were successfully obtained for each CDR based on both NGS and FoldX analysis.

Discussion

The advent of computational methods for antibody design is revolutionizing the field, allowing for the targeted creation of antibodies with specific binding capabilities and stability. In this study, we demonstrate the efficacy of an *in silico* antibody design pipeline, EvolveX, in designing VHHs targeting predefined epitopes on protein targets. Our results highlight the versatility and robustness of the approach, as evidenced by successful sequence design efforts starting from both high and low-affinity complexes.

The success of the *in silico* antibody design pipeline is underscored by the validation experiments conducted for both high and low-affinity complexes of VHH_WT with mouse and human Vsig4, respectively. Through a combination of sequence and structure optimization, and biophysical and structural experimental validation, EvolveX was able to design small, targeted libraries yielding high-affinity VHHs. Notably, the designed VHHs consistently exhibited significantly enhanced thermal, colloidal, and chemical stability, indicating the robustness of the computational design process. Interestingly, comparison of the designed VHHs targeting mouse and human Vsig4 revealed differences in the mutation patterns within the complementarity-determining regions (CDRs), indicating the need for tailored design strategies depending on the target species.

The designed VHHs exhibited not only high affinity but also demonstrated thermodynamic stability comparable to or better than the wild-type antibody. Structural analysis of the VHH-hVsig4 protein complexes confirmed epitope retention and revealed insights into the binding mechanism, providing valuable information for further optimization and development.

Furthermore, FoldX modeling of VHH hit sequences and NGS-identified sequences demonstrated consistency between wet-lab data and *in silico* predicted parameters. Models based on low-affinity hVsig4-VHH complex (5IMK) showed no clear trends or, in some cases, opposite results to

292 expectations. However, the backbone of VHH_WT in the 5IMK complex was initially adapted
293 using ModelX during the design process to include additional starting structures, which could
294 prove highly valuable in identifying high affinity hVsig4 hits. These findings emphasize the
295 critical importance of backbone conformation exploration into the modeling workflow and
296 highlight the need for further refinement of the force field to enhance predictive confidence.

297 Collectively, our study demonstrates the power of computational antibody design in generating
298 VHHs with desired binding properties. The combination of molecular modeling, sequence
299 optimization, and experimental validation offers a systematic and efficient approach for antibody
300 engineering, with broad implications for drug development, diagnostics, and biotechnology.
301 Future work will focus on further refining the design and modelling pipeline, reducing library sizes
302 and exploring its application to additional protein targets and therapeutic modalities.

Main References:

1 Mullard, A. 2023 FDA approvals. *Nat Rev Drug Discov* **23**, 88-95 (2024).

2 Kim, J., McFee, M., Fang, Q., Abdin, O. & Kim, P. M. Computational and artificial intelligence-based methods for antibody development. *Trends Pharmacol Sci* **44**, 175-189 (2023).

3 Sormanni, P., Aprile, F. A. & Vendruscolo, M. Third generation antibody discovery methods: in silico rational design. *Chem Soc Rev* **47**, 9137-9157 (2018).

4 Kiani, A. K. et al. Ethical considerations regarding animal experimentation. *J Prev Med Hyg* **63**, E255-e266 (2022).

5 Luo, S. et al., Antigen-Specific Antibody Design and Optimization with Diffusion-Based Generative Models for Protein Structures. Preprint at https://doi.org/10.1101/2022.07.10.499510 (2022).

6 Bennett, N. R. et al., Atomically accurate de novo design of single-domain antibodies. Preprint at https://doi.org/10.1101/2024.03.14.585103 (2024).

7 Eguchi, R. R. et al., Deep Generative Design of Epitope-Specific Binding Proteins by Latent Conformation Optimization. Preprint at https://doi.org/10.1101/2022.12.22.521698 (2022).

8 Shanehsazzadeh, A. et al., Unlocking *de novo* antibody design with generative artificial intelligence. Preprint at https://doi.org/10.1101/2023.01.08.523187 (2024).

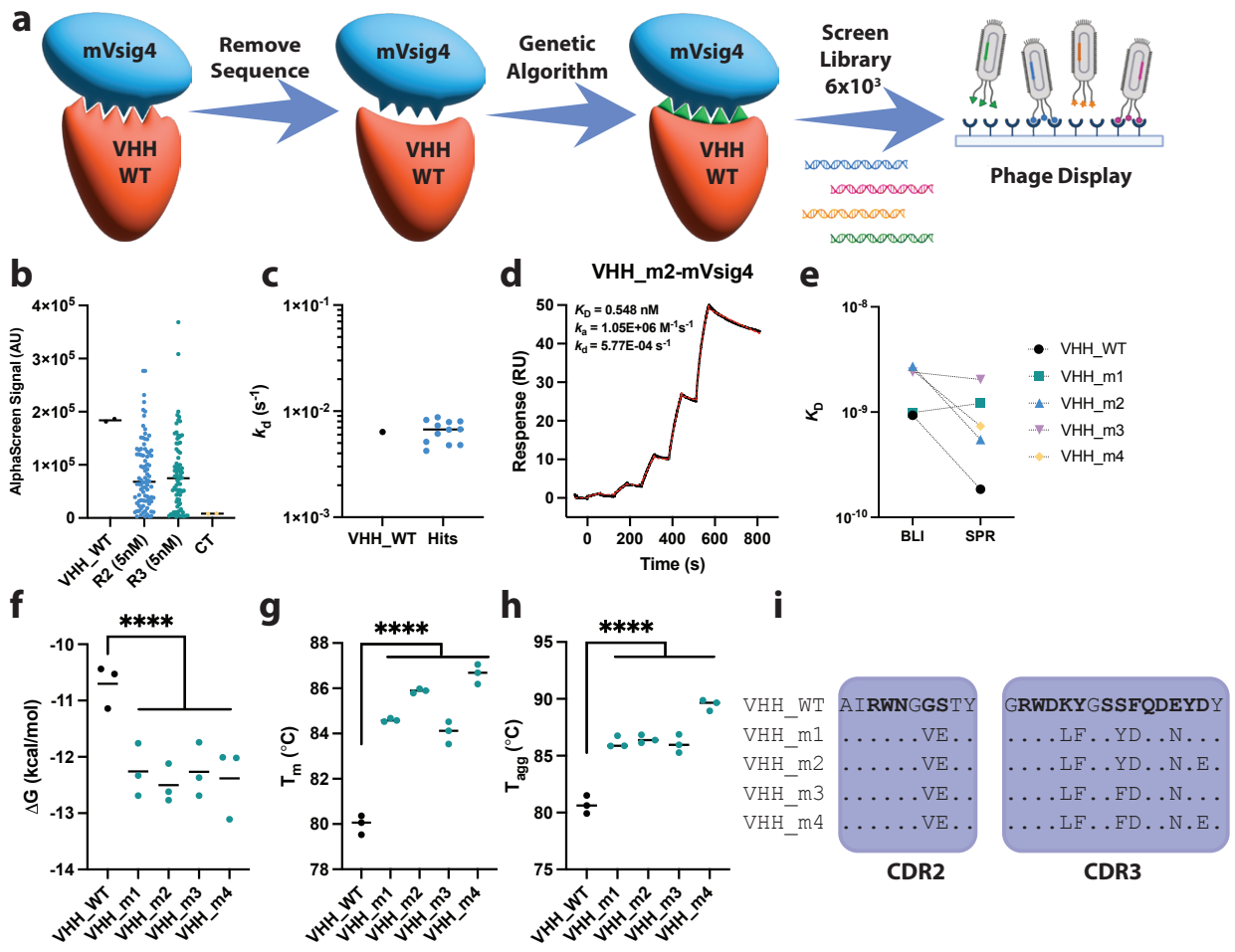
9 Lapidoth, G. D. et al. AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. *Proteins* **83**, 1385-1406 (2015).

10 Hummer, A. M., Abanades, B. & Deane, C. M. Advances in computational structure-based antibody design. *Curr Opin Struct Biol* **74**, 102379 (2022).

326 11 Adolf-Bryfogle, J. et al. RosettaAntibodyDesign (RAbD): A general framework for
327 computational antibody design. *PLoS Comput Biol* **14**, e1006112 (2018).
328 12 Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track
329 neural network. *Science* **373**, 871-876 (2021).
330 13 Nimrod, G. et al. Computational Design of Epitope-Specific Functional Antibodies. *Cell*
331 *Rep* **25**, 2121-2131 e2125 (2018).
332 14 Cianferoni, D., Radusky, L. G., Head, S. A., Serrano, L. & Delgado, J. ProteinFishing: a
333 protein complex generator within the ModelX toolsuite. *Bioinformatics* **36**, 4208-4210
334 (2020).
335 15 Montero-Blay, A. et al. Bacterial expression of a designed single-chain IL-10 prevents
336 severe lung inflammation. *Mol Syst Biol* **19**, e11037 (2023).
337 16 Blanco, J. D., Radusky, L., Climente-Gonzalez, H. & Serrano, L. FoldX accurate structural
338 protein-DNA binding prediction using PADA1 (Protein Assisted DNA Assembly 1).
339 *Nucleic Acids Res* **46**, 3852-3863 (2018).
340 17 Schymkowitz, J. et al. The FoldX web server: an online force field. *Nucleic Acids Res* **33**,
341 W382-388 (2005).
342 18 Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of
343 sequence-dependent and mutational effects on the aggregation of peptides and proteins.
344 *Nat Biotechnol* **22**, 1302-1306 (2004).
345 19 Linding, R., Schymkowitz, J., Rousseau, F., Diella, F. & Serrano, L. A comparative study
346 of the relationship between protein structure and beta-aggregation in globular and
347 intrinsically disordered proteins. *J Mol Biol* **342**, 345-353 (2004).

348 20 Rousseau, F., Schymkowitz, J. & Serrano, L. Protein aggregation and amyloidosis:
349 confusion of the kinds? *Curr Opin Struct Biol* **16**, 118-126 (2006).
350 21 Wen, Y. et al. Structural evaluation of a nanobody targeting complement receptor Vsig4
351 and its cross reactivity. *Immunobiology* **222**, 807-813 (2017).
352 22 Vuchelen, A. et al. (1)H, (13)C and (15)N assignments of a camelid nanobody directed
353 against human alpha-synuclein. *Biomol NMR Assign* **3**, 231-233 (2009).
354 23 Nie, Y. et al. Chemical shift assignments of a camelid nanobody against aflatoxin B(1).
355 *Biomol NMR Assign* **13**, 75-78 (2019).
356 24 Percipalle, M., Hunashal, Y., Steyaert, J., Fogolari, F. & Esposito, G. Structure of
357 Nanobody Nb23. *Molecules* **26** (2021).

358 **Figures:**



359

360 **Fig. 1. Library design, screening, hit selection and validation of VHH's designed with**
361 **EvolveX to bind mouse Vsig4.** (**a**) Graphical representation of the EvolveX *in silico* pipeline,
362 where crystal structures of the high affinity complex between mouse Vsig4 (mVsig4) and
363 VHH_WT were used as templates to *in silico* design sequences of CDR2 and CDR3 that should
364 bind mVsig4 with high affinity, while displaying good stability. A small library was subsequently
365 screened using phage display. Icons of DNA strands and phage display were created with
366 BioRender.com. (**b**) Colonies picked after phage display selection were tested for binding to
367 mVsig4 using AlphaScreen. VHH_WT was used as a positive control, colonies picked after two

368 different rounds of selection were screened. (**c**) Off-rates of the top 12 unique sequences based on
369 AlphaScreen signal to mVsig4, determined using BLI. (**d**) Example SPR binding curves of 1 out
370 of 4 hits binding to mVsig4 with low nanomolar affinity. (**e**) Affinity constants of all 4 hits and
371 VHH_WT binding to mVsig4, determined using BLI and SPR. (**f**) ΔG values determined through
372 curve fitting based on data shown in G. (**g**) Melting temperatures (T_m) for 4 hits and VHH_WT
373 based on Intrinsic Tryptophan Fluorescence (ITF) through a temperature ramp. (**h**) Aggregation
374 onset temperatures (T_{agg}) for 4 hits and VHH_WT based on Right Angle Light Scattering (RALS)
375 through a temperature ramp. One way ANOVA was used in (F-H) (ns $p>0.123$, *$p<0.033$,
376 **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$). (**i**) Sequence alignment indicating the differences in
377 sequence compared to VHH_WT. CDR2 and CDR3 are indicated by blue boxes, positions
378 identical to VHH_WT are indicated with a dot (.), positions explored with the EvolveX pipeline
379 are displayed bold on the sequence of VHH_WT. AbM CDR definition was used.

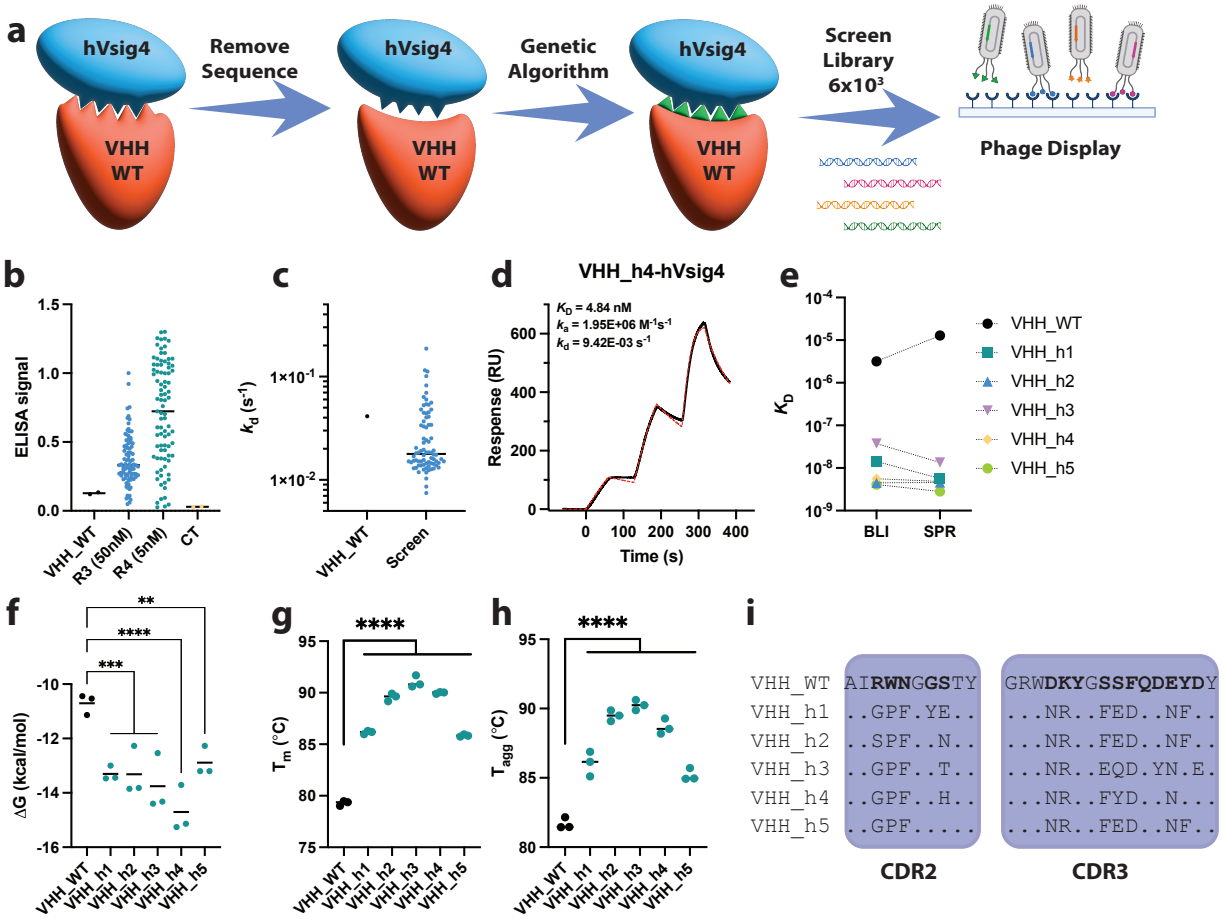


Fig. 2. Library design, screening, hit selection and validation of VHH's designed with EvolveX to bind human Vsig4. (**a**) Graphical representation of the EvolveX *in silico* pipeline, where crystal structures of the low affinity complex between human Vsig4 (hVsig4) and VHH_WT were used as templates to *in silico* design sequences of CDR2 and CDR3 that should bind hVsig4 with high affinity, while displaying good stability. A small library was subsequently screened using phage display. Icons of DNA strands and phage display were created with BioRender.com. (**b**) Colonies picked after phage display selection were tested for binding to hVsig4 using ELISA. VHH_WT was used as a positive control, colonies picked after two different rounds of selection were screened. (**c**) Off-rates of the 82 unique sequences, determined using BLI. (**d**) Example SPR binding curve of 1 out of 5 hits (VHH_h4) binding to hVsig4 with low

391 nanomolar affinity. (**e**) Affinity constants of all 5 hits and VHH_WT binding to hVsig4 determined
392 using BLI and SPR. (**f**) ΔG values determined through curve fitting based on data shown in G. (**g**)
393 Melting temperatures (T_m) for 5 hits and VHH_WT based on Intrinsic Tryptophan Fluorescence
394 (ITF) through a temperature ramp. (**h**) Aggregation onset temperatures (T_{agg}) for 5 hits and
395 VHH_WT based on Right Angle Light Scattering (RALS) through a temperature ramp. One way
396 ANOVA was used in (F-H) (ns $p>0.123$, *$p<0.033$, **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$).
397 (**i**) Sequence alignment indicating the differences in sequence compared to VHH_WT. CDR2 and
398 CDR3 are indicated by blue boxes, positions identical to VHH_WT are indicated with a dot (.),
399 positions explored with the EvolveX pipeline are displayed bold on the sequence of VHH_WT.
400 AbM CDR definition was used.

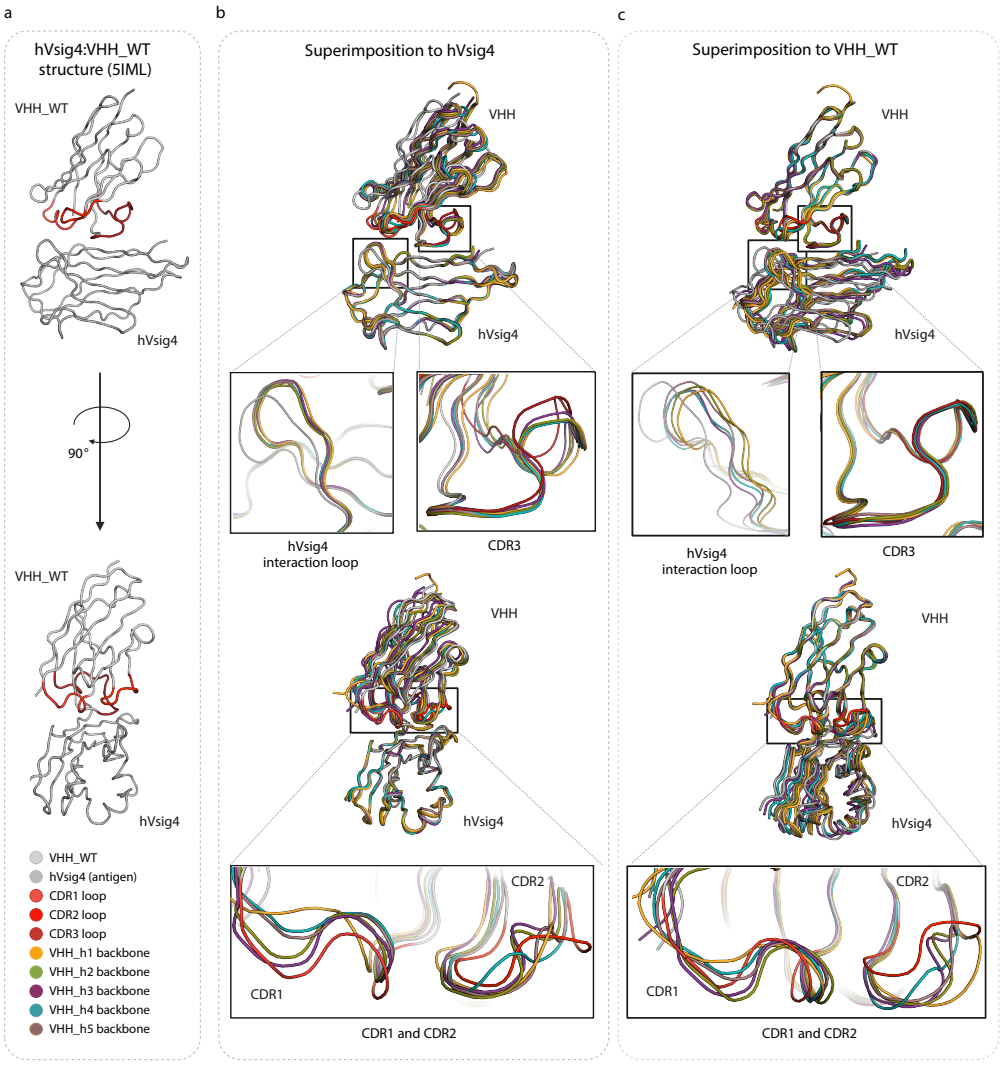


Fig. 3. X-ray crystallography high resolution structures of VHH_h1-h5 show epitope retention and conservation of CDR conformation. (**a**) hVsig4:VHH_WT reference structure (PDB:5IML) represented in 2 views with CDRs shown in red. (**b**) Superimposition of the hVsig4:VHH_WT structure with experimental structures of five in silico designed VHHs in complex with hVsig4 aligned to hVsig4 structure. (**c**) Superimposition of the hVsig4:VHH_WT structure with experimental structures of five in silico designed VHHs in complex with hVsig4 aligned to VHH_WT structure. Superimpositions were generated in Pymol using the align function.

