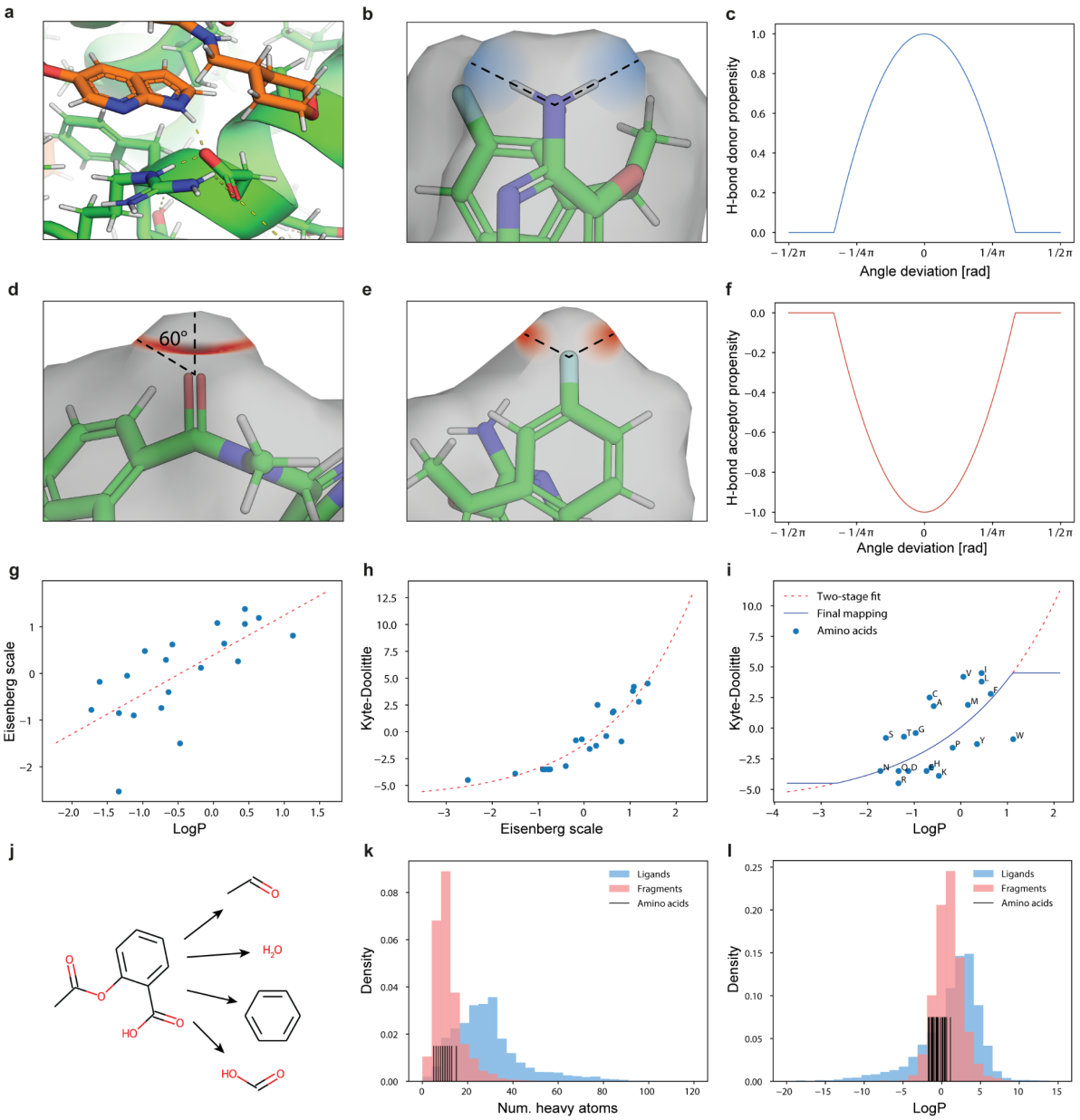
nature portfolio

https://doi.org/10.1038/s41586-024-08435-4

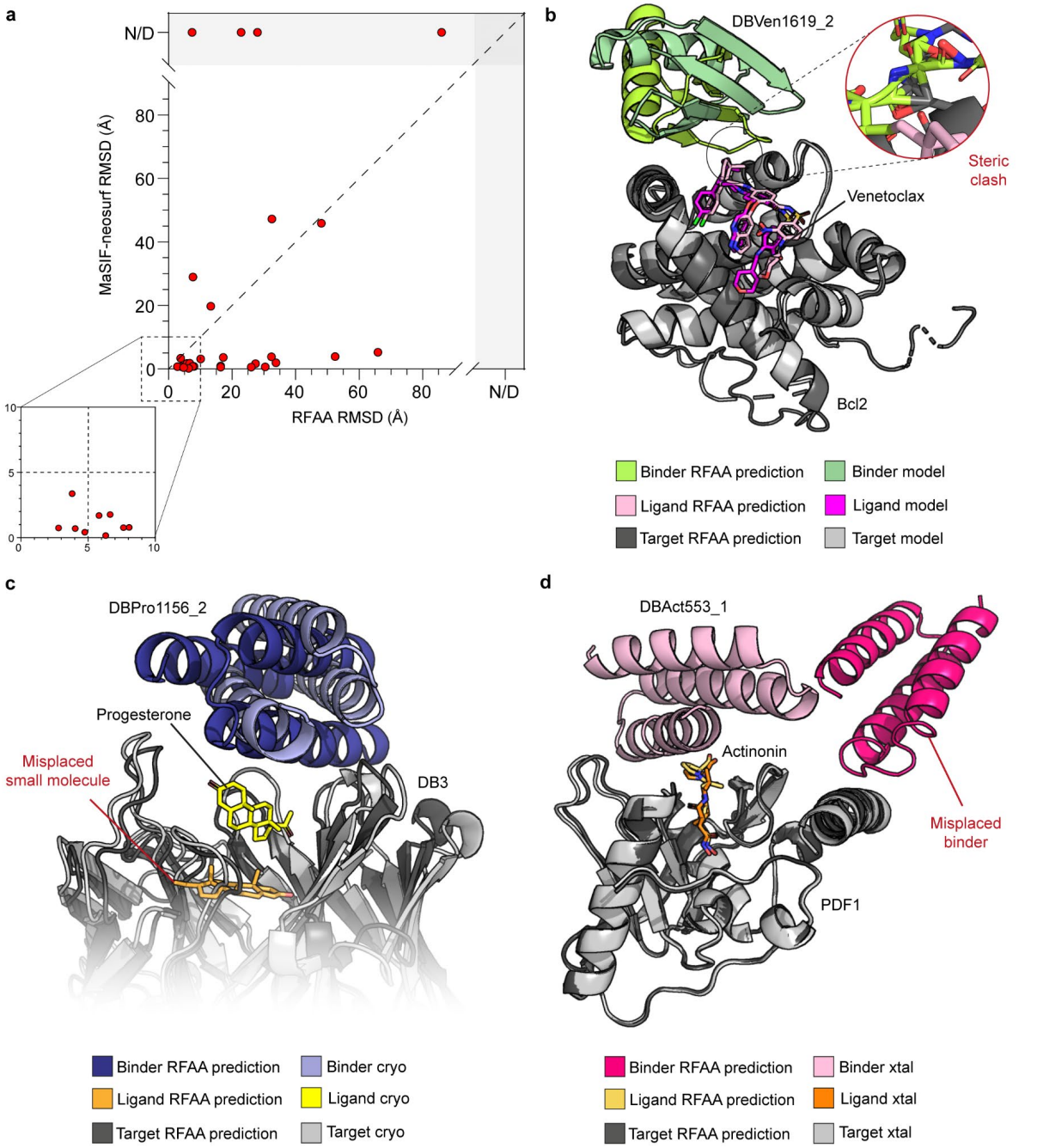
Supplementary information

Targeting protein–ligand neosurfaces with a generalizable deep learning tool

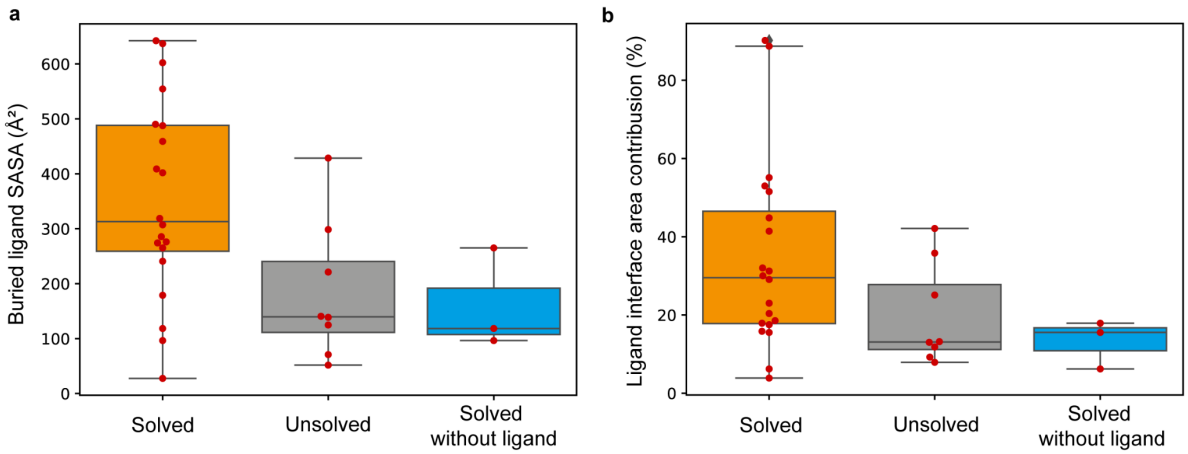
In the format provided by the authors and unedited



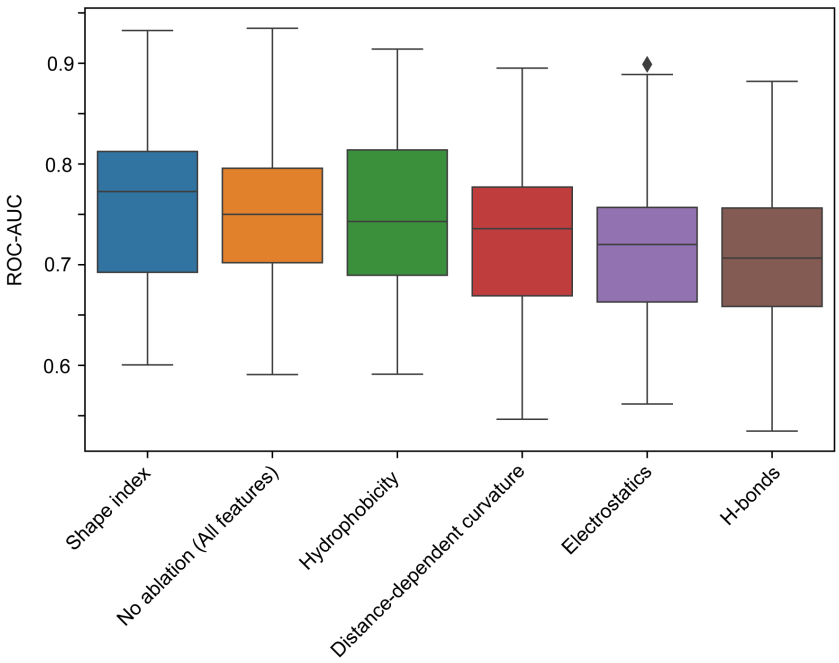
Supplementary Figure 1: MaSIF feature computation for small molecule ligands. a-f. Hydrogen bond propensity is assigned in a direction-dependent manner[33]. Surface points are assigned positive (donor) or negative (acceptor) values based on their distance to the ideal direction (c, f). The optimal position for an acceptor can either lie anywhere on a cone (d) or in one or more unique directions (e). **g-i.** For hydropathy, we convert computational LogP values to the protein-specific Kyte-Doolittle (KD) scale required by MaSIF using the Eisenberg scale as an intermediate (g-h). We also restrict the outputs to be between the minimum and maximum KD values after the mapping leading to the relationship shown in panel (i). **j-l.** Ligands were fragmented into smaller objects (j). The reason is that most ligands are significantly larger than any amino acid (k) and exhibit more extreme LogP values (l). We therefore compute hydrophobicity values based on fragments. This procedure ensures that the new hydropathy feature remains “in-distribution” of the pre-trained MaSIF model.