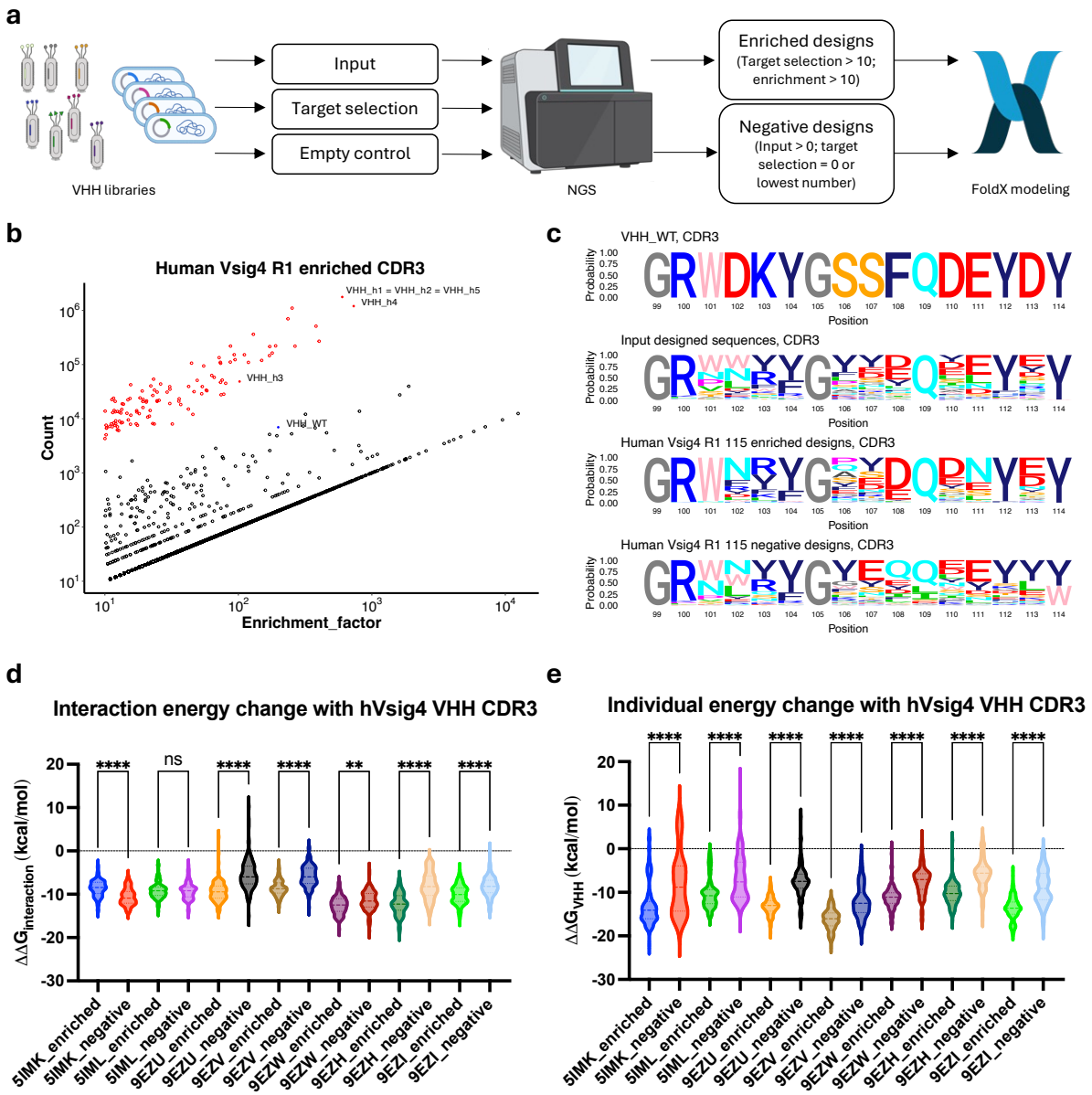


Fig. 4. FoldX modeling of NGS-identified hVsig4 CDR3 designs reveals energy change differences between enriched and negative CDR designs. (**a**) Graphical representation of the NGS preparation, analysis and FoldX modeling workflow. Input, target selection, and empty control libraries were analyzed by NGS. Enriched and negative CDR designs were filtered based on specific criteria and subsequently modeled using the FoldX force field. Icons of phages and NGS sequencer were created with BioRender.com. (**b**) Count and enrichment factor of human

Vsig4 enriched CDR3 sequences (count > 10, enrichment factor > 10, enrichment factor is the ratio of the count in target selection to the count in empty control). Red dots are the sequences designed by EvolveX, red solid dots are the sequences from identified hits, blue solid dot is VHH_WT, and black dots are non-designs or mutants generated during the PCR or phage selection stages. (**c**) Sequence logo plots of VHH_WT CDR3, input CDR3 designed sequences, 115 hVsig4 enriched CDR3 designed sequences and 115 negative CDR3 designed sequences. AbM CDR definition was used. (**d**) Comparison of the interaction energy change ($\Delta\Delta G_{interaction}$) between enriched and negative hVsig4 CDR3 designs. One way ANOVA was run between each group (ns $p>0.123$, *$p<0.033$, **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$). (**e**) Comparison of the VHH individual energy change ($\Delta\Delta G_{VHH}$) between enriched and negative hVsig4 CDR3 designs. One way ANOVA was run between each group (ns $p>0.123$, *$p<0.033$, **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$).

429 # Methods:

430 **Sequence Design: Template generation**

431 The original PDB structures of the wildtype VHH were downloaded for the PDB. Two low affinity
432 complexes of wildtype VHH bound to human Vsig4 (hVsig4; 5IMK and 5IML) and two high
433 affinity complexes of wildtype VHH bound to mouse Vsig4 (mVsig4; 5IMM and 5IMO).
434 Sidechains were subsequently optimized using the FoldX [17] command *RepairPDB*. Backbones
435 were optimized using the ModelX [14-16] command *Vibrate.* It generates 1 by 1 Degree rotations in
436 a range for -20 to 20 Degrees around an imaginary axis along N and C for every amino acid,
437 approximating phi, psi torsional rotations but reducing the number of rotations to one angle. The
438 human bound structure 5imk was the starting point to explore alternative backbones for CDR3
439 using the *Bridging* command in ModelX. ModelX allows to digest customized peptide fragments
440 databases adapted to the different modeling requirements. The peptide fragments were used to
441 graft CDR3 loop conformations generating conformational variability, in our case we created *in*
442 *situ* a database using 120000 X-ray structures with resolution <= 2.5 Å. 18544 backbone moves
443 were generated and selected using these criteria, first they were filtered out those having backbone-
444 backbone clashes and then selected those having a number of contacts bigger than 36, the
445 remaining 85 models that passed the filters were ranked according to their FoldX interaction
446 energy and 14 were selected, resulting in 14 structures with slight movements in CDR3. The 22
447 structures were subjected to *de novo* sequence design of CDR2 and CDR3.

448 **Sequence Design: Explore positions**

449 Each amino acid part of CDR2 or CDR3 (according to Chothia numbering scheme [25]) and within
450 8 Ångström from the ligand was explored separately for each possible amino acid using

451 *BuildModel*, after which the effect of each mutation on stability and interaction energy was
452 calculates using the FoldX commands *BuildModel* and *AnalyseComplex*, respectively.

453 **Sequence Design: Reduce search space**

454 To reduce the sequence space, the positions were filtered. Positions were removed where only the
455 wildtype did not result in major destabilization of the VHH (ddG_binder < 1). Additionally,
456 positions where the wildtype amino acid was best for VHH stability (ddG_binder rank = 1) and
457 where mutations had little effect on interaction energy (ddG_AC variation < 0.1). Positions where
458 the wildtype amino acid is Proline were not considered either. Mutations that destabilize the VHH
459 (ddG_binder > 1 kcal/mol) or reduce the interactions energy (ddG_AC > 1kcal/mol) were removed
460 from the search space.

461 **Sequence design: Genetic algorithm (EvolveX)**

462 Starting sequences were randomly generated based on the remaining search space. Each starting
463 sequence is evolved through a process based on a genetic algorithm [26]. The 50 start points (threads)
464 for each template are used as input for the genetic algorithm. The genetic algorithm was run for
465 500 generations. For each generation 10% of the 50 are recombined with each other and 90% is
466 subjected to a point mutation. Recombination is done by randomizing the identities of each parent
467 for each explored position and splitting the resulting list of mutations at a random point. Point
468 mutations are selected for a random position, while selecting a new amino acid for that position
469 semi-randomly according to a frequency table based on amino acid distributions of paratopes [27].
470 The effect of the mutation event is calculated using FoldX commands *BuildModel*,
471 *AnalyseComplex*, *Stability*, TANGO [18]. Thresholds are set manually for VHH stability
472 (dG_binder) and VHH intraclashes (IntraClash_binder). A mutation in rejected if the difference
473 with the parent is positive and the value of the mutant is above the manually set threshold. When

474 these criteria are met, the main driver is interaction energy (ddG_AC < 0 kcal/mol) which allows
475 the mutant to be accepted and the metropolis criterion is used when the interaction energy is
476 decreased (ddG_AC >= 0 kcal/mol). If the mutant is accepted it is used as input for the next
477 generation, if it is rejected the parent is used as input.

478 **Sequence selection**

479 R was used to select from the 500K accepted sequences. Only sequences are considered that have
480 a satisfactory interaction energy with the ligand (dG_AC < -17.5 kcal/mol), good VHH stability
481 (dG_binder < -5 kcal/mol), low predicted aggregation propensity (total TANGO score < 450). For
482 the remaining models, VHH intracavities (Cav_binder), interface cavities (Cav_Int) and VHH net
483 charge (ChAb) were calculated using YASARA [28]. Z-scores were calculated for all three variables
484 based on their distributions in known VHH complexes in the PDB. Models with Z-scores above 2
485 for any of the three variables were excluded. The three Z-scores were summed. The models with
486 the lowest summed Z-score and designed to bind mouse Vsig4 were selected to be included in the
487 library (AffinityMouse), with the highest summed Z-score were selected as negative control
488 (NegativeControlMouse). The models with the lowest summed Z-score and designed to bind
489 human Vsig4 were selected to be included in the final library (AffinityHuman), 207 with the
490 highest summed Z-score were selected as negative control (NegativeControlHuman).

491 **Vsig4 preparation**

492 Mouse recombinant Vsig4 (mVsig4) used in phage selection, screening, and VHH affinity
493 measurement was bought from Sino Biological (50187-M08H-100) and biotinylated with EZ-Link
494 NHS-PEG4-Biotin (Thermo Scientific, A39259) according to the manufacturer instructions. The
495 mVsig4 material used for crystallization (amino acids 20-139, with C-terminal His tags) was
496 cloned into plasmid pVDS101 (a modified version of pHEN6), expressed in TG1 *E. coli* strain,

497 and purified from the periplasmic extract by a Ni-charged column followed by a size exclusion
498 chromatography (SEC) (Superdex 75 16/600 column).
499 Human recombinant Vsig4 (hVsig4) used in phage selection, screening, and VHH affinity
500 measurement (amino acids 20-137, with C-terminal Avitag and His tags) was cloned into
501 pVDS101 plasmid and expressed in AVB101 *E. coli* strain (AVIDITY) to allow *in vivo*
502 biotinylation, following the suggested protocol. Then the protein was extracted from periplasm
503 and purified by Ni-NTA chromatography. The hVsig4 protein used in crystallization (amino acids
504 20-137, with C-terminal His tags), was cloned into pLTDA104 plasmid (also a modified version
505 of pHEN6), expressed in TG1 *E. coli* strain and purified following the same procedures with
506 mVsig4.

507 **Cloning of the synthetic Vsig4 CDR23 VHH library**

508 The oligo pool containing 12,405 sequences (12,405 CDR pairs, 3200 unique CDR2s, 9884 unique
509 CDR3s), encoding for the region from CDR2 to CDR3 of the VHHs, flanked by two restriction
510 sites (BpiI and Eco91I) was ordered at Twist Biosciences. The oligo pool was amplified with a
511 KAPA HiFi HotStart ReadyMix PCR Kit (Roche, KK2602). The amplified product was then
512 restriction digested and ligated into the phagemid vector pP002ph008 (a modified version of
513 pMECS, which encodes for the non-library encoded parts of VHH_WT, followed by 3xFlag/6xHis
514 tag at the C-terminus of the VHH insertion site) with T4 DNA ligase (Thermo Scientific, EL0011)
515 at 16°C overnight. The purified ligation sample was transformed into the TG1 electrocompetent
516 cells (Lucigen, 60502-2). The resulting library had a size of $>10^7$ CFU and a high insertion rate (>
517 90 %). At last, the phage library was produced with the infection of M13KO7 helper phage
518 (Thermo Fisher Scientific, 18311019) according to the protocol published by Pardon et al[29].

519 **Phage library selection**

520 The VHH displaying phage library was selected against multiple rounds of mVsig4 or hVsig4
521 according to the standard protocol [29] (Supplementary Fig.1 and Supplementary Fig.23). Briefly,
522 $2x10^{11}$ CFU of phages, biotinylated antigens, and streptavidin magnetic beads (Thermo Scientific,
523 88817) were firstly blocked at PBS buffer with 2% of milk for at least 0.5h at room temperature.
524 Then the blocked phages and antigens were co-incubated for 2h at room temperature to allow the
525 binding of VHH-displaying phages with the targets. Non-binding phages were discarded with 5
526 consecutive washing steps with PBS/2% milk, whereas bound phages were captured by the
527 streptavidin beads and eluted by trypsin. The eluted phages were rescued with the exponentially
528 grown TG1 cells for the next round of selection. In the end, the output phage libraries from desired
529 rounds of selection were sub-cloned into the vector pLTDA118 (a modified pHEN6 vector with
530 C-terminal 3xFlag/6xHis tag) and transformed into TG1 cells. Single colonies were randomly
531 picked to grow in the 96-well deep-well plate for the preparation of periplasmic extracts, and
532 Sanger sequenced as previously described [30].

533 **AlphaScreen**

534 AlphaScreen FLAG (M2) detection kit (PerkinElmer, 6760613M) was used for the initial
535 screening of the mVsig4 hits. In short, in a 384-well Optiplate (PerkinElmer, 6007290), 5µL of
536 diluted biotinylated mVsig4, 5µL of diluted periplasmic extracts in a 96-well plate from different
537 clones, and 10µL of anti-FLAG conjugated acceptor beads (50µg/ml) were incubated together for
538 1h at room temperature in the dark. Then 5µL of streptavidin donor beads (100µg/mL) was added
539 and incubated in the dark for 25-30 minutes. The signals were read (excitation 680nm; emission:
540 520-620nm) using the EnVision multilabel plate readers (PerkinElmer).

541 **ELISA screen**

542 Human Vsig4 hits were first screened in an ELISA assay. In brief, 100μL of biotinylated hVsig4
543 (2μg/mL) was coated in the ELISA plate in advance. The plates were blocked with PBS/1% BSA
544 for 2h at room temperature. Subsequently, 100μL of periplasmic extracts were added per well, and
545 incubated for 1.5h at room temperature. Then 100μL/well of anti-FLAG-HRP antibody (Sigma-
546 Aldrich A8592) (1:2500 dilution) was added and incubated for 1 more hour. At last, TMB solution
547 (50μL/well) was used to detect and after 1-5 minutes, 50μL/well of H_2SO_4 was added to stop the
548 reaction. The plates were then read at 450nm of wavelength in a plate reader.

549 **Biolayer interferometry (BLI) off-rate assay**

550 The clones with unique sequences and significant AlphaScreen or ELISA signal (signal/noise, S/N
551 > 3) were further selected, re-arrayed in a 96-well plate, and screened by a biolayer interferometry
552 (BLI) off-rate assay on Octet RED96 (Sartorius). Ni-NTA biosensors (Sartorius, 18-5101) were
553 first used to assess the expression of different clones. After equilibration of the biosensors for at
554 least 10 minutes in 1x Sartorius/Octet kinetic buffer and an extra 120 seconds for the baseline, the
555 biosensors were dipped into 10x diluted periplasmatic extracts for 300 seconds, followed by 300
556 seconds of dissociation in 1x kinetic buffer. Afterward, streptavidin (SA) biosensors (Sartorius,
557 18-5019) were used to measure the off rate of different clones. With the same pre-wet procedure
558 and the baseline setting, biotinylated Vsig4 material was first loaded on the biosensors, after which,
559 the tips were sequentially submerged in baseline wells, association wells with diluted periplasm
560 samples (300s), and finally into the dissociation wells (300s). Data was processed on the Octet
561 RED analysis software.

562 **VHH expression and purification**

563 VHH candidates used for initial affinity measurement were in the vector pLTDA118 (a modified
564 pHEN6 vector with C-terminal 3xFlag/6xHis tag) and expressed in the TG1 cells following the
565 standard protocol [29]. The purification of these VHHs was done by either the Ni-NTA
566 chromatography or the AmMag™ SA Plus System (Genscript, L01013) with Ni magnetic beads
567 (Genscript, L00776). VHHs used in stability study and crystallization were cloned into the vector
568 pLTDA104 (a modified pHEN6 vector with C-terminal 6xHis tag) and expressed in the TG1 cells
569 following the same procedure. These VHHs were then purified by a Ni-charged column followed
570 by size exclusion chromatography (SEC).

571 **Biolayer interferometry (BLI) affinity measurement**

572 Purified VHHs were used to determine the affinity with mVsig4 and hVsig4 on Octet RED96
573 (Sartorius). Briefly, streptavidin (SA) biosensors (Sartorius, 18-5019) were first equilibrated for at
574 least 10 minutes in 1x Sartorius/Octet kinetic buffer. Subsequently, the tips were dipped into the
575 biotinylated Vsig4 (3-5µg/mL) target for 30-60 seconds. Afterward, the tips were sequentially
576 submerged in the 1x kinetic buffer for the baseline (120s), VHH dilution samples for the
577 association (300s), and in the end back to the 1x kinetic buffer for the dissociation (300s). Fitting
578 and binding kinetics determination was performed with a 1:1 model interaction on the Octet RED
579 analysis software.

580 **Surface plasmon resonance (SPR) affinity measurement**

581 Binding kinetics and affinity of various VHHs for the human and murine Vsig4 were evaluated by
582 surface plasmon resonance on a BIAcore T200 instrument (Cytiva) with a running buffer
583 composed of 10mM HEPES, 150mM NaCl & 0.005% Tween 20. The assay format involved ligand
584 capture on a Sensor S Sensor Chip SA chip. Briefly, the biotinylated ligands were immobilized by

585 non-covalent capture (binding to streptavidin) following the instructions provided with the Cytiva
586 chip. mVSIG4 was captured on flow cell 2, and hVSIG4 was captured on flow cell 3, leaving flow
587 cell 1 as a subtractive reference. The capture level of the mVsig4 was targeted between 100 and
588 200 resonance units, and the hVsig4 was targeted between 1200 and 1300 resonance units. A serial
589 dilution of the VHHs was flowed over the immobilized Vsig4 (50µL/min for 1 minute) and
590 allowed to dissociate for 4 minutes. The capture surface was regenerated with a 60-s injection of
591 3 M MgCl2 (50 µL/min for 1 min). A 3-fold concentration dilution series of each VHH variant
592 ranging was used to analyze binding to mVsig4. For the VHH_WT, VHH_m1-VHH_m4 the
593 concentration series ranged from 0.25nM to 20nM. For the VHH_h1-VHH_h5 the concentration
594 dilution series ranged from 6.2nM to 500nM. A 3-fold concentration dilution series of each VHH
595 variant ranging was used to analyze binding to hVsig4. For the WT and VHH_m1, the
596 concentration dilution series ranged from 3.7nM to 300nM. For VHH_h1-VHH_h5 and VHH_m4
597 the concentration dilution series ranged from 0.25nM 20nM. For the VHH_m2 and VHH_m3, the
598 concentration dilution series ranged from 2.1nM to 167nM. All sensograms were analyzed using
599 a 1:1 Langmuir binding model with software supplied by the manufacturer to calculate the kinetics
600 and binding constants.

601 **Thermal stability measurement**

602 The temperature at which 1.0 mg/mL VHHs unfold (T_m, unfolding temperature) and aggregate
603 (T_{agg}, aggregation onset temperature) using an UNcle device (Unchained Labs, CA, USA). More
604 specifically, the T_m and T_{agg} were determined simultaneously via intrinsic fluorescence (ITF) and
605 static light scattering (SLS) measurements, respectively, during gradual heating from 15 to 95°C
606 at 0.3°C/min. The excitation wavelength of 280 nm was used for both measurements, with the
607 emission wavelength ranging from 250 to 720 nm for the intrinsic fluorescence and corresponding

608 to 266 nm for the SLS. The intrinsic fluorescence values between 320 and 370 were used to
609 calculate the barycentric mean values (BCMs), resulting in a smooth curve. Curve fitting with R-
610 studio was used to calculate the T_m value corresponding to the inflection point of the sigmoid
611 curve. The T_{agg} value was determined using manually set thresholds, to allow for proper
612 comparison between variants, ignoring baseline noise. All samples were measured at least in
613 triplicate.

614 **Chemical stability measurement**

615 Chemical denaturation was performed using both Urea (0-9.2M) and Guanidine Hydrochloride
616 (GdnHCl, 0-5.5M) using a gradient of 24 steps. Pipetting the different steps was done using a robot
617 (Opentrons OT-2) from a protein stock solution of 1 mg/mL in PBS with a final concentration of
618 0.04 mg/mL and freshly prepared denaturant stock solution of 10M Urea or 6M GdnHCl. All stock
619 solutions were filtered using 0.22 μM filter. After pipetting, plates were spun down in a centrifuge
620 at 1000g for 2 minutes. The plates were incubated for 2 hours at 25ºC with a lid to avoid
621 evaporation of the sample. Samples were measured using the standard protocol on the SUPR-CM
622 (ProteinStable).

623 **Crystallization of the VHH:hVsig4 complexes**

624 For VHH_h1, VHH_h2 and VHH_h3 complexes with hVsig4, 1.2 molar excess of VHH was
625 added to hVsig4, incubated for 15 minutes at room temperature and purified by SEC on a Superdex
626 75 Increase 10/300 (Cytiva) column using HBS pH 7.4 buffer. Fractions containing the complex
627 were pooled and concentrated to 4 - 6 mg/ml. Crystallization screens were set up using a Mosquito
628 liquid handling robot (TTP Labtech) in sitting-drop format using 100 nl of protein mixed with 100
629 nl of mother liquor in SwissSci 96-well triple drop plates incubated at 287 or 294 K.
630 VHH_h2:hVsig4 complex hit was obtained by micro-seeding using crystal seeds of

631 VHH_h1:hVsig4 complex grown in condition 0,17M ammonium sulfate, 15% glycerol, 30,5%
632 PEG 4000. Seed stock was prepared by crushing the crystals with a loop in original mother liquor
633 drop. Crystal seeds were added using a cat whisker to drops containing 100 nl of protein mixed
634 with 100 nl of mother liquor in multiple conditions. Plate was incubated at 287K. Best diffracting
635 crystal of this complex was obtained in 0,1M Na-citrate pH 5,6, 20% 2-propanol and 21% PEG
636 4000. Crystals were cryoprotected in mother liquor containing glycerol or ethylene glycol prior to
637 being cryocooled in liquid nitrogen. Diffraction data was collected at 100 K at the P13 microfocus
638 beamline at PETRA III (Hamburg, Germany).
639 For VHH_h4 and VHH_h5 complexes with hVsig4, VHH and hVsig4 were mixed in a 1:1 molar
640 ratio respectively, incubated for 40 minutes at room temperature and purified by size exclusion
641 chromatography (SEC) on a Superdex 75 Increase 10/300 (Cytiva) column using 20mM Tris pH
642 7.5, 5% Glycerol and 150mM NaCl buffer. Fractions containing the complex were pooled and
643 concentrated to 3.5-4.5 mg/ml. The crystallization screens were set up using the hanging drop
644 vapor-diffusion technique by mixing 1 µl of protein solution with 1 µl of mother liquor. For this
645 purpose crystallization conditions consisting of 0.1M Bis-Tris pH 6.0-6.5 and 25-30% PEG 3350
646 were used, based on the previously reported conditions that yielded crystals of native
647 VHH_WT:hVsig4 complexes [21]. Plates were incubated at 293K. Obtained crystals were
648 cryoprotected in mother liquor containing 15% glycerol prior to being cryocooled in liquid
649 nitrogen. Diffraction data was collected at 100 K at the Xaloc beamline at ALBA (Barcelona,
650 Spain). Detailed crystallography setup and data collection statistics are summarized in
651 Supplementary Table.7.

652 **Determination of crystallographic structures of the VHH:hVsig4 complexes**

653 Datasets for complexes of VHH_h1, VHH_h2 and VHH_h3 were integrated and scaled using XDS
654 [31] and AIMLESS and the data quality was analyses by Phenix Xtriage and datasets for complexes
655 of VHH_h4 and VHH_h5 with hVsig4 were processed with autoPROC [32].

656 Initial phases for the VHH:hVsg4 complexes were obtained by maximum-likelihood molecular
657 replacement with Phaser [33] using the coordinates of native structures of hVsig4 and VHH (PDB
658 ID 5IML) as a search model. A model for each of the VHH variants was manually rebuilt in Coot
659 [34] and the refinement of coordinates and atomic displacement parameters for all the structures was
660 performed in Refmac [35] and Phenix refine [36]. Model and map validation tools in Coot and the
661 PHENIX suite were used throughout the workflow to guide improvement and validate the quality
662 of crystallographic models. Refinement statistics are summarized in Supplementary Table.7.

663 **NMR spectroscopy**

664 The NMR samples contained 0.8-1 mM [^{13}C,^{15}N] or ^{15}N VHH in 20 mM sodium phosphate 50
665 mM NaCl pH 6.5, 0.02% NaN_3, and 10 % D_2O for the lock. All NMR spectra were acquired at
666 313 K on a Bruker Avance III HD 800 MHz spectrometer, equipped with a triple-resonance TCI
667 cryoprobe for enhanced sensitivity. Three-dimensional experiments were acquired with a non-
668 uniform sampling (25-50%) as implemented in TopSpin 3.7 (Bruker). The NMR data were
669 processed in TopSpin 3.7 (Bruker) or MddNMR [37] and NMRPipe [38], and analyzed in CCPNMR
670 [39]. Assignments of protein backbone resonances were obtained from [^{13}C,^{15}N]-labelled samples of
671 C-terminally triFlag- and His-tagged VHH_WT and VHH_h4 by establishing intra- and inter-
672 residue connectivities in 3D HNCACB, HN(CO)CACB, HNCO, HN(CA)CO, and
673 HBHA(CO)HN spectra at the corresponding peak frequencies in the 2D [^{1}H,^{15}N]-HSQC spectrum.
674 The assignments were subsequently transferred to [^{1}H,^{15}N]-HSQC spectra of ^{15}N-labelled C-

675 terminally His-tagged VHH_WT and VHH_h4 constructs and verified by 3D ^{15}N-NOESY-HSQC
676 (mixing time 120 ms). The assigned chemical shift lists were deposited in Biological Magnetic
677 Resonance Data Bank (BMRB) entries 52361 (VHH_WT) and 52360 (VHH_h4). The VHH
678 secondary structure (Supplementary Fig.16 and Supplementary Fig.18) was predicted from the
679 backbone chemical shifts using the CSI function and the DANGLE module [40] in CCPNMR [39] or
680 the web-based RCI [41].
681 Relaxation measurements were performed on ^{15}N-labelled samples of C-terminally His-tagged
682 VHH_WT and VHH_h4. Longitudinal (R_1) and transverse (R_2) relaxation rates were obtained
683 from [^{1}H,^{15}N] HSQC-based pseudo-3D spectra with interscan delays of 2.5 s. The inversion
684 recovery experiment (R_1) had relaxation delays of 10, 100, 1000, 400, 600, 800, 500, 1200, 200,
685 50, 100, 300, 900 ms (in this order), while Carr-Purcell-Meiboom-Gill (CMPG) measurments (R_2)
686 employed relaxation delays of 17, 68, 136, 170, 85, 34, 51, 34, 119, 153, 102 ms (in this order).
687 Steady-state, heteronuclear {^{1}H}-^{15}N NOEs were measured in an interleaved experiment with and
688 without ^{1}H saturation and the relaxation delay of 6 s and performed in duplicate to estimate the
689 experimental errors. For each analysed [^{1}H,^{15}N]-HSQC peak, the R_1 and R_2 values were obtained
690 from exponential decays of signal intensities in CCPNMR [39], while {^{1}H}-^{15}N NOEs were
691 determined as ratios of peak intensities in the spectra with and without ^{1}H saturation. Relaxation
692 analysis was performed with Modelfree v 4.20 [42] using FastModelfree setup [43]. Initial estimates of
693 the diffusion tensor parameters were obtained with quadric_diffusion [44] from the experimental
694 R_2/R_1 data that excluded the residues with high R_2 values ($R_2 \geq \langle R_2 \rangle + \sigma_{R2}$) as likely exhibiting
695 significant internal motions. The hydrodynamic parameters and predicted R_2/R_1 ratios were
696 calculated with HYDRONMR [45] from the atomic coordinates of VHH_WT taken from the X-ray
697 structure of its complex with Vsgi4 (5IMM).

698 **FoldX modelling and analysis of Vsig4 top VHHs**

699 Sidechains in all nine available structures (PDBID: 5IMM and 5IMO for VHH_mVsig4, and
700 5IMK, 5IML, 9EZU, 9EZV, 9EVW, 9EZH, and 9EZI for VHH_hVsig4) were initially optimized
701 by using the FoldX command *RepairPDB*. Structure models of all VHH candidates including
702 VHH_WT, in complex with their respective target were generated using the FoldX *BuildModel*
703 function, by first mutating the explored residues to Alanine and then mutating to the corresponding
704 amino acids. The interaction energy ($\Delta G_{interaction}$) and VHH individual energy (ΔG_{VHH}) were
705 calculated using FoldX *AnalyseComplex* function and visualized in GraphPad Prism. Triplicate
706 modelling was performed for each complex. Differences between the candidates and VHH_WT
707 within each group were analyzed using two-way ANOVA in GraphPad Prism.

708 **NGS of Vsig4 libraries**

709 Library plasmids from the input, all rounds of selection output, and the empty control (none-target
710 selection) outputs generated in parallel with each round of selection, were used to perform NGS
711 sample preparation. In brief, 0.5ng of plasmid template was used in a 25 µL first-round of PCR
712 reaction following the guidelines of KAPA HiFi HotStart ReadyMix PCR Kit (Roche, KK2602).
713 During the first round, gene of interests (GOIs, sequence from CDR2 to CDR3) were enriched,
714 and unique molecule identifiers (UMIs) were introduced. The second round of PCR, conducted at
715 VIB Nucleomics Core (Leuven, Belgium) incorporated the sequencing adapters, and the samples
716 were subsequently pooled and sequenced on AVITI 2x300 Medium (PE 250) with 5% PhiX.

717 Raw data with around 30 millions of reads from the input, 10-20 millions of reads from R1
718 selection outputs and 1-2 millions of reads from R2-R4 selection outputs were acquired,
719 demultiplexed, and initially processed with Galaxy website. Paired reads were merged using
720 PEAR, and only reads with the correct length were retained. Additionally, UMI information were

721 used to deduplicate. DNA sequences from CDR2 to CDR3 were extracted and translated to protein
722 sequences. The count and enrichment factor (ratio of counts in the target selection output to counts
723 in the empty control output; 1 was used if the count in the empty control output is 0) of CDR2 and
724 CDR3 sequences from the input and R1 outputs were analyzed separately. Enriched CDR designs
725 were filtered from R1 selection for both mVsig4 and hVsig4 based on specific criteria (count >
726 10, enrichment factor > 10). An equivalent number of negative CDR designs were identified,
727 defined as sequences present in the input but absent or minimally represented in the R1 selection
728 outputs. Plots of the counts and enrichment of CDR sequences as well as CDR sequence logos,
729 were generated using R.

730 **FoldX modelling and analysis of NGS enriched VHH sequences**

731 Sidechains in all nine available structures (PDBID: 5IMM and 5IMO for VHH_mVsig4, and
732 5IMK, 5IML, 9EZU, 9EZV, 9EVW, 9EZH, and 9EZI for VHH_hVsig4) were initially optimized
733 by using the FoldX command *RepairPDB*. Structure models incorporating enriched or negative
734 CDR sequences were generated using the FoldX *BuildModel* function, by first mutating the
735 explored residues to Alanine and then mutating to the corresponding amino acids. The interaction
736 energy change ($\Delta\Delta G_{interaction}$) and VHH individual energy change ($\Delta\Delta G_{VHH}$) were calculated using
737 FoldX *AnalyseComplex* function. Differences between the models with enriched sequences and
738 models with negative sequences were analyzed using one-way ANOVA in GraphPad Prism.

739 **Data availability:** Crystallographic coordinates and final structure factors have been deposited in
740 the Protein Data Bank (PDB) with access codes: 9EZU (hVSIG4-VHH_h1), 9EZV (hVSIG4-
741 VHH_h2), 9EZW (hVSIG4-VHH_h3), 9EZH (hVSIG4-VHH_h4) and 9EZI (hVSIG4-VHH_h5).
742 The assigned chemical shift lists were deposited in Biological Magnetic Resonance Data Bank
743 (BMRB) entries 52361 (VHH_WT) and 52360 (VHH_h4).

744 **Code availability:** The code written to perform the sequence design by EvolveX is available at
745 https://github.com/SwitchLab-VIB/EvolveX/

746 **Methods References:**

747 25 Chothia, C. & Lesk, A. M. Canonical structures for the hypervariable regions of
748 immunoglobulins. *J Mol Biol* **196**, 901-917 (1987).

749 26 Katoch, S., Chauhan, S. S. & Kumar, V. A review on genetic algorithm: past, present, and
750 future. *Multimed Tools Appl* **80**, 8091-8126 (2021).

751 27 Akbar, R. et al. A compact vocabulary of paratope-epitope interactions enables
752 predictability of antibody-antigen binding. *Cell Rep* **34**, 108856 (2021).

753 28 Krieger, E. & Vriend, G. YASARA View - molecular graphics for all devices - from
754 smartphones to workstations. *Bioinformatics* **30**, 2981-2982 (2014).

755 29 Pardon, E. et al. A general protocol for the generation of Nanobodies for structural biology.
756 *Nat Protoc* **9**, 674-693 (2014).

757 30 Rue, L. et al. Novel Human/Non-Human Primate Cross-Reactive Anti-Transferrin
758 Receptor Nanobodies for Brain Delivery of Biologics. *Pharmaceutics* **15** (2023).

759 31 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

760 32 Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta*
761 *Crystallogr D Biol Crystallogr* **67**, 293-302 (2011).

762 33 Cao, L. et al. Design of protein-binding proteins from the target structure alone. *Nature*
763 **605**, 551-560 (2022).

764 34 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*
765 *Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

766 35 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures
767 by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255
768 (1997).

769 36 Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography.
770 *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018).
771 37 Orekhov, V. Y. & Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-
772 dimensional decomposition. *Prog Nucl Magn Reson Spectrosc* **59**, 271-292 (2011).
773 38 Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on
774 UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
775 39 Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a
776 software pipeline. *Proteins* **59**, 687-696 (2005).
777 40 Cheung, M. S., Maguire, M. L., Stevens, T. J. & Broadhurst, R. W. DANGLE: A Bayesian
778 inferential method for predicting protein backbone dihedral angles and secondary structure.
779 *J Magn Reson* **202**, 223-233 (2010).
780 41 Berjanskii, M. V. & Wishart, D. S. A simple method to predict protein flexibility using
781 secondary chemical shifts. *J Am Chem Soc* **127**, 14970-14971 (2005).
782 42 Mandel, A. M., Akke, M. & Palmer, A. G., 3rd. Backbone dynamics of Escherichia coli
783 ribonuclease HI: correlations with structure and function in an active enzyme. *J Mol Biol*
784 **246**, 144-163 (1995).
785 43 Cole, R. & Loria, J. P. FAST-Modelfree: a program for rapid automated analysis of
786 solution NMR spin-relaxation data. *J Biomol NMR* **26**, 203-213 (2003).
787 44 Lee, L. K., Rance, M., Chazin, W. J. & Palmer, A. G., 3rd. Rotational diffusion anisotropy
788 of proteins from simultaneous analysis of 15N and 13C alpha nuclear spin relaxation. *J*
789 *Biomol NMR* **9**, 287-298 (1997).

45 Garcia de la Torre, J., Huertas, M. L. & Carrasco, B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* **147**, 138-146 (2000).

Acknowledgments:

Thanks for support to the Unit of Structural Biology (IBMB-CSIC), to the Xaloc beamline at ALBA and to the Crystallography Platform at the Barcelona Science Park (PCB). We thank the staff of the EMBL beamline P14 (PETRA III, Hamburg, Germany) for beamtime allocation and technical support.

Funding: The Switch Laboratory was supported by the Flanders Institute for Biotechnology (VIB, grant no. C0401 to FR and JS) and the Fund for Scientific Research Flanders (FWO; SBO project grant S000722N to FR). Savvas N. Savvides's lab was supported by the Flanders Institute for Biotechnology (VIB, grant no. C0101 to SS). D.V acknowledges financial support from MICINN (Spain) through the Juan de la Cierva-Formación program (JDC2022-048697-I).

Author contributions:

Conceptualization: RvdK, SNS, MD, JS, LSP, FR

Methodology: RvdK, ZZ, IM, JDB, DC, ANV, SNS, MD, JS, LSP, FR

Investigation: RvdK, ZZ, IM, DV, TG, KM, JDB, DC, GO, CC, ANV

Visualization: RvdK, ZZ, IM, ANV

Funding acquisition: RvdK, JS, FR, LSP, SNS, MD, NG

Project administration: RvdK, JS, FR, LSP, SNS, MD

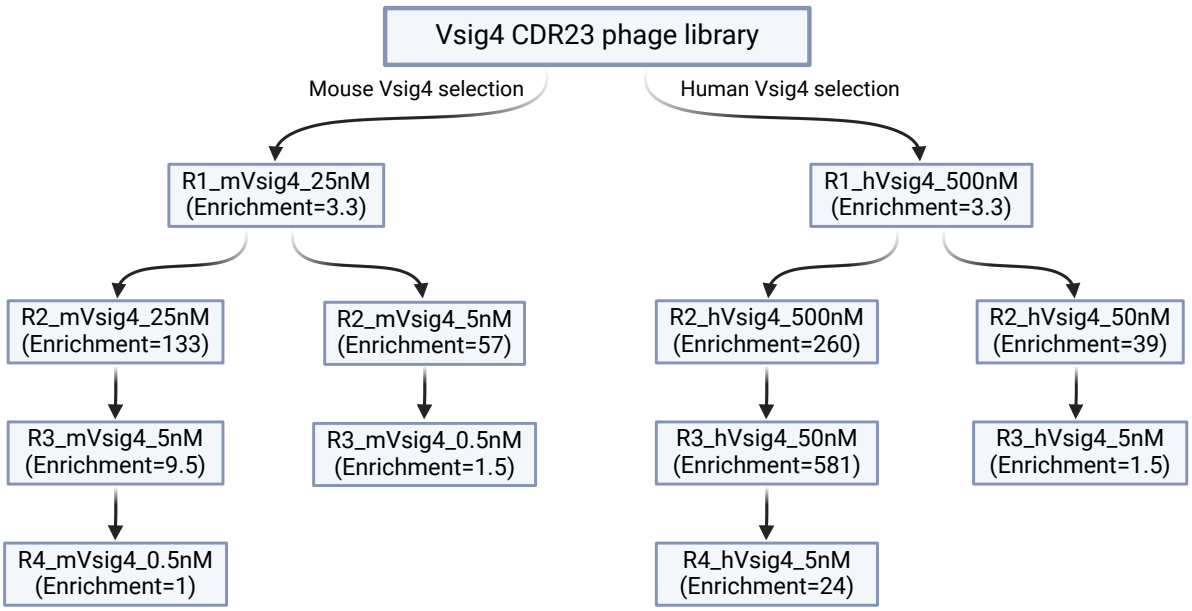
Supervision: RvdK, JS, FR, LSP, SNS, MD

Writing – original draft: RvdK, ZZ

Writing – review & editing: RvdK, ZZ, IM, DV, TG, KM, JDB, DC, GO, CC, ANV, SNS, MD, JS, LSP, FR

Competing interests: RvdK, JS, FR, LSP, JDB, DC. are co-inventors on E.U. provisional patent number EP 24170283.6 which covers the computational antibody design pipeline described here.