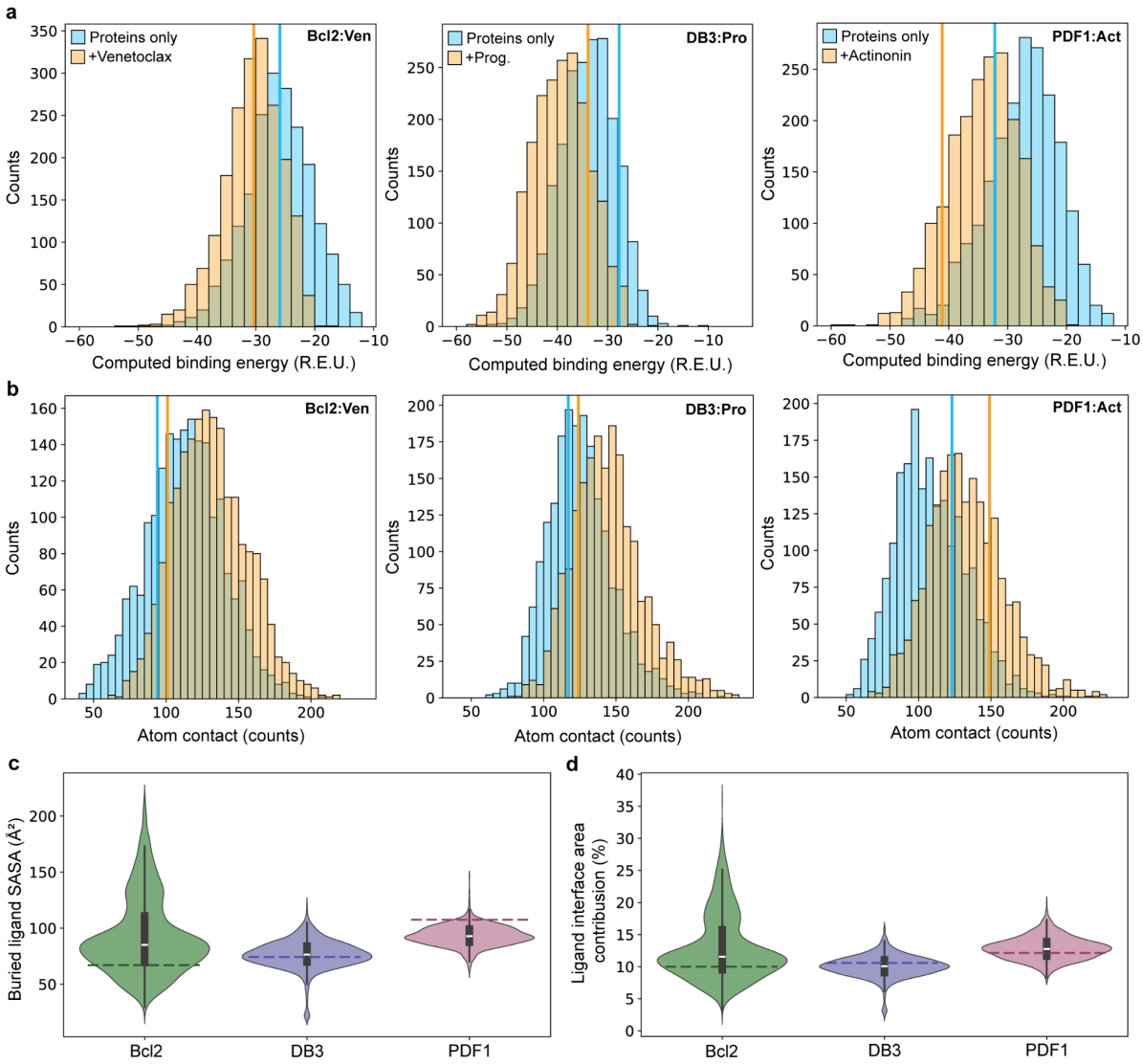
Supplementary Figure 2: Comparison between RoseTTAFold All-Atom prediction and validated structures. a. Comparison of the root mean square deviation (RMSD) obtained by MaSIF-neosurf compared to RoseTTAFold All-Atom (RFAA). Measurements were performed on the 28 complexes used in the benchmark dataset (see Supplementary Table 4). Outputs that were not successfully determined are labeled as N/D. RMSD below 5Å were considered as successful poses. **b-d.** Experimentally-validated structures DBVen1619_2 (b), DBPro1156_2 (c) & DBAct553_1 (d), were aligned based on the respective target protein. Validated binders were represented by their computational models, cryoEM structure or crystal structure respectively. Misfolded or misplaced regions are highlighted with red. All RFAA predictions were performed with 30 recycles with the recommended settings.