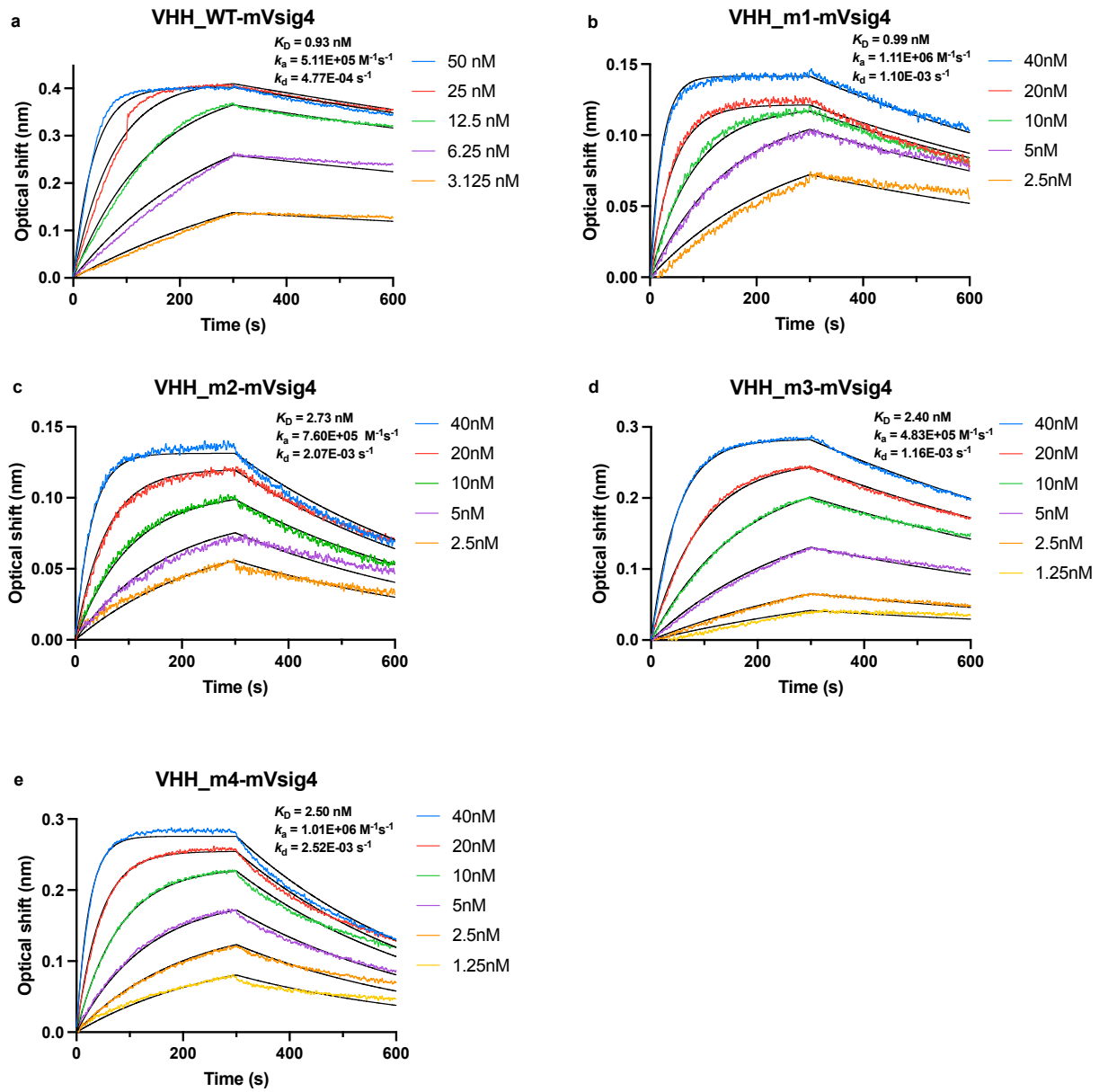
816 **Supplementary Figures:**



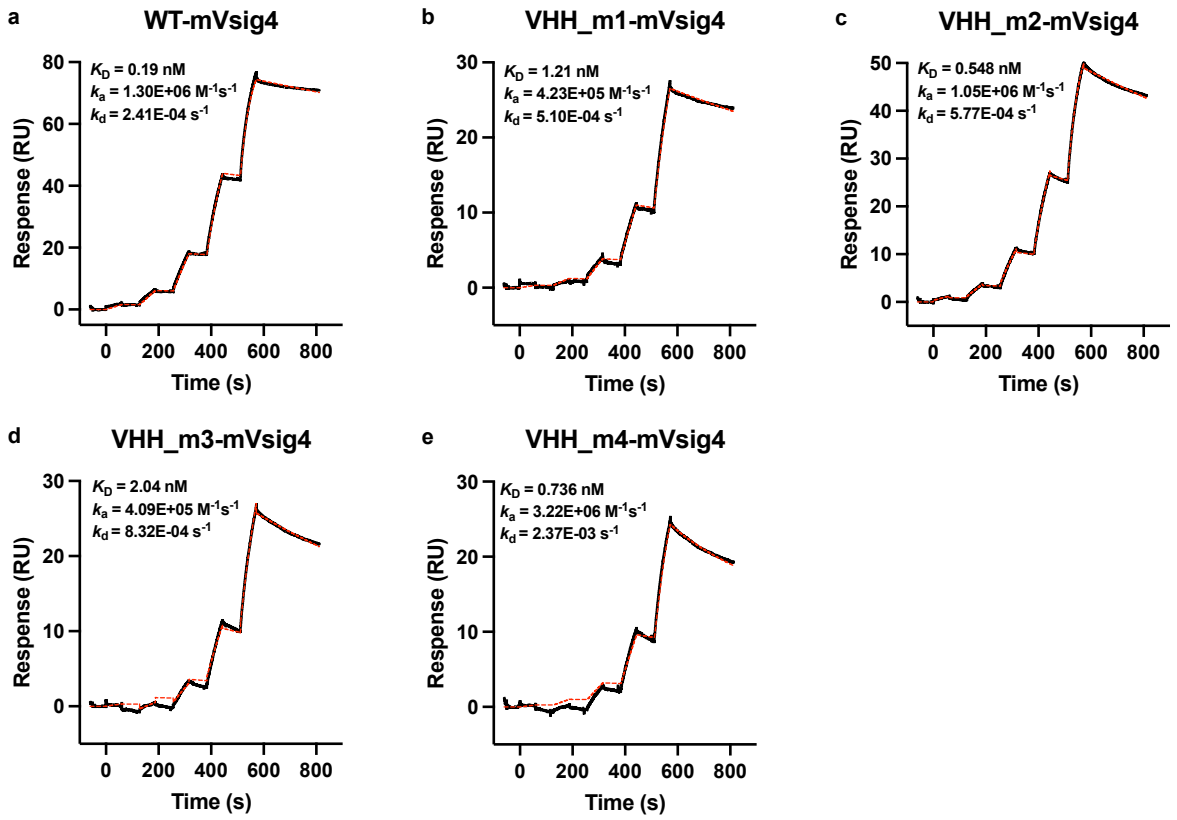
817

818 **Supplementary Fig.1.** Phage library selection rounds for the combined Vsig4 CDR23 library.

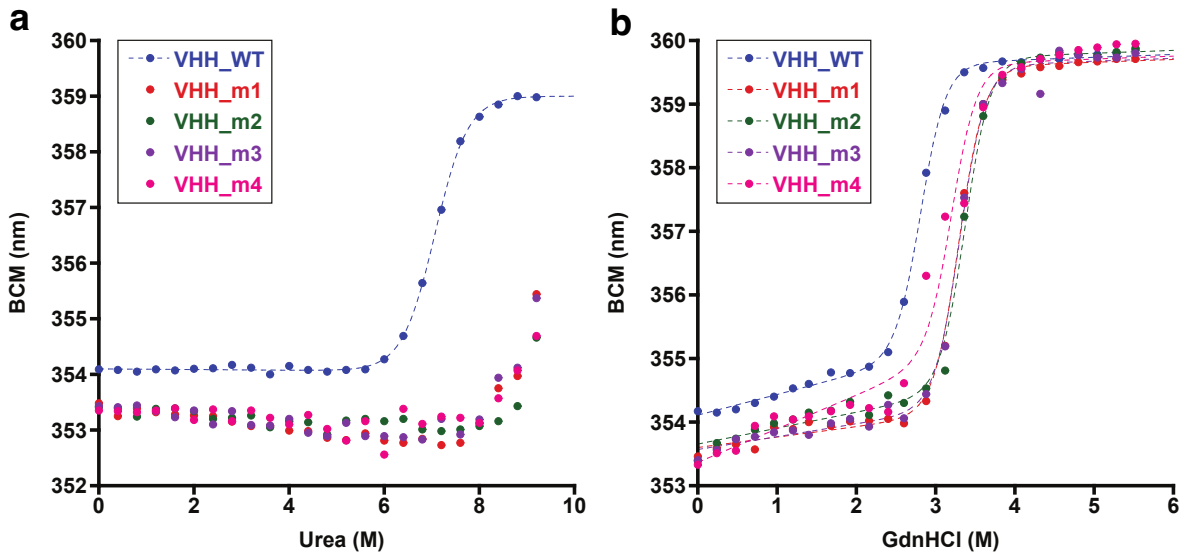
819 Graphical designs were created with BioRender.com.



Supplementary Fig.2. The binding curves (colored) of VHH_WT, VHH_m1-VHH_m4 against mVsig4 (BLI) with fitting curves (black).



Supplementary Fig.3. The binding curves (black) of VHH_WT, VHH_m1-VHH_m4 against mVsig4 (SPR) with fitting curves (red).

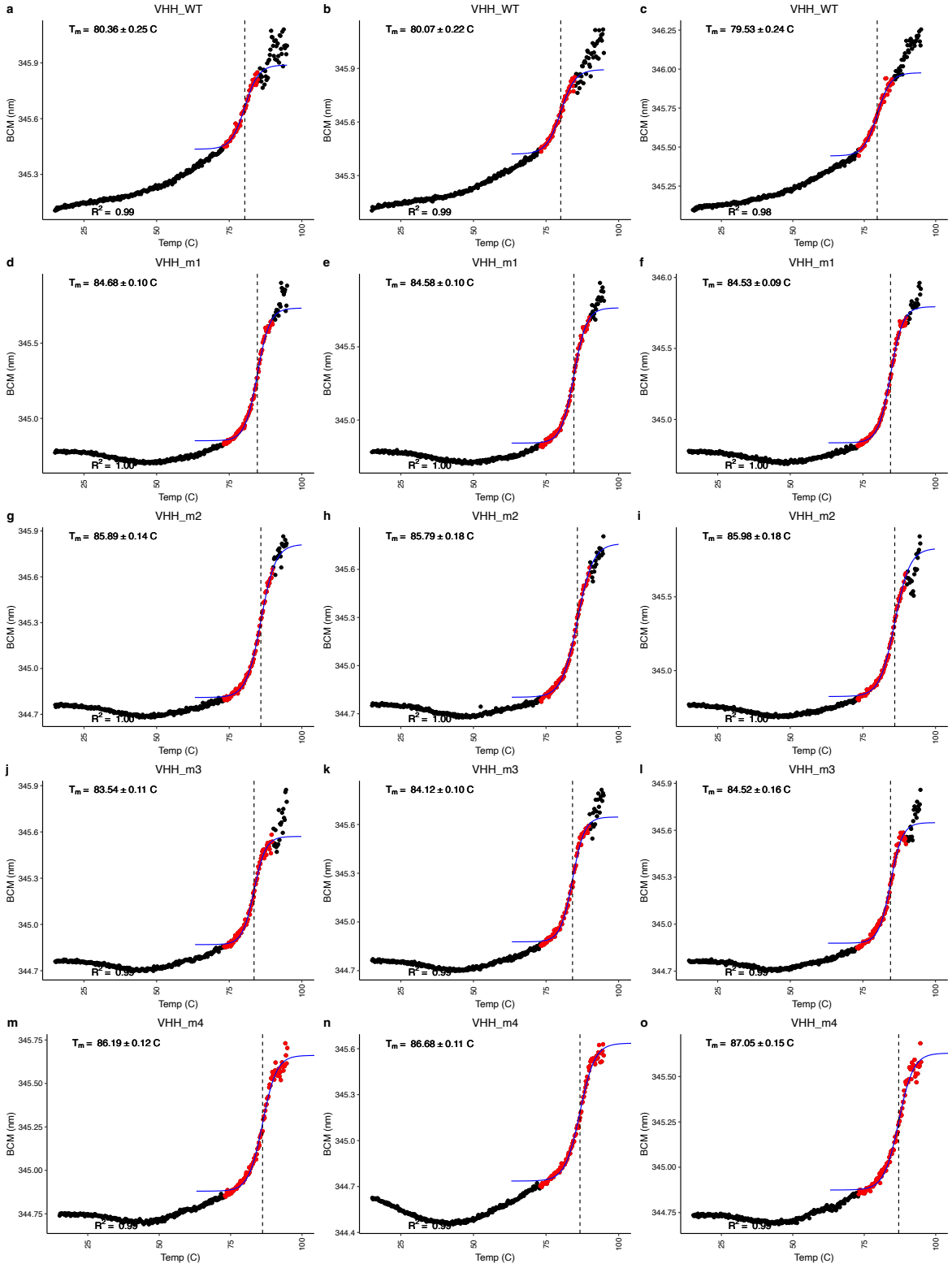


826

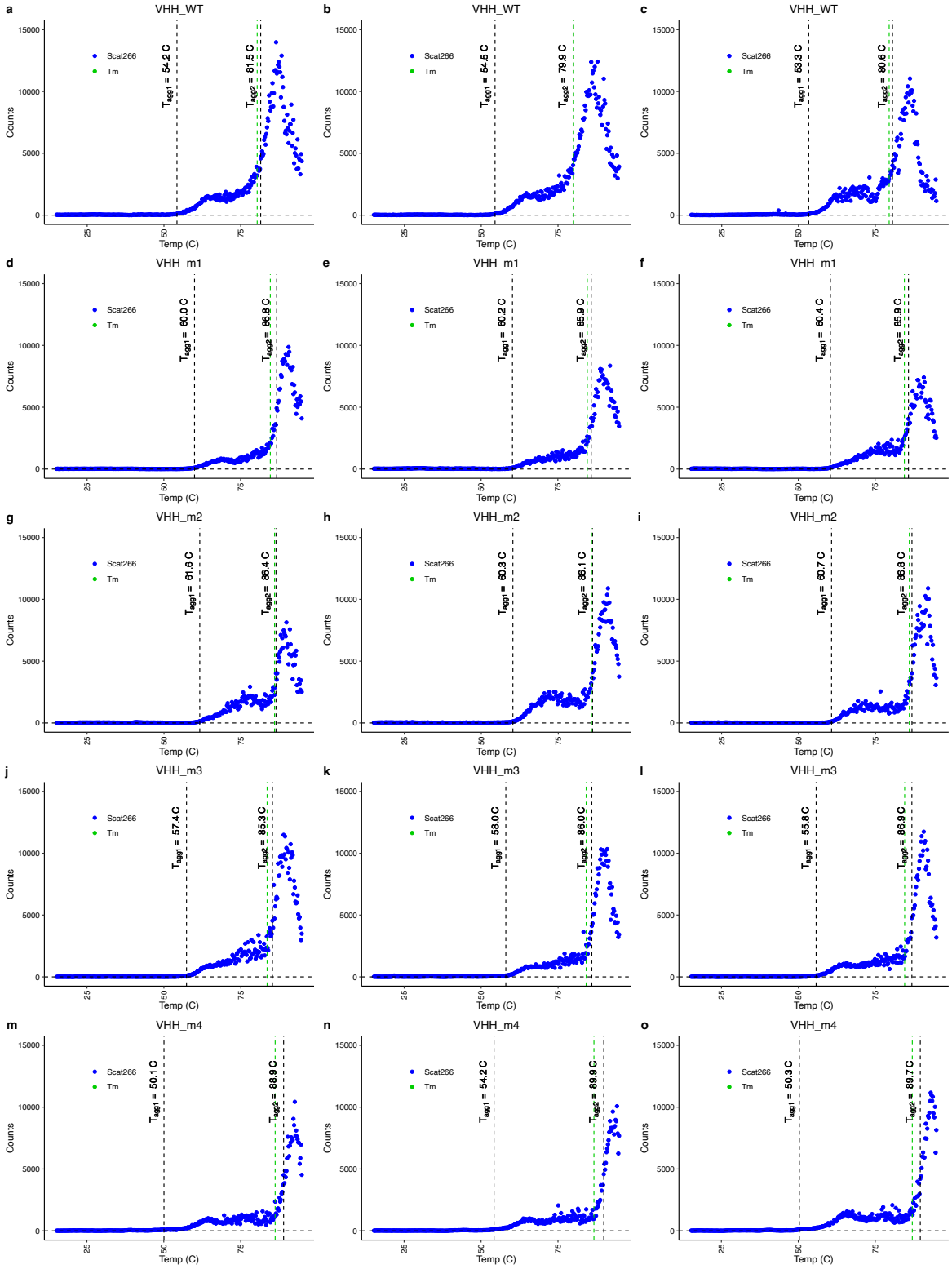
827 **Supplementary Fig.4. Representative chemical denaturation curves of VHH_WT, VHH_m1-**

828 **m4.** (**a**) Urea denaturation. (**b**) GndHCl denaturation. Fits were done using Kaleidagraph. Average

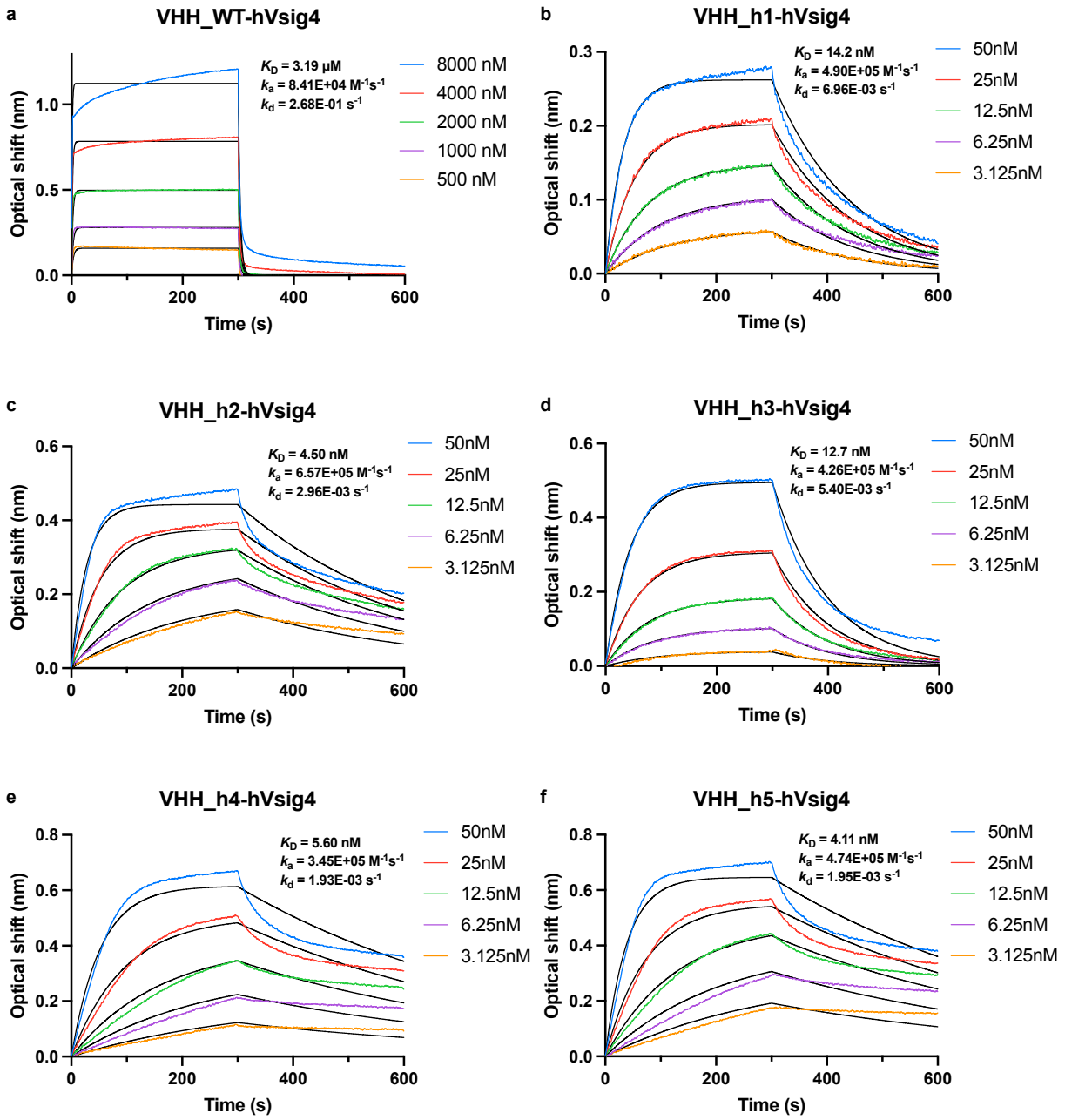
829 ΔG is listed in Supplementary Table.3.



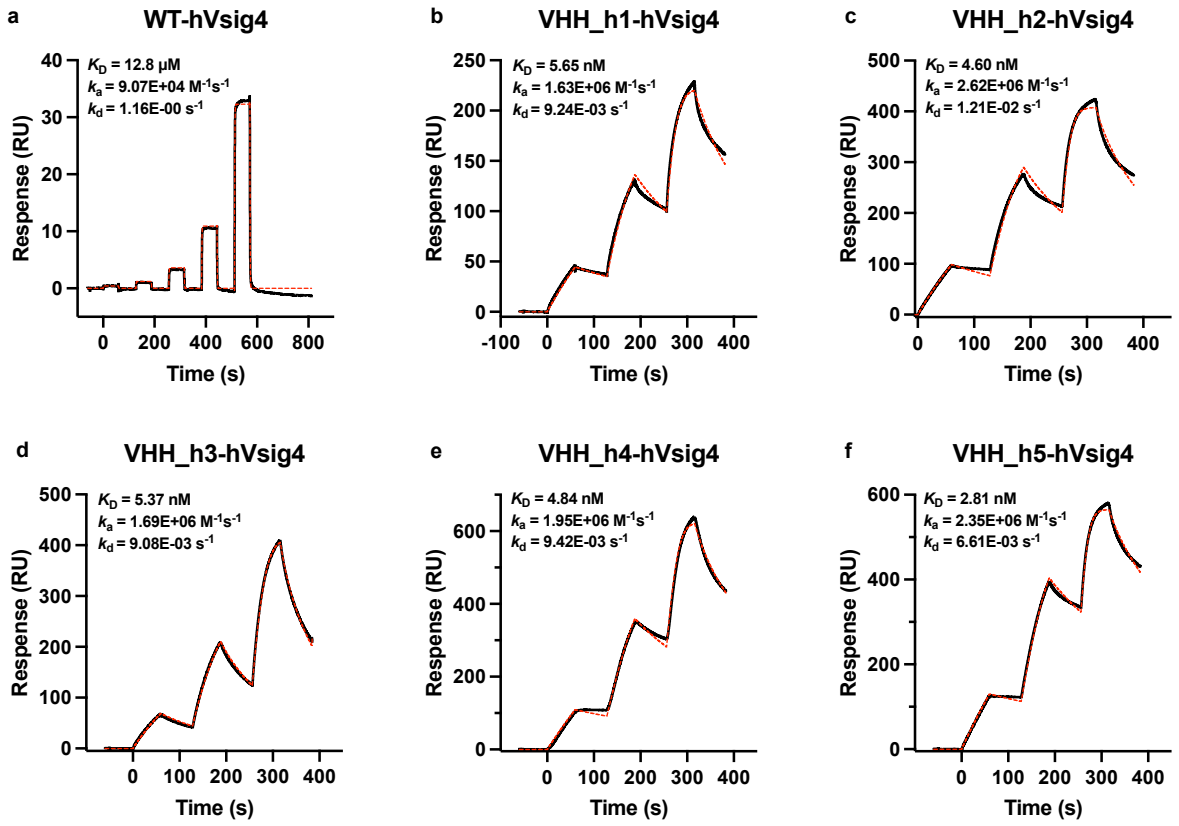
Supplementary Fig.5. T_m determination based on Barycentric Mean calculation from Intrinsic Tryptophan Fluorescence (ITF) during a temperature ramp for VHH_WT, VHH_m1-VHH_m4.



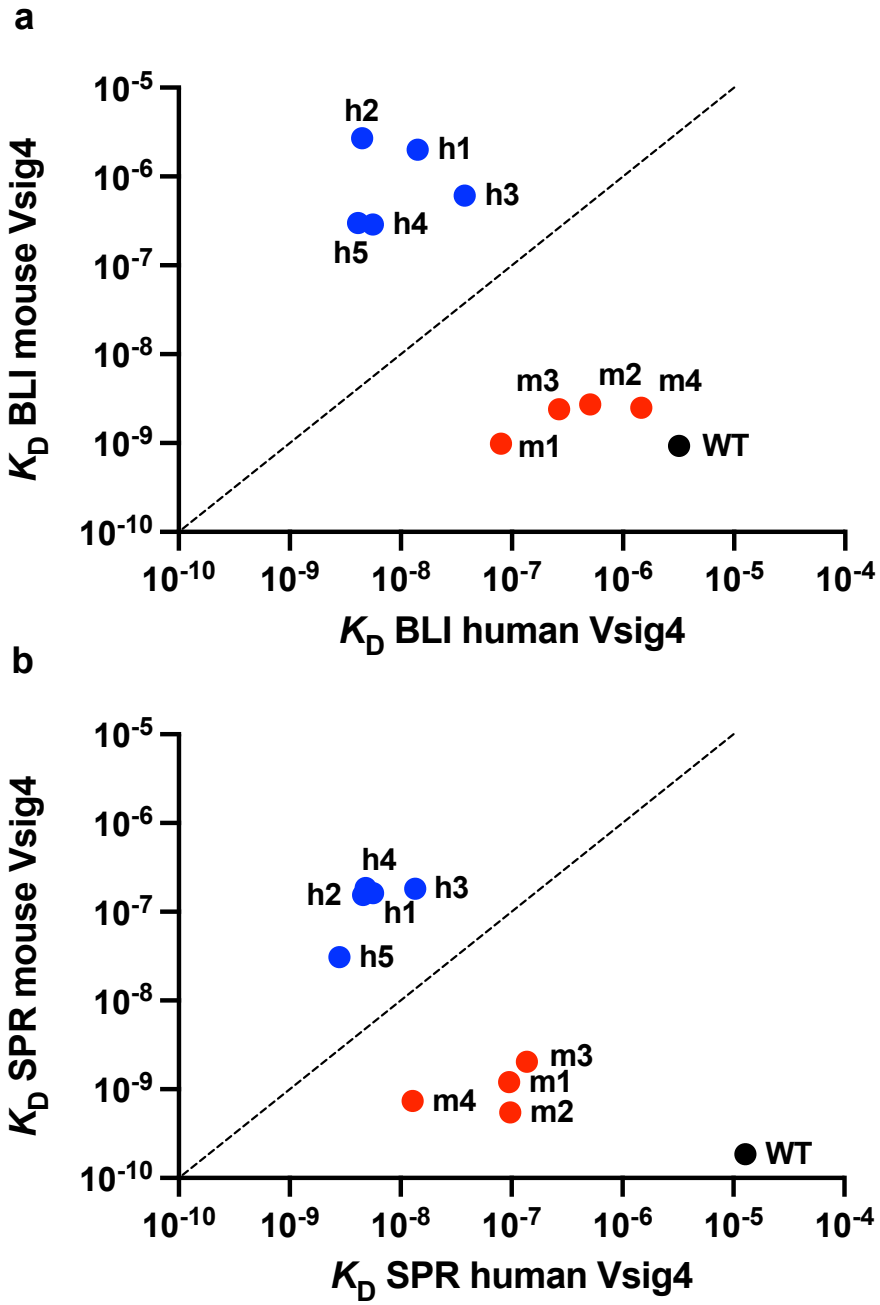
Supplementary Fig.6. T_{agg} determination based on Scatting at 266 nm from Right Angle Light Scattering (RALS) during a temperature ramp for VHH_WT, VHH_m1-VHH_m4.



Supplementary Fig.7. The binding curves (colored) of VHH_WT, VHH_h1-VHH_h5 against hVsig4 (BLI) with fitting curves (black).



Supplementary Fig.8. The binding curves (black) of VHH_WT, VHH_h1-VHH_h5 against hVsig4 (SPR) with fitting curves (red).

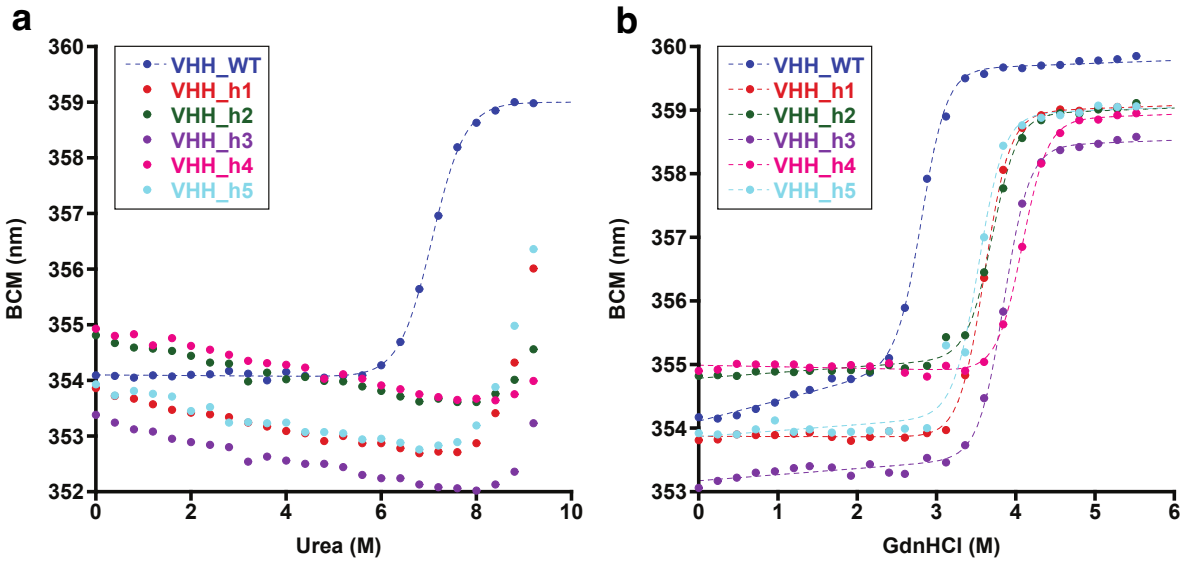


842

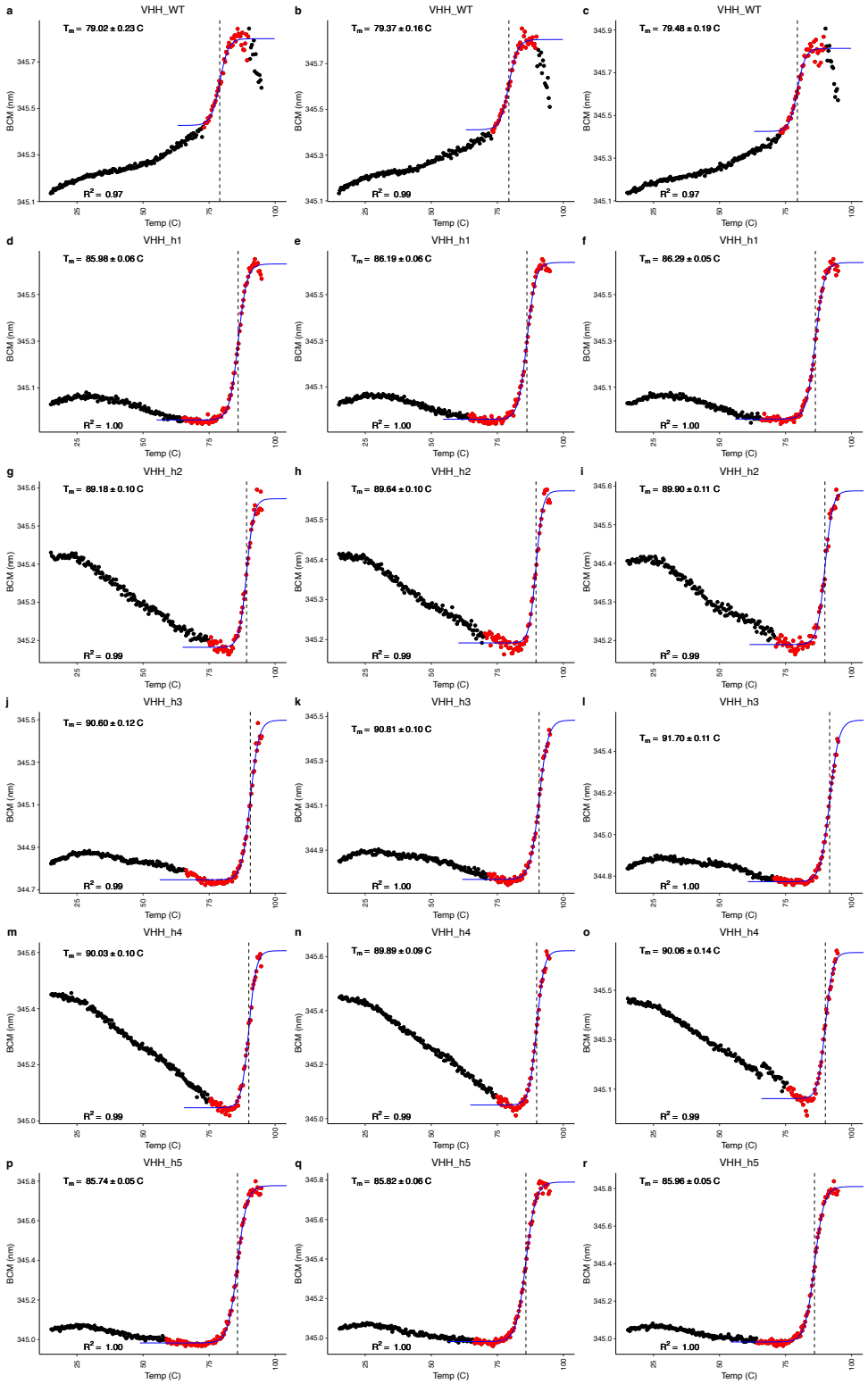
843 **Supplementary Fig.9. Binding of VHH_WT, VHH_m and VHH_h hits to mVsig4 and**

844 **hVsig4.** (**a**) Affinity constants measured by BLI of all Vsig4 hits to both mVsig4 and hVsig4. (**b**)

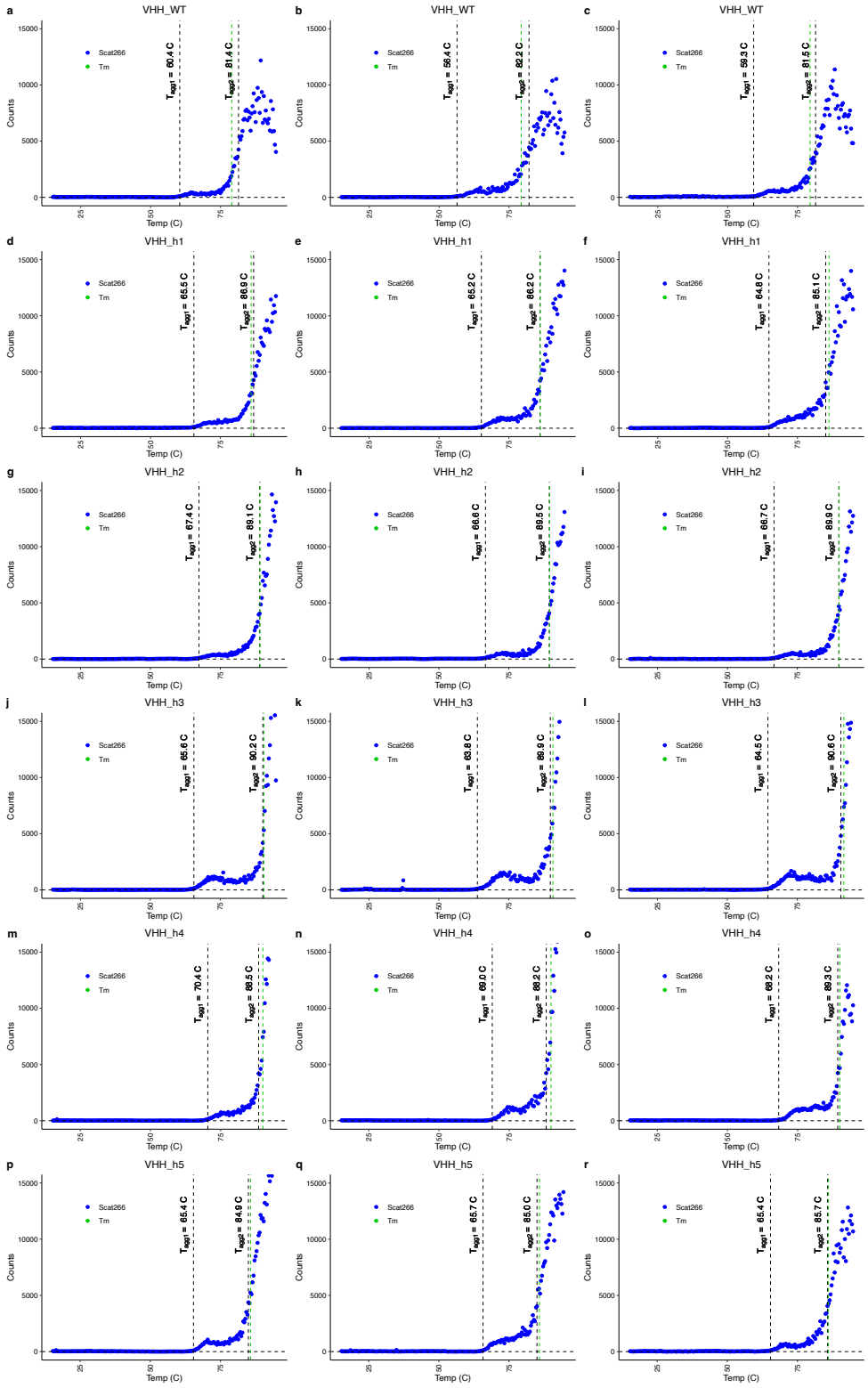
845 Affinity constants measured by SPR of all Vsig4 hits to both mVsig4 and hVsig4.



Supplementary Fig.10. Representative chemical denaturation curves of VHH_WT, VHH_h1-h5. (**a**) Urea denaturation. (**b**) GndHCl denaturation. Fits were done using Kaleidagraph. Average ΔG is listed in **Supplementary Table.6**.



Supplementary Fig.11. T_m determination based on Barycentric Mean calculation from Intrinsic Tryptophan Fluorescence (ITF) during a temperature ramp for VHH_WT, VHH_h1-VHH_h5.



853

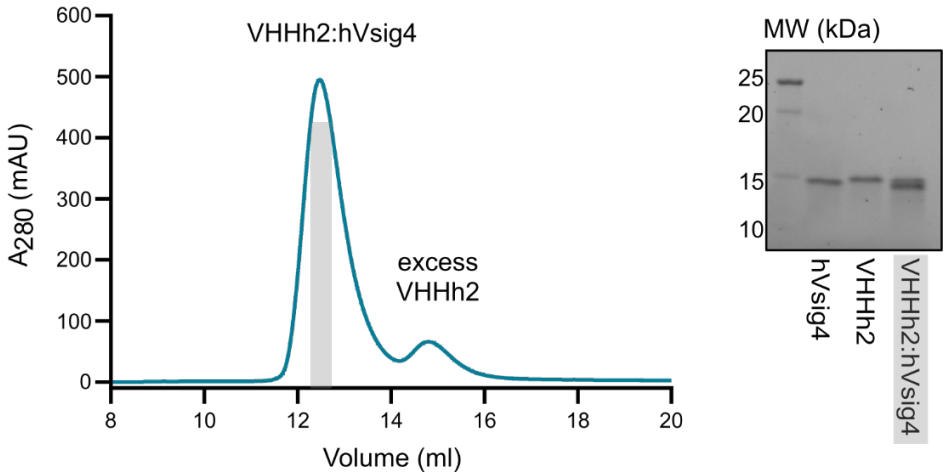
854 **Supplementary Fig.12.** T_{agg} determination based on Scatting at 266 nm from Right Angle Light

855 Scattering (RALS) during a temperature ramp for VHH_WT, VHH_h1-VHH_h5.

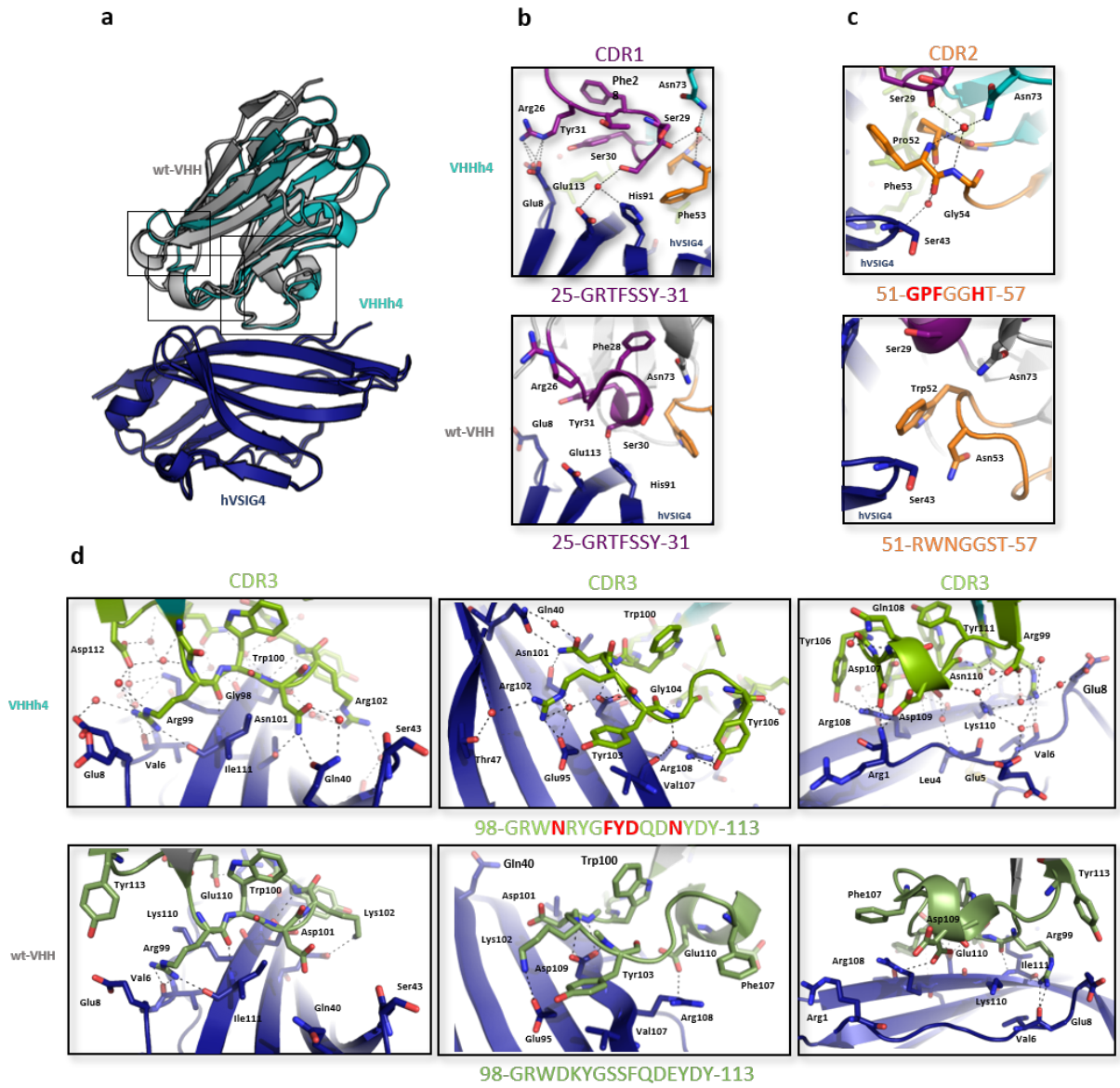
a)