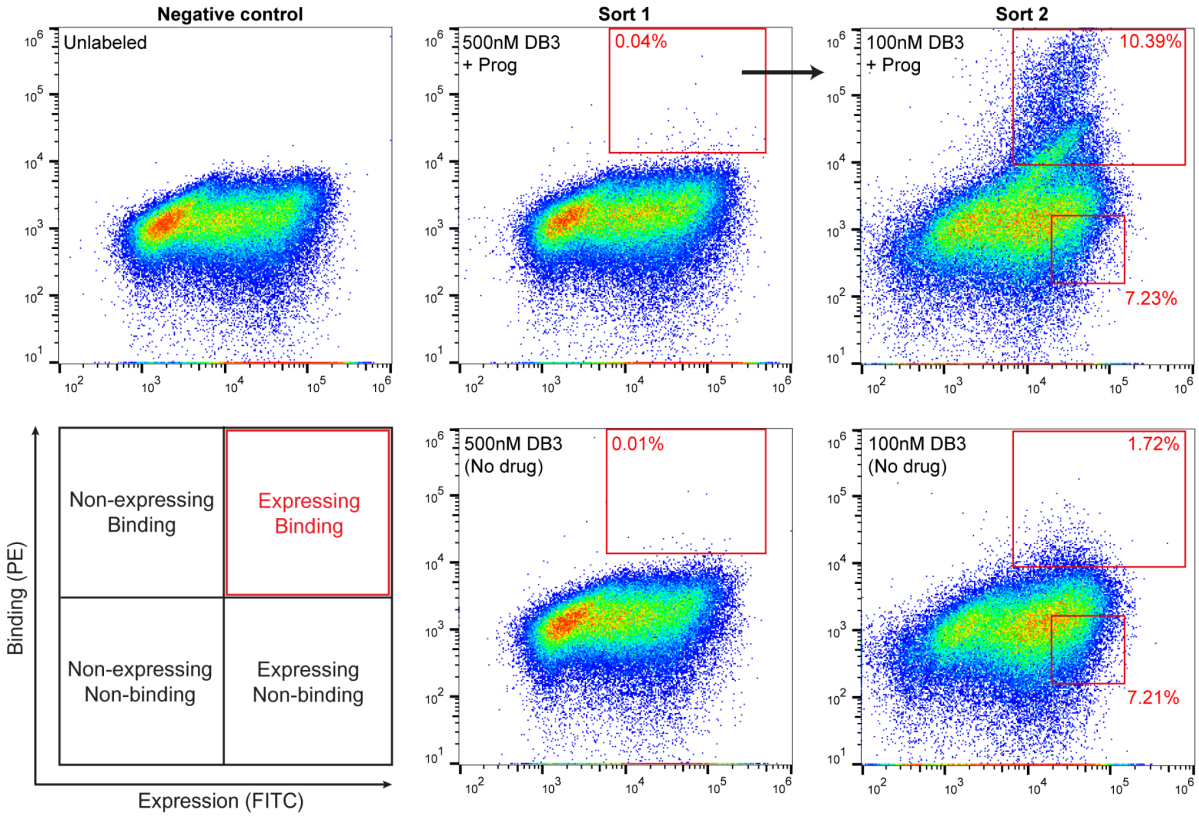
Supplementary Figure 3: Ligand interface area contribution in the benchmark dataset. a-b. Buried solvent-accessible surface area (SASA) contribution of the ligand in absolute value (a) or in percentage of the total interface surface area (b) for the 28 protein-ligand complexes used in the benchmark dataset of known ternary complexes. Complexes were categorized based on the benchmark outcome, namely successfully solved complexes with the ligand (orange), unsolved complexes with the ligand (gray) or solved without the ligand (blue). Boxes represent quartiles and whiskers show data points within 1.5x the interquartile range. Data represent all independent protein-ligand complexes.



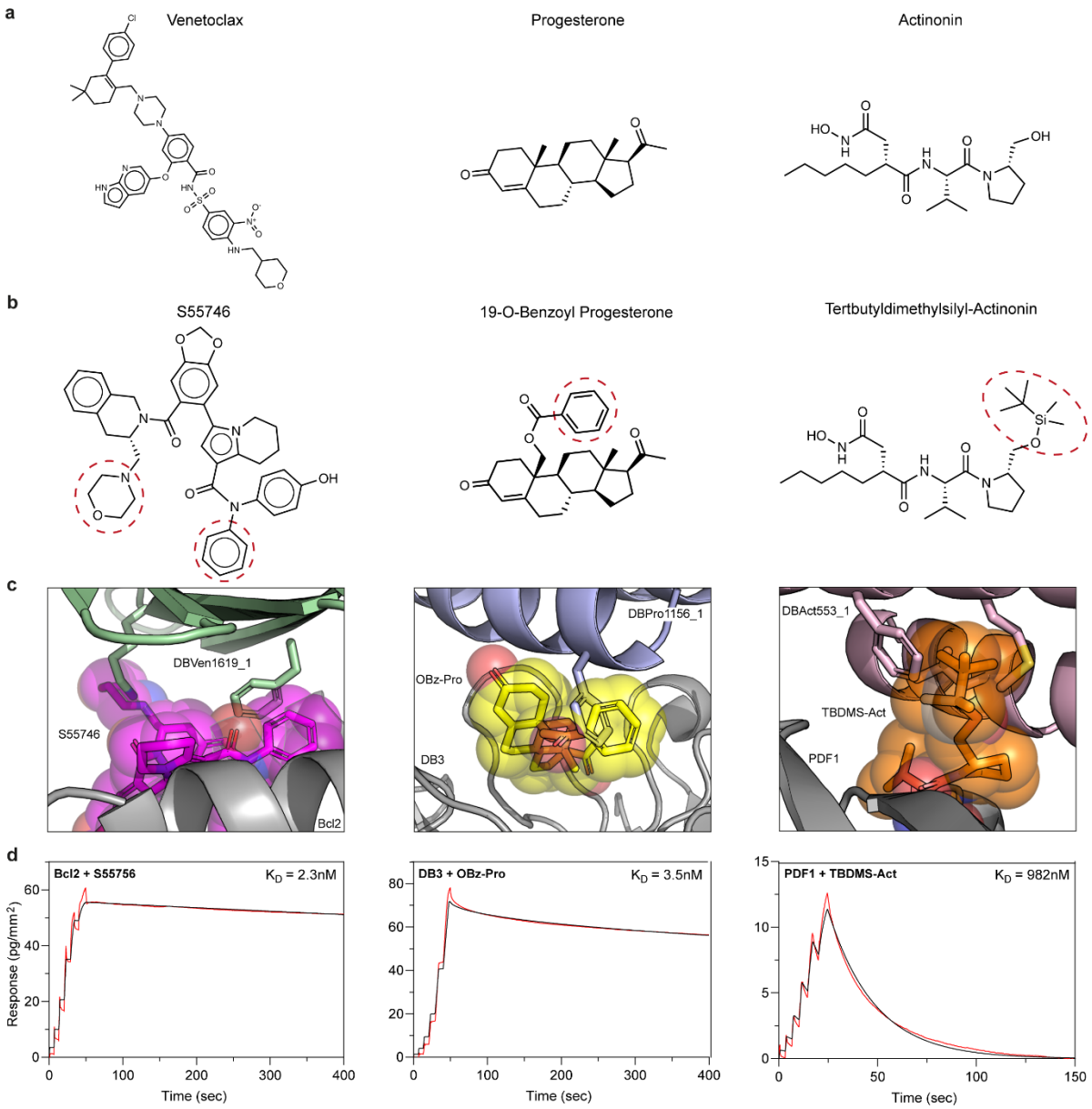
Supplementary Figure 4: MaSIF feature ablation study. Discriminative power of MaSIF-neosurf with retrained models ablating one feature at a time. Performance was assessed on the dataset of 28 complexes as for the computational benchmark. It is measured as the area under the receiver operating characteristic (ROC-AUC) while using interacting patches (center points at most 1 Å apart) as the positive category and an equal number of random non-interacting pairs (distance between center points >5 Å) as negatives. The classification criterion is the Euclidean distance between learned descriptor vectors. Boxes represent quartiles and whiskers show data points within 1.5x the interquartile range. Data outside this range are shown as flier points. Data represent all 28 independent protein-ligand complexes.



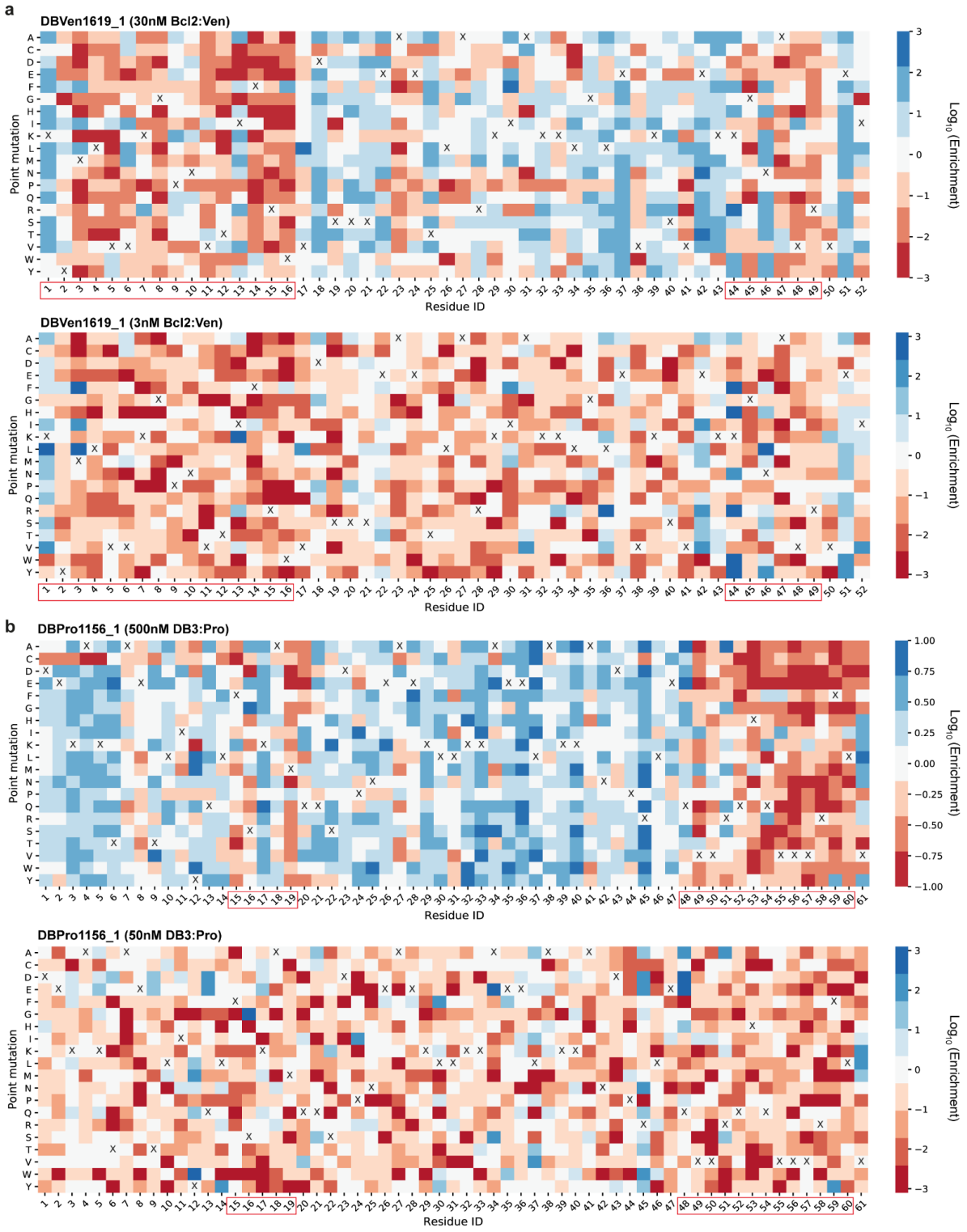
Supplementary Figure 5: Neosurface properties captured by the designed binders. a-b. Computed binding energy (ΔΔG, panel a) and number of atomic contacts (b) for all designs targeting Bcl2:Venetoclax (left), DB3:Progesterone (middle) and PDF1:Actinonin (right) complexes. Calculations were done in absence (blue) and presence (orange) of the respective small molecules. Atom contacts were defined based on the Van der Waals radii (r_{VdW} + 0.2 Å tolerance) of each pair of atoms. Vertical lines represent the identified binder for each targeted complex. **c-d**. Buried solvent-accessible surface area (SASA) of the ligand (c) and ligand contribution with respect to the total buried SASA (d) for each target protein-ligand complex. Dashed lines represent the identified binder for each targeted complex. Inner boxes represent the quartiles and whiskers show data points within 1.5x the interquartile range. 1995 computational designs were plotted for Bcl2:Venetoclax, 1998 for DB3:Progesterone and 1997 for PDF1:Actinonin.