Protein	MW_{calc} (kDa) Protein
hVsig4 (20-137)	14.46
VHHh1	14.85
VHHh2	14.76
VHHh3	14.77
VHHh4	14.80
VHHh5	14.70

b)

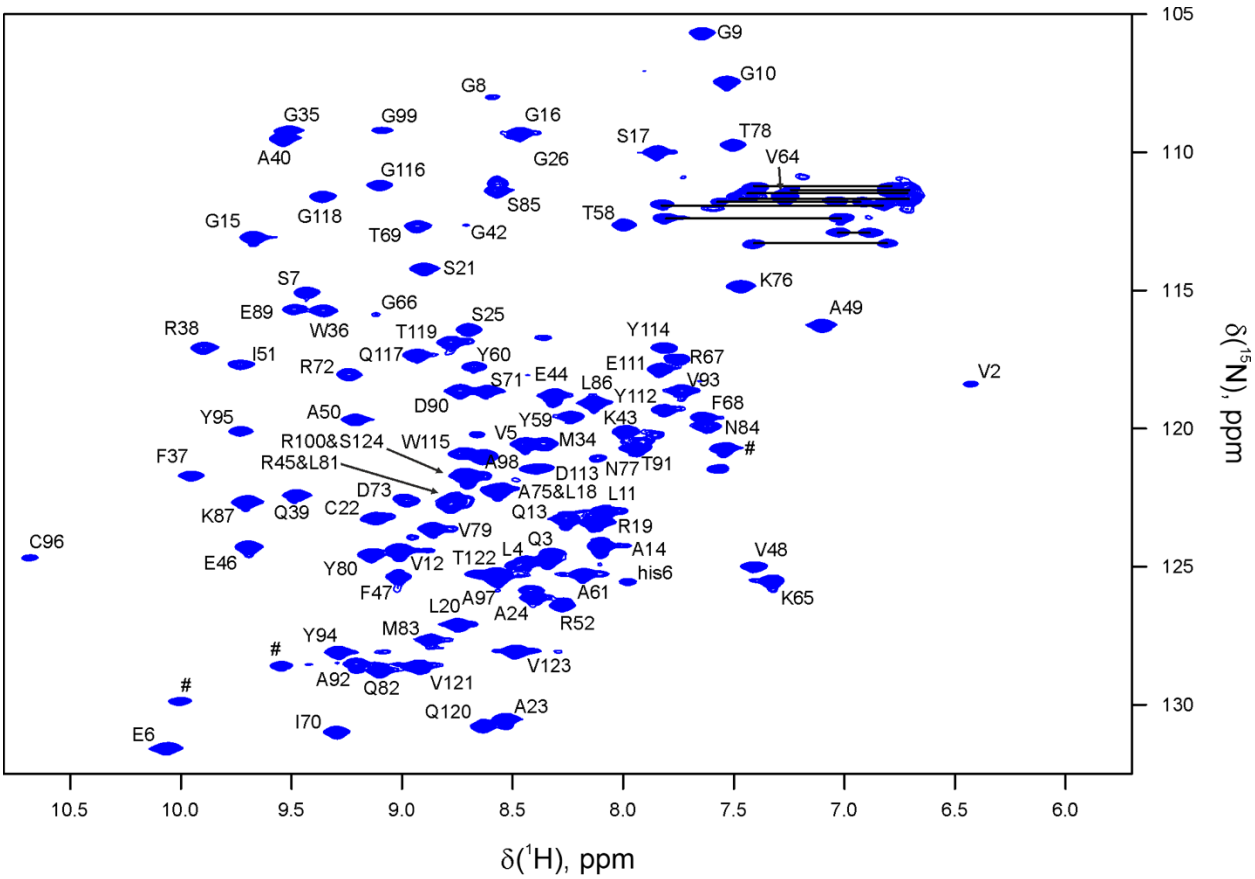


Supplementary Fig.13. Representative SEC elution profiles for VHH_h2:hVsig4 complex reconstitution. (**a**) Overview of molecular masses of all the VHHs and hVsig4. (**b**) Representative SEC elution profile for the purification of the VHH_h2:hVsig complex using excess of VHH_h2 on a Superdex 75 Increase 10/300 GL column and representative SDS-PAGE.

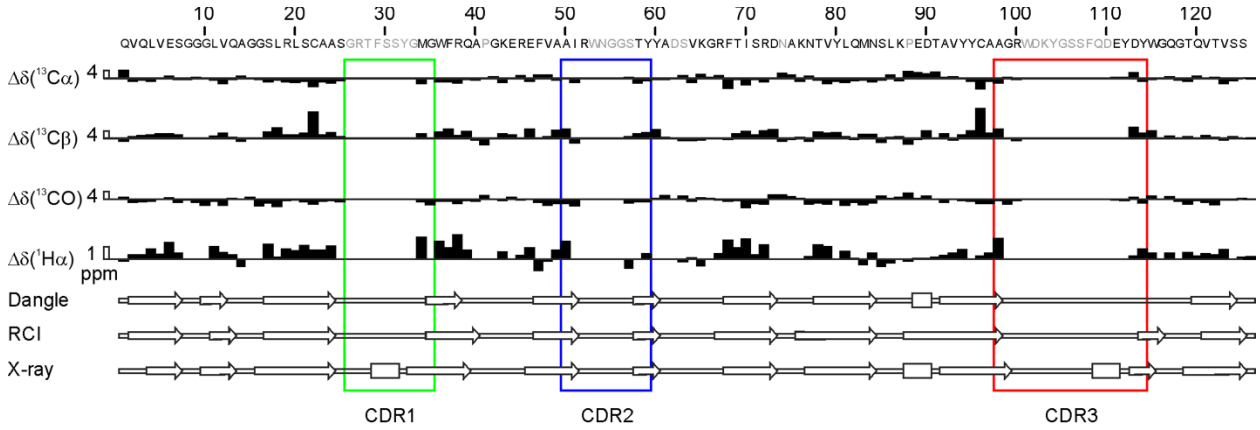


Supplementary Fig.14. Structure of the hVsig4:VHH_h4 complex compared to the hVsig4:VHH_WT complex. (**a**) Cartoon representation of the determined X-ray structure for VHH_h4 fragment in complex with hVsig4 superimposed to hVsig4:VHH_WT complex (PDB 5IML) based on the structural alignment of the two hVsig4 structures. VHH_h4 is colored in teal, VHH_WT in grey and hVsig4 in dark blue. (**b**) Details of the hVsig4:VHH_h4 (up) and hVsig:VHH_WT (down) interface focusing on CDR1 loop. CDR1 loops of VHHh4 and VHH_WT are coloured in deep purple. (**c**) Details of the hVsig4:VHH_h4 (up) and hVsig:VHH_WT (down)

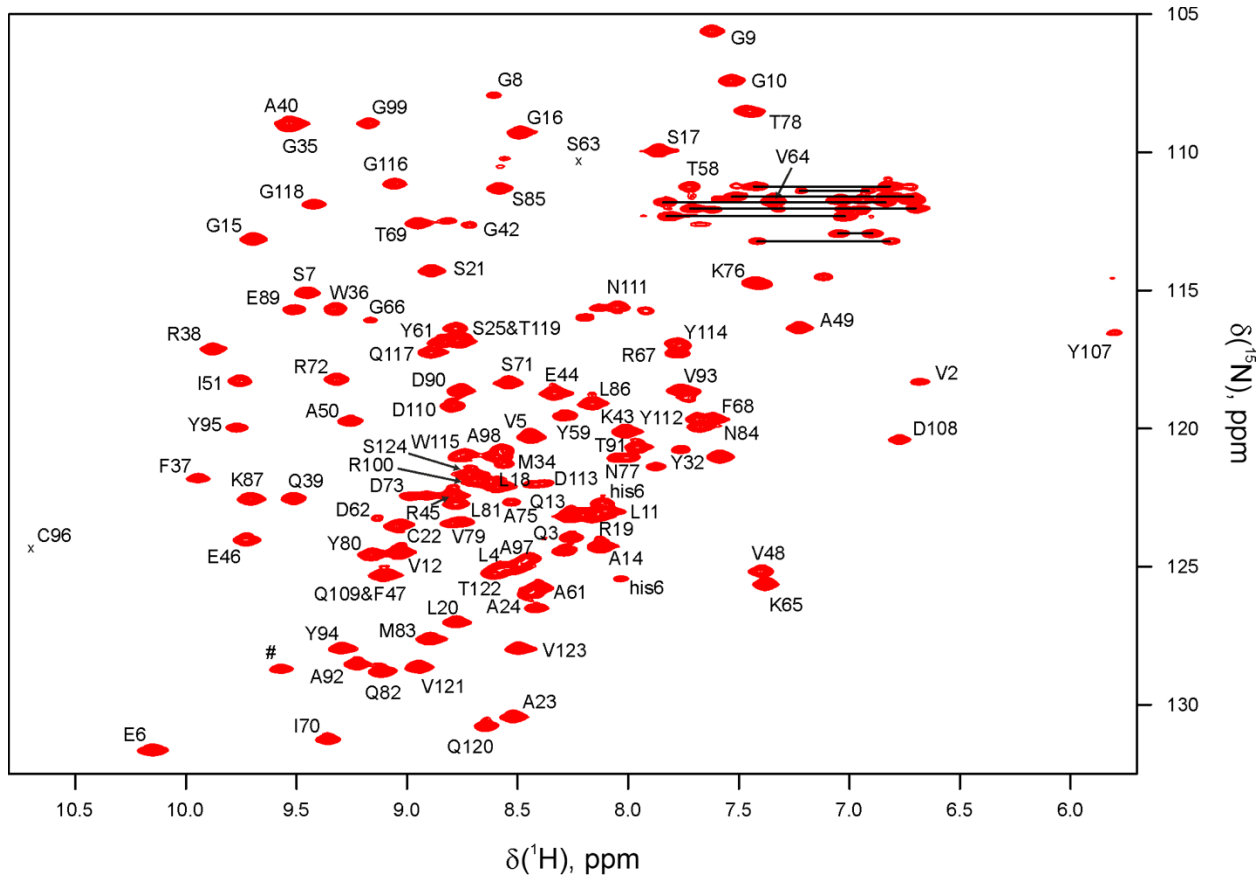
869 interface focusing on CDR2 loop. CDR2 loops of VHH_h4 and VHH_WT are coloured in orange.
870 (**d**) Details of the hVsig4:VHH_h4 (up) and hVsig:VHH_WT (down) interface focusing on CDR3
871 loop with 3 different perspectives. Selected interface residues are labelled and shown as sticks.
872 CDR3 loops of VHH_h4 and VHH_WT are coloured in green (split pea). Dashed lines represent
873 salt bridges and hydrogen bonds. Red spheres represent water molecules which are part of the
874 interface.



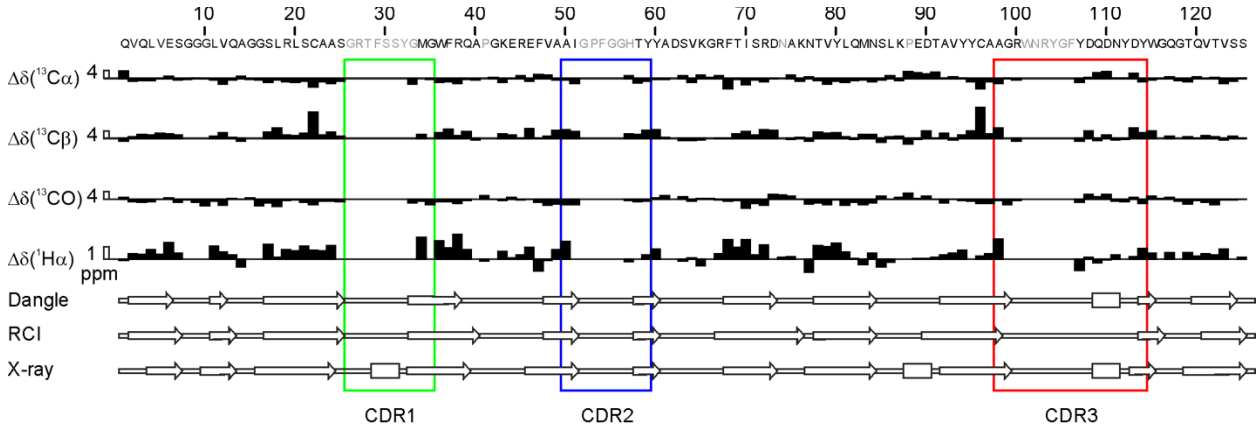
Supplementary Fig.15. [^{1}H,^{15}N] HSQC spectrum of VHH_WT. The labels show assigned backbone amide resonances, tryptophan $N_{\varepsilon1}H_{\varepsilon1}$ indoles (#), a His-tag peak (his6), and sidechain NH_2 resonances of Asn and Gln residues (horizontal lines). The experiments were performed in 20 mM sodium phosphate 50 mM NaCl pH 6.5 at 313 K.



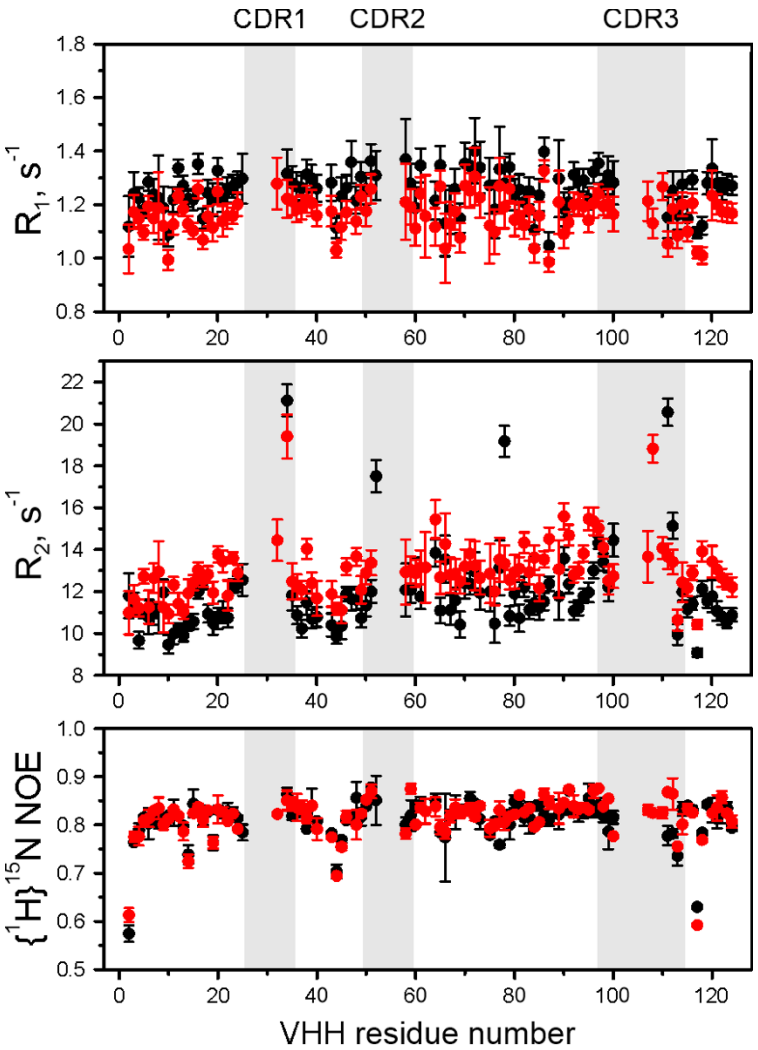
Supplementary Fig.16. Extent of the NMR backbone assignment and secondary structure prediction of VHH_WT. Threshold deviations from random-coil ^{13}Cα, ^{13}Cβ, ^{13}CO, and ^{1}Hα chemical shifts (filled bars), with the leading open bars for scale (values in ppm). The secondary structure (rectangles – α helices, arrows – β strands) predicted by DANGLE, RCI, or seen in the X-ray structure of the VHH_WT bound to mouse Vsgi4 (5immBA). The residues that exhibited no backbone amide resonances in the HSQC spectrum are in grey.



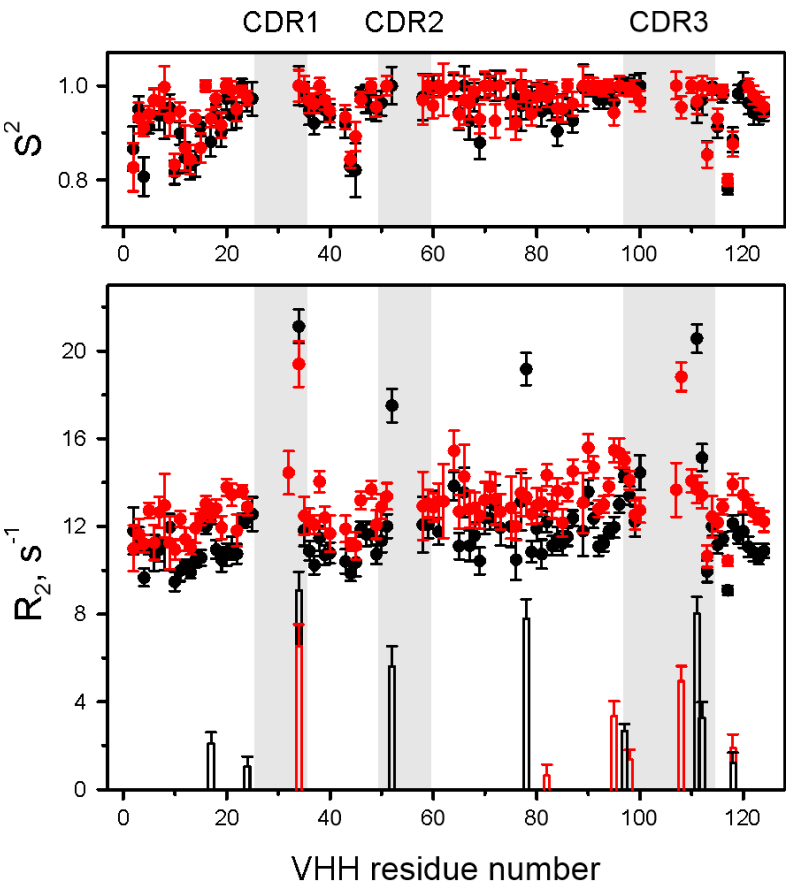
Supplementary Fig.17. [^{1}H,^{15}N] HSQC spectrum of VHH_h4. The labels show assigned backbone amide resonances, tryptophan $N_{\varepsilon1}H_{\varepsilon1}$ indole (#), His-tag peaks (his6), and sidechain NH_2 resonances of Asn and Gln residues (horizontal lines). Positions of S63 and C96 backbone amide signals (not seen at this contour level) are indicated by crosses. The experiments were performed in 20 mM sodium phosphate 50 mM NaCl pH 6.5 at 313 K.



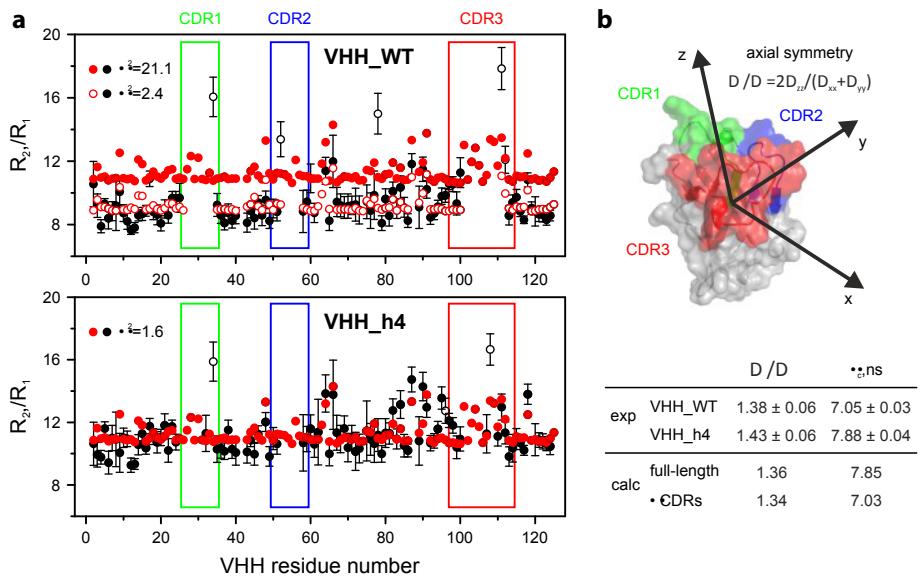
Supplementary Fig.18. Extent of the NMR backbone assignment and secondary structure prediction of VHH_h4. Threshold deviations from random-coil ^{13}Cα, ^{13}Cβ, ^{13}CO, and ^{1}Hα chemical shifts (filled bars), with the leading open bars for scale (values in ppm). The secondary structure (rectangles – α helices, arrows – β strands) predicted by DANGLE, RCI, or seen in the X-ray structure of the VHH_WT bound to mouse Vsgi4 (5immBA). The residues that exhibited no backbone amide resonances in the HSQC spectrum are in grey.