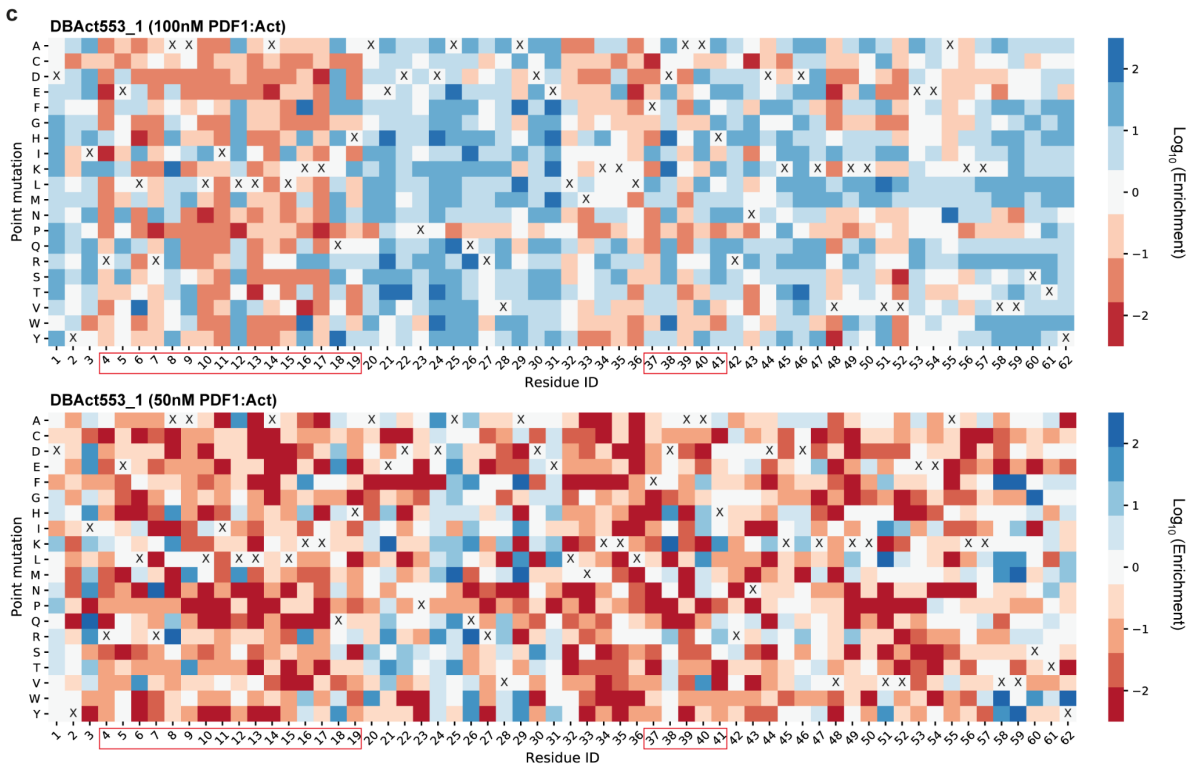


Supplementary Figure 6: Representative flow cytometry graphs of the binder screening. Yeast surface display screening of the first and second sort of the library against DB3:Progesterone in presence (+Prog) and absence of the small molecule. Yeasts labeled with secondary antibodies but without any ligand were used as a negative control to set the gates. In sort 1, binding population from the selected gates was used for a second sort. In sort 2, yeasts were sorted for both binding (upper gate) and non-binding (lower gate) populations and used for next-generation sequencing.

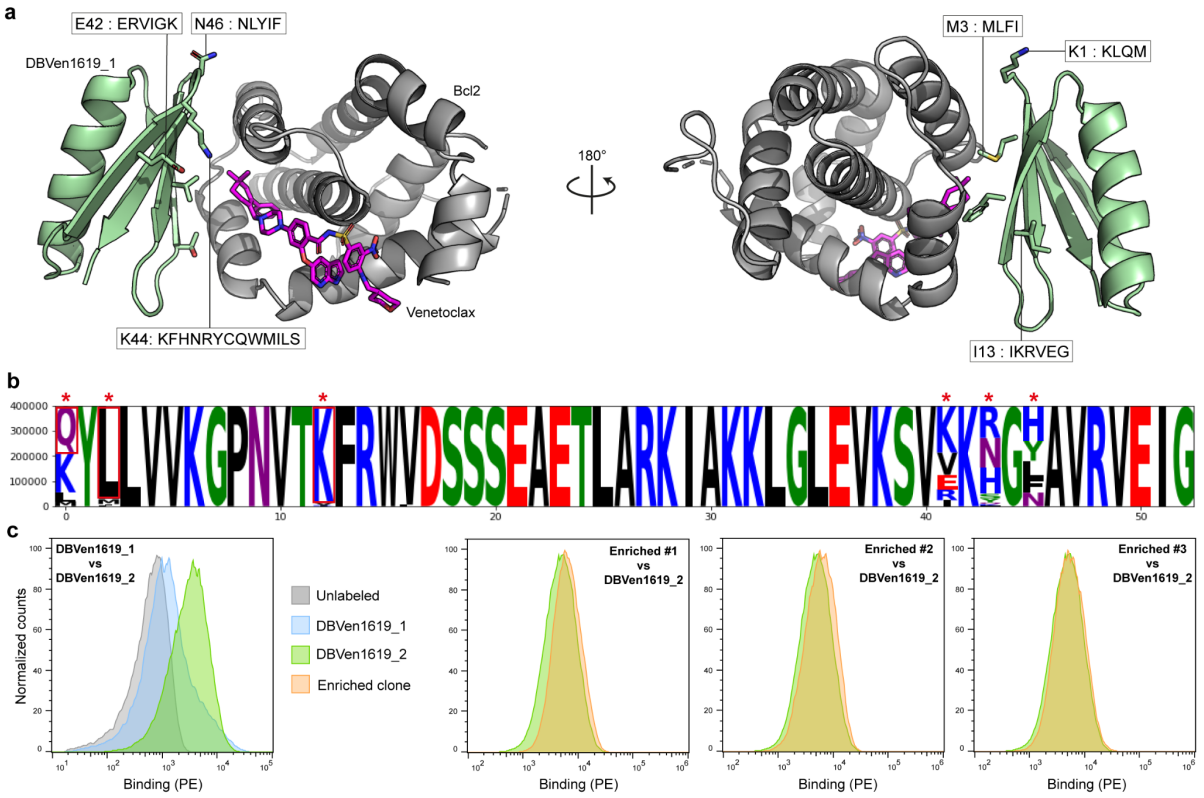


Supplementary Figure 7: Binding control of small molecules analogs. a. Original small molecule used for each drug:protein complex. **b.** Small molecule derivative binding to the same target but introducing a clash with the designed binder. Steric clashes are indicated with red dashed circles. **c.** Crystal structure (S55746; PDB: 6GL8) or computational models (19-O-Benzoyl Progesterone and Tertbutyldimethylsilyl-Actinonin) of the small molecule analogs relaxed by Gnina with their respective target protein in absence of the designed binders. Complexes were then overlapped with the computational model of the designed binders to identify clashing regions. **d.** WaveRAPID binding kinetics of each small molecule analog to their respective target protein performed by Grating-Coupled Interferometry (GCI). Measurements are indicated in red and fit curves in black.