Supplementary Fig.19. ^{15}N NMR relaxation parameters of VHHs. Spin-lattice (R_1) and spin-spin (R_2) relaxation rates, and heteronuclear steady-state nuclear Overhauser effect ({^{1}H}-^{15}N NOE) of the backbone amide atoms of VHH_WT (black) and VHH_h4 (red). The grey areas indicate CDR regions. Experiments were performed in 20 mM NaPi 50 mM NaCl pH 6.5 at 313 K.

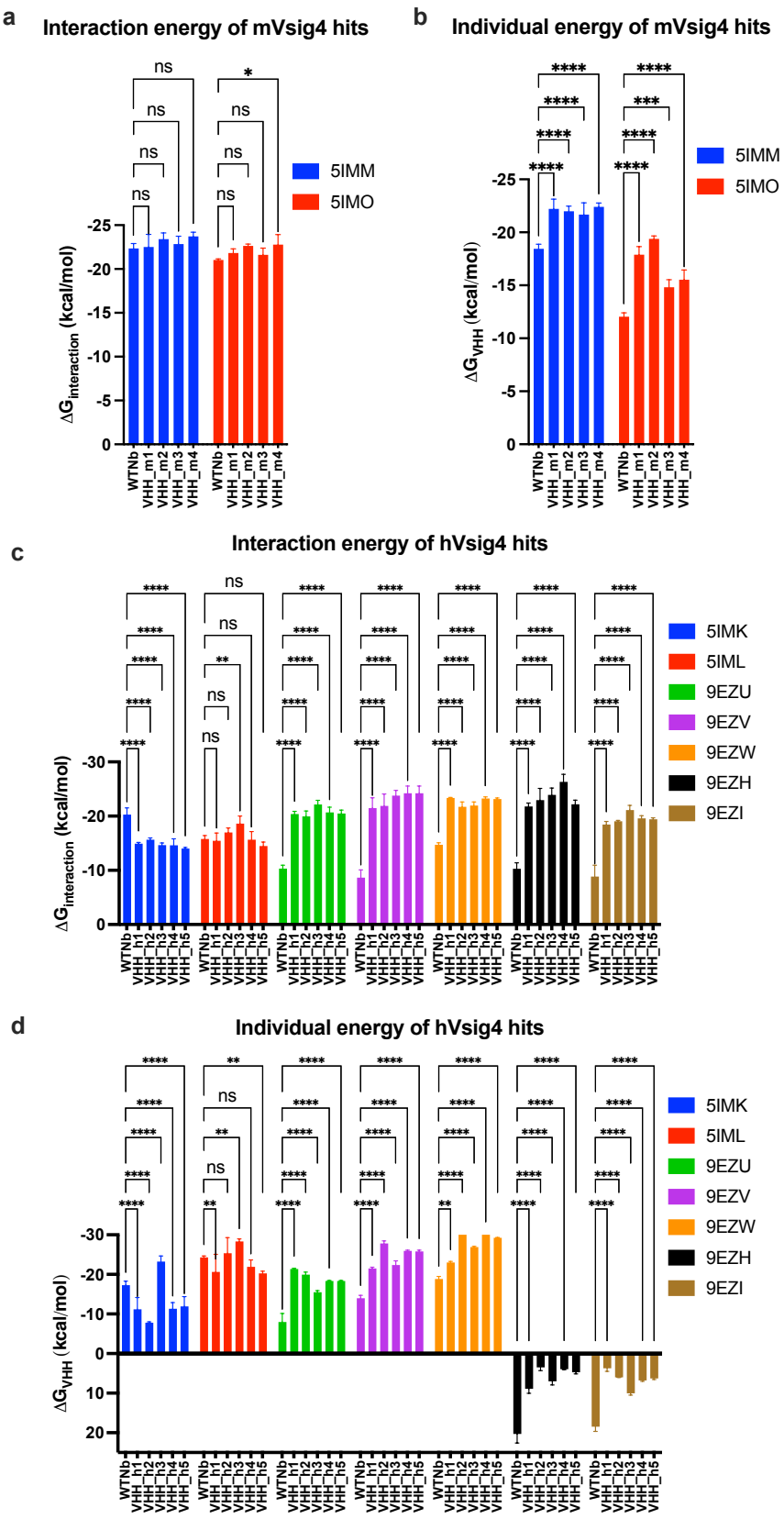


Supplementary Fig.20. Backbone dynamics of VHHs. (top) S^2 order parameters for the N-H vectors obtained from the model-free analysis of the NMR relaxation data in Supplementary Fig.19. (bottom) Measured R_2 relaxation rates (circles) and $R_{2,ex}$ exchange contributions obtained from the model-free analysis (bars). The data for the VHH_WT and VHH_h4 are in black and red, respectively. The grey areas indicate CDR regions. S2 metric indicates very limited ps-ns dynamics, with the smallest values for the N-terminus and loops centered on residues 10, 44, and 117 (top). Interestingly, the few VHH_WT CDR residues that do exhibit HSQC peaks show considerable exchange contributions to transverse relaxation rates (R2,ex), suggesting that they undergo a conformational exchange on the µs-ms timescale (bottom). The R2,ex contributions are reduced in the VHH_h4 (bottom), suggesting a much less extensive conformational exchange of the mutant CDRs, which explains its more favorable spectral properties.



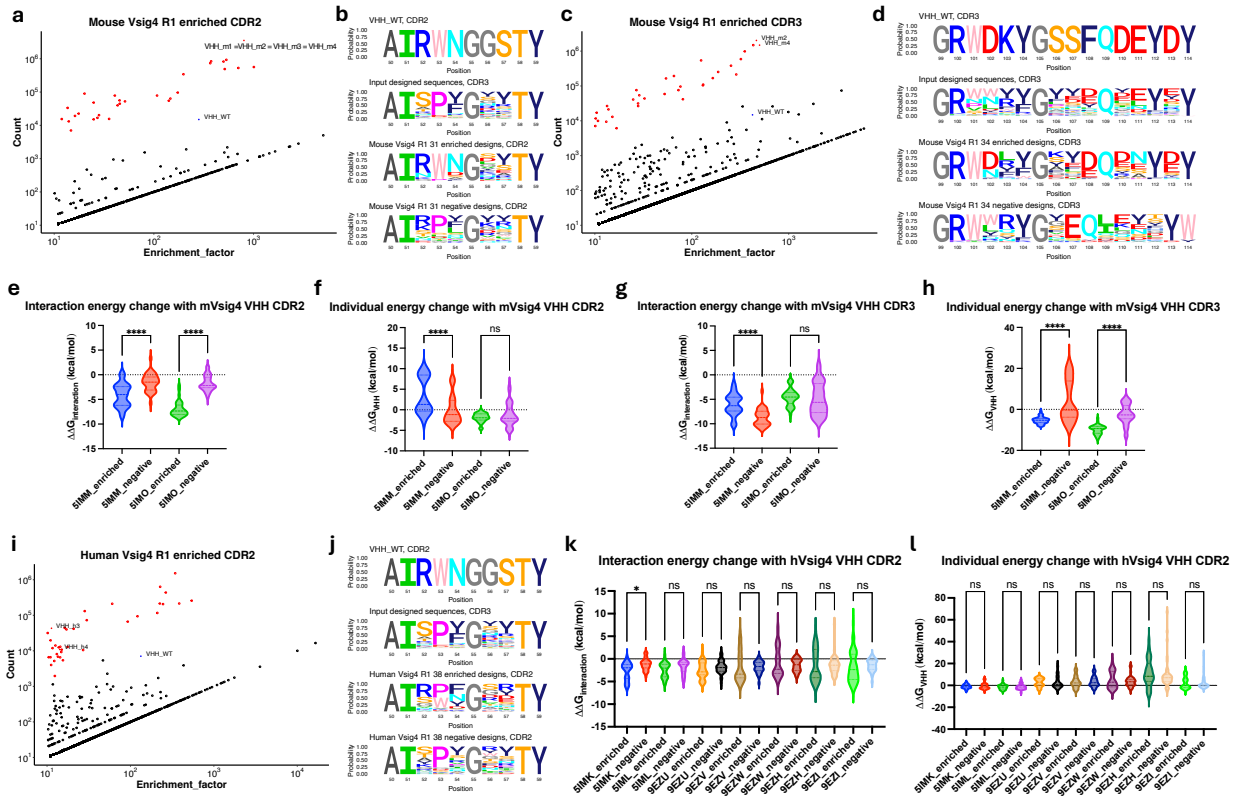
Supplementary Fig.21. Hydrodynamic properties of VHHs. (**a**) Measured (black) and predicted (red) ratios of ^{15}N R_1 and R_2 NMR relaxation rates for the VHH_WT (top) and VHH_h4 (bottom). R_2/R_1 values were calculated from the X-ray structure of the full-length VHH_WT or with the CDRs removed (filled and open red circles, respectively). VHH residues with R_2 exchange contributions (black open circles) were excluded from the χ^2 calculations. (**b**) Orientation of the diffusion tensor (coordinate system) in the VHH molecular frame. The structure of the VHH_WT is taken from the X-ray structure of its complex with the mouse Vsgi4 (5immBA), with CDRs indicated. The relaxation properties of both VHHs are well accounted for by an axially symmetric diffusion model, with the diffusion tensor parameters given in (b). Calculated from the X-ray structure of VHH bound to mouse Vsgi4, the hydrodynamic parameters of VHH_h4 are in excellent agreement with those derived from experiment: closely matching values of rotational correlation time (τc) and high similarity of experimental and calculated R2/R1 profiles ($\chi 2$ = 1.6) are particularly noteworthy (a, bottom). At the same time, the measured R2/R1 and experimentally derived τc values for the VHH_WT are considerably lower than expected (a, top), indicating a faster molecular tumbling compared to VHH_h4. To simulate a molecular system where

934 conformationally exchanging CDR loops do not contribute to the overall protein tumbling, we
935 removed the VHH_WT CDR residues that show no HSQC peaks from the X-ray structure and re-
936 calculated the hydrodynamic properties. Both R2/R¬ and τc parameters of this reduced set
937 (ΔCDRs) are in excellent agreement with the experimental VHH_WT values (a, top). Overall, it
938 appears that free VHH_h4 in solution maintains the bound-like protein conformation, with closely
939 packed CDR loops undergoing limited motion, while the CDRs of VHH_WT are largely
940 disordered and involved in extensive conformational exchange.


a
Interaction energy of mVsig4 hits
b
Individual energy of mVsig4 hits
5IMM
5IMO
ΔG interaction (kcal/mol)
ΔG VHH (kcal/mol)
WTNb
VHH_m1
VHH_m2
VHH_m3
VHH_m4
c
Interaction energy of hVsig4 hits
d
Individual energy of hVsig4 hits
5IMK
5IML
9EZU
9EZV
9EZW
9EZH
9EZI
VHH_h1
VHH_h2
VHH_h3
VHH_h4
VHH_h5

941

942 **Supplementary Fig.22. FoldX modeling of VHH hit sequences reveals energy differences between**
943 **VHH_WT and VHH candidates.** (**a**) Comparison of the interaction energy ($\Delta G_{interaction}$) between VHH_WT and
944 mVsig4 hits on 5IMM and 5IMO models. (**b**) Comparison of the VHH individual energy (ΔG_{VHH}) between
945 VHH_WT and mVsig4 hits on 5IMM and 5IMO models. (**c**) Comparison of the interaction energy ($\Delta G_{interaction}$)
946 between VHH_WT and hVsig4 hits on 5IMK, 5IML, 9EZU, 9EZV, 9EZW, 9EZH and 9EZI models. (**d**) Comparison
947 of the VHH individual energy (ΔG_{VHH}) between VHH_WT and hVsig4 hits on 5IMK, 5IML, 9EZU, 9EZV, 9EZW,
948 9EZH and 9EZI models. Triplicate modelling was performed and two-way ANOVA was run between VHH_WT and
949 VHH candidates in all four plots (ns $p>0.123$, *$p<0.033$, **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$).



Supplementary Fig.23. FoldX modeling of NGS-identified CDR designs reveals energy change differences between enriched and negative CDR designs. (**a**) Count and enrichment factor of mVsig4 enriched CDR2 sequences (count > 10, enrichment factor > 10, enrichment factor is the ratio of the count in target selection to the count in empty control). Red dots are the sequences designed by EvolveX, red solid dots are the sequences from identified hits, blue solid dot is VHH_WT, and black dots are non-designs or mutants generated during the PCR or phage selection stages. The same annotation is omitted in the (c) and (i) below. (**b**) Sequence logo plots of VHH_WT CDR2, 31 mVsig4 enriched CDR2 designed sequences and 31 negative CDR2 designed sequences. AbM CDR definition was used. The same annotation is omitted in the (d) and (j) below. (**c**) Count and enrichment factor of mVsig4 enriched CDR3 sequences. (**d**) Sequence logo plots of VHH_WT CDR3, 34 mVsig4 enriched CDR3 designed sequences and 34 negative CDR3 designed sequences. (**e**) Comparison of the interaction energy change ($\Delta\Delta G_{interaction}$) between enriched and negative mVsig4 CDR2 designs. (**f**) Comparison of the VHH individual energy change ($\Delta\Delta G_{VHH}$) between enriched and negative mVsig4 CDR2 designs. (**g**) Comparison of the interaction energy change ($\Delta\Delta G_{interaction}$) between enriched and negative mVsig4 CDR3 designs. (**h**) Comparison of the VHH individual

964 energy change ($\Delta\Delta G_{VHH}$) between enriched and negative mVsig4 CDR3 designs. (**I**) Count and enrichment factor of

965 hVsig4 enriched CDR2 sequences. (**j**) Sequence logo plots of VHH_WT CDR2, 38 hVsig4 enriched CDR2 designed

966 sequences and 38 negative CDR2 designed sequences. (**k**) Comparison of the interaction energy change

967 ($\Delta\Delta G_{interaction}$) between enriched and negative hVsig4 CDR2 designs. (**l**) Comparison of the VHH individual energy

968 change ($\Delta\Delta G_{VHH}$) between enriched and negative hVsig4 CDR2 designs. One way ANOVA was used in (e-h and k-

969 l) (ns $p>0.123$, *$p<0.033$, **$p<0.002$, ***$p<0.0002$, ****$p<0.00001$).

Supplementary Tables:

Supplementary Table.1. 36 unique sequences of mVsig4 hits with average of AlphaScreen signal/noise (S/N) ratio.

Sequence alignment	Average of AS S/N
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	21.57
.....S.....RH.AYD.Y.....	46.17
.....NY.....RH.AYD.Y.....	22.33
.....VE.....RH.AYD.Y.....	18.59
.....VE.....LF..FD..N.....	16.27
.....VE.....LF.KYD..N.....	16.21
.....VE.....I.....LF..YD..N.E.....	13.93
.....VE.....LF..YD..N.....	12.06
.....VE.....LF..YD..N.E.....	12.02
.....VE.....LF..FD..N.E.....	11.35
.....RY.....LF..FD..N.E.....	11.01
.....R.....VE.....LF..FD..N.E.....	10.70
.....V.....LF.KYD..N.E.....	10.39
.....VE.....LF.KFD..N.E.....	8.78
.....L.....VE.....LF..FD..N.E.....	8.49
.....VE.....T.....LF..YD..N.E.....	8.41
.....VE.....LF.KYD..N.E.....	7.27
.....T.....	5.78
.....DY.....LF..YD..N.E.....	2.13
.....DE.....LF..YD..N.E.....	1.94
.....L...VE.....LF..FD..NFE.....	1.42
.....DE.....LF..FD..N.E.....	1.37
.....SEY.YA.....VWN..FGE.W..S.....	1.34
.....K...YY.....F..AFE.ES.....	1.31
.....SA.....LF..FD..N.E.....	1.30
.....G.D.EY.....RH.AYD.Y.....	0.90
.....G.D.EY.....L.....RH.AYD.Y.....	0.72
.....SPD.NL.....SLRF.AAL.E.LE.....	0.65
.....SEY.YG.....VWN..FGE.W..S.....	0.59
.....SEY.YG.....RYF.NFD.N..E.....	0.54
.....SEY.YA.....PVNF..YYE.YD.N.....	0.52
.....SPY.Y.....AEY..YA..YDFE.....	0.49
.....DPY.DV.....NI..NYD.IL.E.....	0.41
.....A.DY.....FF.YDE.ND.K.....	0.40
.....SPS.SR.....F..AFE.ES.....	0.35
.....SPY.DY.....YR..WE.I.E.....	0.35
.....HPY.SL.....VWE..Y.S.E.FE.....	0.35

Supplementary Table.2. 12 unique sequences of mVsig4 hits selected for the BLI off-rate test.

Sequence alignment	k_d (s^{-1})
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	6.38E-03
.....VE.....LF..YD..N.....	4.22E-03
.....VE.....LF..YD..N.E.....	4.77E-03
.....VE.....LF..FD..N.....	4.80E-03
.....VE.....LF..FD..N.E.....	5.13E-03
.....VE.....LF.KYD..N.....	6.13E-03
.....VE.....LF.KYD..N.E.....	6.70E-03
.....V.....LF.KYD..N.E.....	6.75E-03
.....RY.....LF..FD..N.E.....	7.32E-03
.....VE.....LF.KFD..N.E.....	7.89E-03
.....NY.....RH.AYD.Y.....	8.01E-03
.....S.....RH.AYD.Y.....	8.26E-03
.....VE.....RH.AYD.Y.....	8.77E-03

Supplementary Table.3. Biophysical characterization data of VHH_m1-VHH_m4.

Name	CDR2	CDR3	K_D BLI (nM) (mVsig4)	K_D SPR (nM) (mVsig4)	T_m (°C)	T_{agg} (°C)	ΔG_{Urea} (kcal/mol)	ΔG_{GdnHCl} (kcal/mol)
VHH_WT	AI**RWN**G**GS**TY	GR**WDKY**G**SSFQDE**Y**D**Y	0.9	0.2	80.0	80.7	-10.2	-10.7
VHH_m1	.....VE..	....LF..YD..N.D.	1.0	1.2	84.6	86.2	☆	-12.3
VHH_m2	.....VE..	....LF..YD..N.E.	2.7	0.5	85.9	86.5	☆	-12.5
VHH_m3	.....VE..	....LF..FD..N.D.	2.4	2.0	84.0	86.0	☆	-12.3
VHH_m4	.....VE..	....LF..FD..N.E.	2.5	0.7	86.6	89.5	☆	-12.4

(Note: CDR definition: AbM, ☆ means it was not fully denatured under the most extreme conditions)

Supplementary Table.4. 112 unique sequences of hVsig4 hits with average of ELISA signal/noise (S/N) ratio.

Sequence alignment	Average of ELISA S/N
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	4.33
.....GPF..H.....NR..FYD..N.....	37.30
.....GPF.....NR..FED..NF.....	35.92
.....GPF..I.....NR..FED..NF.....	35.66
.....PF..D.....NR..FED..NF.....	34.98
.....PF.NY.....NR..FED..NF.....	34.87
.....GPF.SL.....NR..FED..NF.....	34.18
.....SPF..T.....NR..FED..NF.....	33.44
.....PF.AN.....NR..FYD..N.....	32.50
.....GPF.NG.....A.....NR..FED..NF.....	32.00
.....GPF.DK.....NR..FED..NF.....	31.56
.....SPF..N...S.....NR..FED..NF.....	31.22
.....PF..D.....NR..FYD..N.....	31.22
.....GPF.....NR..FEDR.NF.....	31.00
.....APF.YT.....NR..FED..NF.....	29.28
.....SPF..N.....NR..FED..NF.....	29.06
.....GPF..T.....T....NR..FYE..N.....	28.69
.....GPF.SR.....NR..FED..NF.....	27.89
.....GPF..Y.....NR..FYD..N.....	27.56
.....GPF.R.....I.....NR..FED..NF.....	27.41
.....R..G.....NR..FED..NF.....	27.06
.....GPF..T.....NR..EQD.YN.E.....	26.50
.....SPF..A.....NR..FED..NF.....	26.34
.....GPF.DR.....NY..QYE.TS.....	25.48
.....NPF..H.....NR..FED..NF.....	25.41
.....GPF.RG.....NR..FED..NF.....	25.19
.....NPF..Y.....NR..FYD..N.....	24.24
.....SPF..Q.....NR..FED..NF.....	23.75
.....SPF.EN.....NR..FYD..N.....	23.33
.....APF.LT.....NR..FED..NF.....	22.22
.....GPF.DR.....NY..QFD.YS.E.....	21.97
.....VPF..D.....NR..FYD..N.....	21.70
.....PF.WD.....NR..FED..NF.....	21.37
.....FS.AY.....NR..FED..NF.....	21.03
.....GPF.YE.....NR..FED..NF.....	20.70
.....APF.DR.....RY..EDD.EN.E.....	20.52
.....TPF.EW.....NR..FED..NF.....	20.15
.....TPF.SL.....NR..FED..NF.....	19.19
.....GPY.SY.....NR..FED..NF.....	19.07
.....SPF..R.....NR..FED..NF.....	18.94
.....SPY..T.....NR..FED..NF.....	18.51
.....APF.YD.....NR..FED..NF.....	18.41
.....APF.HK.....NR..EQD.YN.E.....	18.13
.....TPF.D.....NR..FED..NF.....	17.63
.....SPF.RY.....NR..FED..NF.....	17.38
.....VE.....NY..FTD..NFE.....	16.93
.....APY.SY.....NR..FED..NF.....	16.91
.....VE.....LF..YD..N.E.....	16.51
.....S.R.AR.....NR..FED..NF.....	15.82
.....SPF..W.....NR..FED..NF.....	15.63
.....GPF.SY.....NF..FED..NF.....	15.15
.....SPF.SH.....NR..FED..NF.....	14.84
.....SPF.SH.....NR..FYD..N.....	14.81
.....APF.L.....NR..FED..NF.....	14.74
.....SPF..R.....NR..FYD..N.....	14.66

Sequence alignment	Average of ELISA S/N
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	4.33
.....SPY.SW.....NR..FED..NF.....	13.78
.....VPY.RY.....NR..FED..NF.....	13.74
.....SPF.SL.....NR..FED..NF.....	13.70
.....SPF.YK.....NR..FED..NF.....	13.63
.....SSF.RR.....NR..FED..NF.....	12.56
.....TPY..Y.....NR..FED..NF.....	12.52
.....GGY.KY.....M.....NR..FYD..N.....	12.48
.....SPF.SY.....LRF.YYD.EN.E.....	12.30
.....APY.DY.....NR..FED..NF.....	12.30
.....N...K.....NR..FYD..N.....	12.28
.....APF.R.....LRF.YYD.EN.E.....	11.96
.....SPF..R.....NY..FTD..NFE.....	11.93
.....VE.....LF..FD..N.E.....	11.92
.....SPF..L.....NR..FYD..N.....	11.88
.....TPF.VE.....YR..FED..NF.....	11.82
.....SPY.YY.....L.....NR..FED..NF.....	11.67
.....NPY.RI.....NR..FYD..N.....	11.65
.....SPF.H.....NR..FED..NF.....	11.38
.....PF.WD.....NR..FYE..N.....	11.33
.....GPY..T.....NR..PTD.EN.....	11.30
.....TPF.IG.....NR..FED..NF.....	11.22
.....SYD.SY.....RY..EDD.RN.E.....	10.44
.....VPY.RR.....NR..FED..NF.....	10.29
.....VPY.TR.....NR..FED..NFG.....	10.26
.....GPF..R.....NY..FTD..NFE.....	10.03
.....VE.....NL..KYE..S.....	9.96
.....APF.RH.....NR..FYD..N.....	9.93
.....S.NQ.....NR..FED..NFG.....	9.85
.....NPY.RR.....NR..FED..NF.....	9.63
.....SPF.YT.....NR..FED..NF.....	9.59
.....NPY.RR.....NR..FED..NF....R..R.....	9.44
.....SPF.VT.....NR..FED..NFG.....	9.30
.....GPF.RQ.....NR..FED..NFG.....	9.28
.....PF.....NY..FTD..NFE.....	9.11
.....APY.SF.....NR..FED..NF.....	8.52
.....W.....SPF..N.....NR..FED..NF.....	8.48
.....SPF..Y.....LRF.YYD.EN.E.....	8.37
.....APF.YE.....RY..EDD.EN.E.....	7.97
.....YPY.NR.....NR..FED..NF.....	7.44
.....SPF.YQ.....LRF.YYD.EN.E.....	7.15
.....NPF.QN.....NF..YYE.KNH.....	6.52
.....SPW.YA.....NR..FED..NFG.....	6.38
.....TPF.YQ.....NR..FYD..N.....	6.26
.....SPF.RQ.....NR..AND....S.....	6.22
.....PF.SE.....NYF.PYD..N.E.....	5.91
.....GPF.SF.....RY..EDD.RN.E.....	5.19
.....SPF.EN.....NY..FTD..NFE.....	4.11
.....GPF.S.....NF..YYE.KNH.....	4.04
.....NPF..Y.....RY..AND....E.....	3.67
.....SPF..V.....NNR..FYS....E.....	3.30
.....SPF.EL.....HYF.E.D.EN.E.....	2.96
.....QPF.H.....LN..R-E..Q.E.....	2.94
.....DY.....LF..YD..N.E.....	2.69
.....SPF.LE.....LF.KYD..N.E.....	2.22
.....SPW.YK.....NR..PTD.EN.E.....	1.82
.....SPF.LT.....NY..FTD..NFE.....	1.81
.....SPF..R.....LF..YD..N.E.....	1.00
.....YPY.NY.....NF..YED.NL.E.....	0.81

Supplementary Table.5. 56 unique sequences of hVsig4 hits selected for the BLI off-rate test.

Sequence alignment	k_d (s^{-1})
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	4.14E-02
.....SPF..N.....NR..FED..NF.....	7.48E-03
.....GPF..T.....NR..EQD.YN.E.....	8.59E-03
.....GPF.YE.....NR..FED..NF.....	1.03E-02
.....GPF.R.....I.....NR..FED..NF.....	1.19E-02
.....SPF..W.....NR..FED..NF.....	1.24E-02
.....GPF.RG.....NR..FED..NF.....	1.26E-02
.....GPF..T.....T....NR..FYE..N.....	1.30E-02
.....SPF..Q.....NR..FED..NF.....	1.30E-02
.....TPF.SL.....NR..FED..NF.....	1.31E-02
.....SPF.RY.....NR..FED..NF.....	1.36E-02
.....PF.NY.....NR..FED..NF.....	1.38E-02
.....NPF..H.....NR..FED..NF.....	1.38E-02
.....PF.WD.....NR..FED..NF.....	1.40E-02
.....SPF..N...S.....NR..FED..NF.....	1.43E-02
.....GPF..H.....NR..FYD..N.....	1.45E-02
.....APF.YE.....RY..EDD.EN.E.....	1.45E-02
.....APF.YD.....NR..FED..NF.....	1.46E-02

Sequence alignment	k_d (s^{-1})
QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYGMGWFRQAPGKEREFVAAIRWNGGSTYYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCAAGRWDKYGSSFQDEYDYWGQGTQVTVSS	4.14E-02
.....APF.LT.....NR..FED..NF.....	1.48E-02
.....APF.YT.....NR..FED..NF.....	1.49E-02
.....GPF.NG.....A.....NR..FED..NF.....	1.51E-02
.....PF..D.....NR..FYD..N.....	1.51E-02
.....GPF.....NR..FEDR.NF.....	1.51E-02
.....SPF..A.....NR..FED..NF.....	1.51E-02
.....GPF.SY.....NF..FED..NF.....	1.51E-02
.....SPF..T.....NR..FED..NF.....	1.53E-02
.....GPF.....NR..FED..NF.....	1.55E-02
.....R..G.....NR..FED..NF.....	1.56E-02
.....GPF.SR.....NR..FED..NF.....	1.59E-02
.....FS.AY.....NR..FED..NF.....	1.59E-02
.....SPF..R.....NR..FED..NF.....	1.63E-02
.....APF.HK.....NR..EQD.YN.E.....	1.74E-02
.....GPF.DK.....NR..FED..NF.....	1.75E-02
.....GPF..I.....NR..FED..NF.....	1.76E-02
.....GPF.SL.....NR..FED..NF.....	1.81E-02
.....PF..D.....NR..FED..NF.....	1.84E-02
.....VPF..D.....NR..FYD..N.....	1.85E-02
.....S.R.AR.....NR..FED..NF.....	1.85E-02
.....SPF.EN.....NR..FYD..N.....	1.91E-02
.....PF.AN.....NR..FYD..N.....	1.93E-02
.....GPF..Y.....NR..FYD..N.....	1.96E-02
.....VE.....NY..FTD..NFE.....	2.43E-02
.....SPY..T.....NR..FED..NF.....	2.51E-02
.....GPY.SY.....NR..FED..NF.....	2.53E-02
.....NPY.RR.....NR..FED..NF.....	3.30E-02
.....VPY.RR.....NR..FED..NF.....	3.41E-02
.....NPY.RI.....NR..FYD..N.....	3.69E-02
.....TPF.D.....NR..FED..NF.....	4.09E-02
.....APY.SY.....NR..FED..NF.....	4.41E-02
.....GPF.DR.....NY..QYE.TS.....	4.81E-02
.....TPF.EW.....NR..FED..NF.....	5.29E-02
.....DY.....LF..YD..N.E.....	5.35E-02
.....VE.....LF..YD..N.E.....	6.01E-02
.....APF.DR.....RY..EDD.EN.E.....	6.35E-02
.....NPF..Y.....NR..FYD..N.....	6.98E-02
.....GPF.DR.....NY..QFD.YS.E.....	8.23E-02
.....VPY.TR.....NR..FED..NFG.....	1.15E-01

Supplementary Table.6. Biophysical characterization data of VHH_h1-VHH_h5.

Name	CDR2	CDR3	K_D BLI (nM) (hVsig4)	K_D SPR (nM) (hVsig4)	T_m (°C)	T_{agg} (°C)	ΔG_{Urea} (kcal/mol)	ΔG_{GdnHCl} (kcal/mol)
VHH_WT	AI**RWN**G**GS**TY	GRW**DK**YG**SSF**Q**DEY**DY	3190	12800	79.3	81.7	-10.2	-10.70
VHH_h1	..GPF.YE..	...NR..FED..NF..	14.2	5.7	86.2	86.1	☆	-13.30
VHH_h2	..SPF..N..	...NR..FED..NF..	4.5	4.6	89.6	89.5	☆	-13.31
VHH_h3	..GPF..T..	...NR..EQD.YN.E.	37.6	13.5	91.0	90.2	☆	-13.75
VHH_h4	..GPF..H..	...NR..FYD..N...	5.6	4.8	90.0	88.7	☆	-14.70
VHH_h5	..GPF.....	...NR..FED..NF..	4.1	2.8	85.8	85.2	☆	-12.89

(Note: CDR definition: AbM, ☆ means it was not denatured under the most extreme conditions)

986 **Supplementary Table.7.** Crystallization data for VHH_h1-5 in complex with hVsig4.

Protein complex	**hVsig4-VHH_h1**	**hVsig4-VHH_h2**	**hVsig4-VHH_h3**	**hVsig4-VHH_h4**	**hVsig4-VHH_h5**
Crystallization conditions	0,17M ammonium sulfate 15% glycerol 25,5% PEG 4000	0,1M Na-citrate pH 5,6 20% 2-propanol 21% PEG 4000	0,1M Na-citrate pH 5,6 20% 2-propanol 21% PEG 4000	0,1M Bis-Tris pH 6.0 24% PEG3350	0,1M Bis-Tris pH 6.5 30% PEG3350
Temperature (K)	287	287	294	293	293
Protein concentration (mg/ml)	4,2	4,2	5,6	3,80	4,5
Cryoprotectant	none	25% ethylene glycol	20% ethylene glycol	15% glycerol	15% glycerol
Data collection[a]					
X-ray source	PETRA III (P13)	PETRA III (P13)	PETRA III (P13)	ALBA (Xaloc)	ALBA (Xaloc)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, *b*, *c* (Å)	50.651, 51.404, 89.649	52.291 53.069 87.531	60.867 129.940 144.280	25.7, 51.0, 175.2	25.7, 50.9, 175.9
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Unique reflections	109581 (17482)	43885 (7006)	89487 (14083)	18076 (928)	17284 (825)
Wavelenght (Å)	0.9762	0.9762	0.9762	0.9793	0.9793
Resolution (Å)	51.40–1.05 (1.11-1.05)	53.07-1.45 (1.54-1.45)	129.94-1.90 (2.01-1.90)	87.62-1.94 (1.97-1.94)	87.95-1.97 (2.00-1.97)
R_{meas} (%)	10.0 (102.8)	8.8 (218.6)	19.8 (247.1)	9.6 (109.4)	11.3 (99.2)
$\langle I/\sigma \rangle$	12.40 (1.59)	17.77 (1.13)	13.61 (1.19)	16.0 (2.5)	13.5 (2.4)
Completeness (%)	99.90 (99.42)	99.90 (100)	98.3 (97.1)	100 (100)	100 (100)
Redundancy	13.3 (11.9)	13.3 (13.3)	13.9 (13.5)	12.6 (12.8)	12.5 (12.3)
CC (½) (%)	99.7 (75.7)	100 (42.3)	99.8 (45.6)	99.8 (91.0)	99.8 (90.9)
Wilson B factor ($Å^2$)	15.02	27.10	36.57	31.67	31.15
Refinement[b]					
Resolution (Å)	24.71 - 1.05 (1.06 - 1.05)	33.56 - 1.45 (1.502 - 1.45)	64.97 – 1.90 (1.92 – 1.90)	44.05-1.94 (2.009-1.94)	44.04-1.97 (2.04-1.97)
No. reflections	109570 (10760)	43876 (4327)	89397 (8824)	17989 (1772)	17159 (1630)
R_{work} (%)	15.67	17.34	19.08	19.63	19.75
R_{free} (%)	17.57	21.51	22.83	24.01	24.30
No. non-H atoms	2328	2162	8376	2060	1985
Protein	2069	1948	7810	1937	1918
Ligand/ion	5	36	47	6	-
Water	254	178	519	117	67
B-factors ($Å^2$)	20.06	27.74	35.86	54.69	48.46
Protein	18.42	26.52	35.62	54.99	48.61
Ligand/ion	17.25	40.08	47.68	49.02	-
Water	33.45	38.61	38.35	49.97	44.27
R.m.s. deviations					
Bond lengths (Å)	0.012	0.002	0.007	0.008	0.009
Bond angles (°)	1.25	0.65	0.87	1.07	1.06
Ramachandran favored (%)	97.52	96.64	96.79	95.3	95.3
Ramachandran allowed (%)	2.48	3.36	3.21	4.7	4.7
Ramachandran outliers (%)	0	0	0	0	0
PDB ID	9EZU	9EZV	9EZW	9EZH	9EZI

987 [a] Each data set was collected from a single crystal. Values in parentheses correspond to the highest resolution shell. Values reported by XDS.

988 [b] Values reported by Phenix

989 **Supplementary Table.8.** Interactions at the VHH_WT:hVSIG4 interfaces compared to the

990 interactions at the VHH_h:hVSIG4 interfaces.

hVsig4:VHH_WT interface (5IML)							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_WT	hVsig4		VHH_WT	hVsig4		VHH_WT
Ile 3	**CDR1:**			**CDR3:**			**CDR3**
Pro 7	Thr 27	Ile 111	-	Arg 99	Glu 95	-	Lys 102
Glu 8	Ser 29	Val 6	-	Arg 99	Arg 108	-	Glu 110
Lys 36	Ser 30	Val 6	-	Arg 99			
Leu 38	Tyr 31	Asp 109	-	Asp 101			
Gln 40	**CDR2:**	Asp 109	-	Lys 102			
Ser 43	Arg 51	Asp 109	-	Tyr 103			
Asp 44	Trp 52						
Pro 45	Asn 53						
His 91	**CDR3:**						
Thr 93	Trp 100						
Val 107	Gly 104						
Lys 110	Phe 107						
Glu 113	Asp 109						
	Glu 110						
	Tyr113						

hVsig4:VHH_h1 interface							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_h1	hVsig4		VHH_h1	hVsig4		VHH_h1
Ile 3	**CDR1:**			**CDR1:**			**CDR1:**
Pro 7	Thr 27	Glu 8	-	Arg 26	Glu 8	-	Arg 26
Leu 38	Ser 30						**CDR3**
Gly 42	Tyr 31			**CDR3:**	Ile 111	-	Arg 99
Ser 43	**CDR2:**	Ile 111	-	Arg 99	Glu 95	-	Arg 102
Pro 45	Pro 52	Val 6	-	Arg 99	Arg 108	-	Asp 107
His 91	Phe 53	Gln 40	-	Asn 101	Arg 1	-	Asp 107
Val 107	Gly 54	Thr 93	-	Asn 101	Arg 1	-	Asp 109
Lys 110	Glu 56	Glu 95	-	Arg 102			
	CDR3:	Asp 109	-	Asn 101			
	Trp 100	Asp 109	-	Arg 102			
	Gly 104	Asp 109	-	Tyr 103			
	Asn 110	Arg 108	-	Asp 107			
	Asp 112	Arg 1	-	Asp 107			
		Arg 1	-	Asp 109			

hVsig4:VHH_h2 interface							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_h2	hVsig4		VHH_h2	hVsig4		VHH_h2
Ile 3	**CDR1:**			**CDR1:**			**CDR3**
Pro 7	Ser 30	Glu 8	-	Thr 27	Glu 95	-	Arg 102
Leu 38	**CDR2:**	Glu 8	-	Tyr 31	Arg 108	-	Asp 107
Gly 42	Pro 52			**CDR3:**	Arg 1	-	Asp 107
Ser 43	Phe 53	Ile 111	-	Arg 99	Arg 1	-	Asp 109
Pro 45	Gly 54	Val 6	-	Arg 99			
His 91	Glu 56	Gln 40	-	Asn 101			
Val 107	**CDR3:**	Thr 93	-	Asn 101			
Lys 110	Trp 100	Glu 95	-	Arg 102			
	Gly 104	Asp 109	-	Asn 101			
	Asn 110	Asp 109	-	Arg 102			
	Asp 112	Asp 109	-	Tyr 103			
		Arg 108	-	Asp 107			
		Arg 1	-	Asp 107			
		Arg 1	-	Asp 109			

hVsig4:VHH_h3 interface							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_h3	hVsig4		VHH_h3	hVsig4		VHH_h3
Ile 3	**CDR1:**			**CDR1:**			**CDR1**
Glu 5	Thr 27	Glu 8	-	Arg 26	Glu 8	-	Arg 26
Lys 36	Ser 30						**CDR3**
Leu 38	Tyr 31			**CDR3:**	Glu 95	-	Arg 102
Gly 42	**CDR2:**	Ile 111	-	Arg 99	Arg 108	-	Asp 107
Ser 43	Pro 52	Val 6	-	Arg 99	Lys 110	-	Glu 112
Pro 45	Phe 53	Gln 40	-	Asn 101			
His 91	Gly 54	Thr 93	-	Asn 101			
Val 107	**CDR3:**	Glu 95	-	Arg 102			
Glu 113	Trp 100	Asp 109	-	Arg 102			
Arg 115	Gly 104	Asp 109	-	Tyr 103			
	Tyr 109	Arg 108	-	Asp 107			
	Tyr 113	Lys 110	-	Asn 110			
		Lys 110	-	Glu 112			

hVsig4:VHH_h4 interface							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_h4	hVsig4		VHH_h4	hVsig4		VHH_h4
Ile 3	**CDR1:**			**CDR1:**			**CDR1**
Lys 36	Ser 30	Glu 8	-	Arg 26	Glu 8	-	Arg 26
Leu 38	Tyr 31						**CDR3**
Gly 42	**CDR2:**			**CDR3:**	Glu 95		Arg 102
Ser 43	Phe 53	Ile 111	-	Arg 99	Arg 1	-	Asp 107
Pro 45	Pro 52	Val 6	-	Arg 99	Arg 108	-	Asp 107
His 91	Gly 54	Val 6	-	Arg 99	Arg 1	-	Asp 109
Val 106	Gly 55	Gln 40	-	Asn 101		-	
Val 107	**CDR3:**	Thr93	-	Asn 101			
Lys 110	Trp 100	Asp 109	-	Asn 101			
	Gly 104	Glu 95	-	Arg 102			
	Asn 110	Asp 109	-	Arg 102			
	Asp 112	Asp 109	-	Tyr 103			
		Arg 108	-	Tyr 106			
		Arg 1	-	Asp 107			
		Arg 108	-	Asp 107			
		Arg 1	-	Asp 109			

hVsig4:VHH_h5 interface							
Van der Waals contacts		Potential H-bonds			Potential salt bridges		
hVsig4	VHH_h5	hVsig4		VHH_h5	hVsig4		VHH_h5
Ile 3	**CDR1:**			**CDR1:**			**CDR1**
Lys 36	Ser 30	Glu 8	-	Tyr 31	Glu 8	-	Arg 26
Leu 38	**CDR2:**			**CDR3:**			**CDR3**
Gly 42	Pro 52	Ile 111	-	Arg 99	Glu 95	-	Arg 102
Ser 43	Phe 53	Val 6	-	Arg 99	Arg 1	-	Asp 107
Pro 45	Gly 54	Pro 7	-	Arg 99	Arg 108	-	Asp 107
His 91	**CDR3:**	Gln 40	-	Asn 101	Arg 1	-	Asp 109
Val 107	Trp 100	Thr93	-	Asn 101			
	Gly 104	Asp 109	-	Asn 101			
	Asp 112	Glu 95	-	Arg 102			
	Tyr 113	Asp 109	-	Arg 102			
		Asp 109	-	Tyr 103			
		Arg 1	-	Asp 107			
		Arg 108	-	Asp 107			
		Arg 1	-	Asp 109			
		Lys 110	-	Asn 101			

991