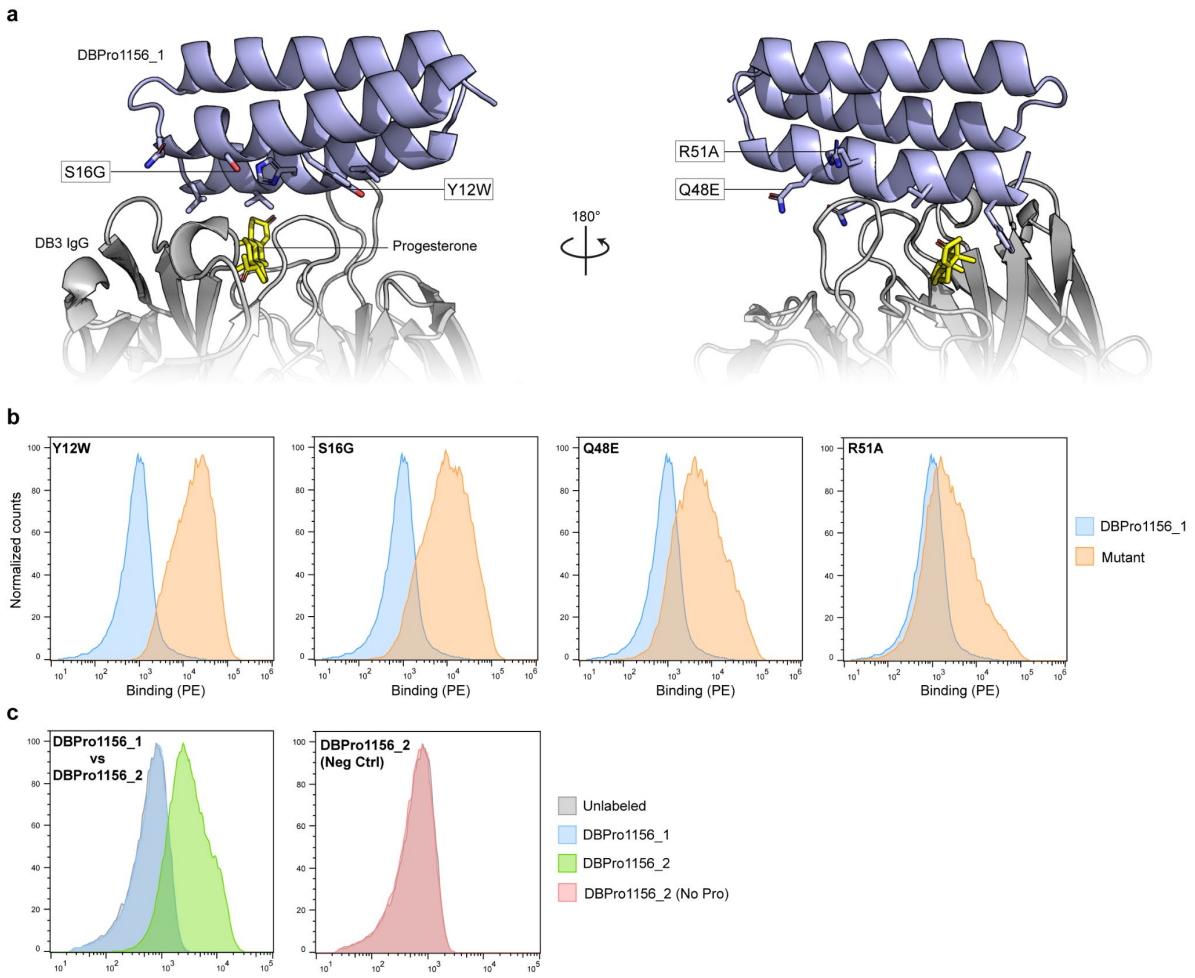

a
DBVen1619_1 (30nM Bcl2:Ven)
DBVen1619_1 (3nM Bcl2:Ven)
Point mutation
Residue ID
Log10 (Enrichment)
b
DBPro1156_1 (500nM DB3:Pro)
DBPro1156_1 (50nM DB3:Pro)



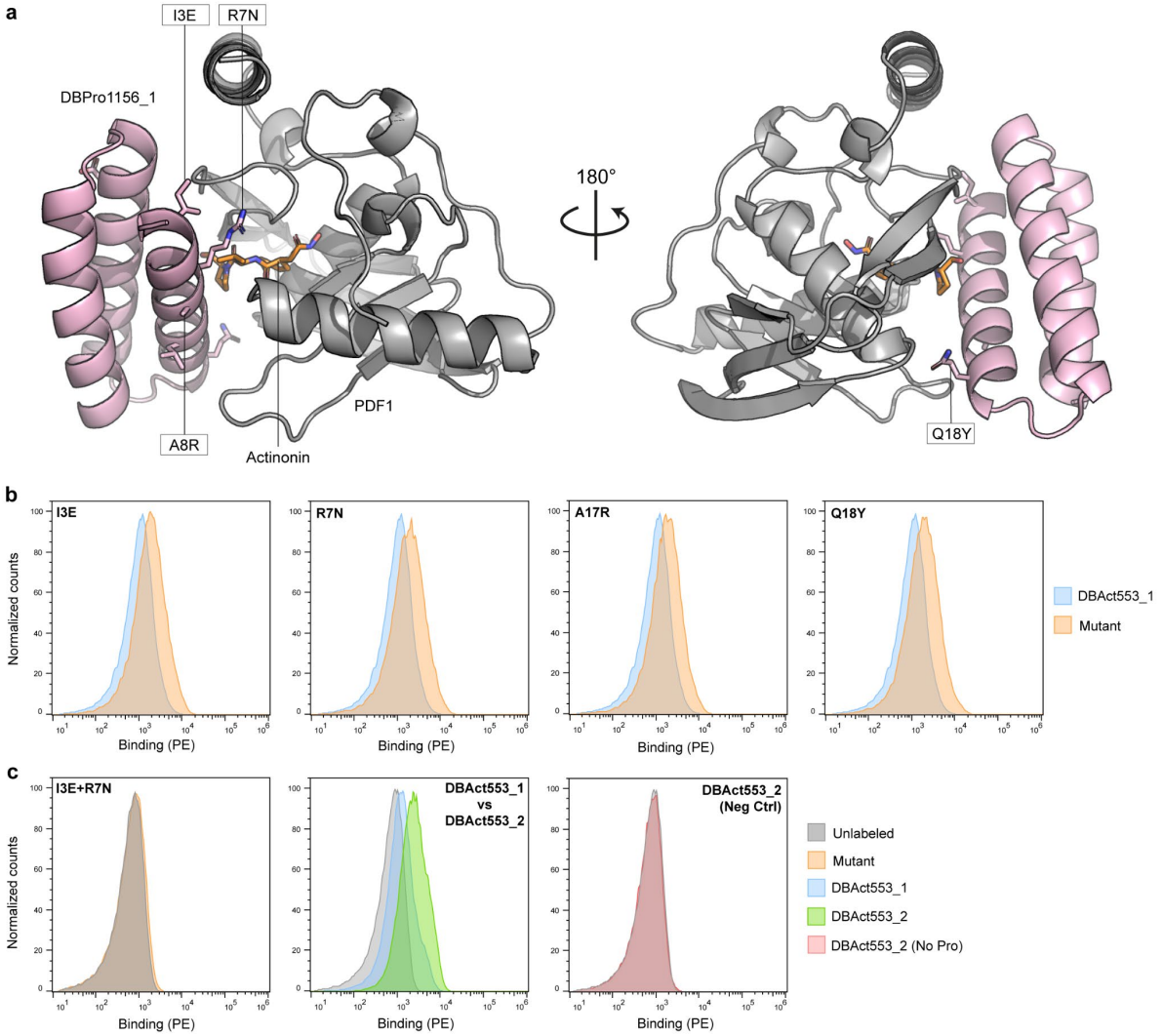
Supplementary Figure 8: Full data of the site-saturation mutagenesis. a-c. Heatmaps represent the logarithmic value of the enrichment score (Counts in binding population divided by counts in non-binding population) of each mutation at every position. Sorting has been performed following the gating strategy presented in Supplementary Fig. 6. Native amino acids are marked with a cross (X) and near-interface positions are highlighted with a red box. Sorts were performed with a high concentration and a low concentration of target complex to focus on deleterious and beneficial mutations respectively. Site saturation mutagenesis were performed on DBVen1619_1 for the Bcl2:Venetoclax complex (a), DBPro1156_1 for the DB3:Progesterone complex (b) and DBAct553_1 for the PDF1:Actinonin complex (c).



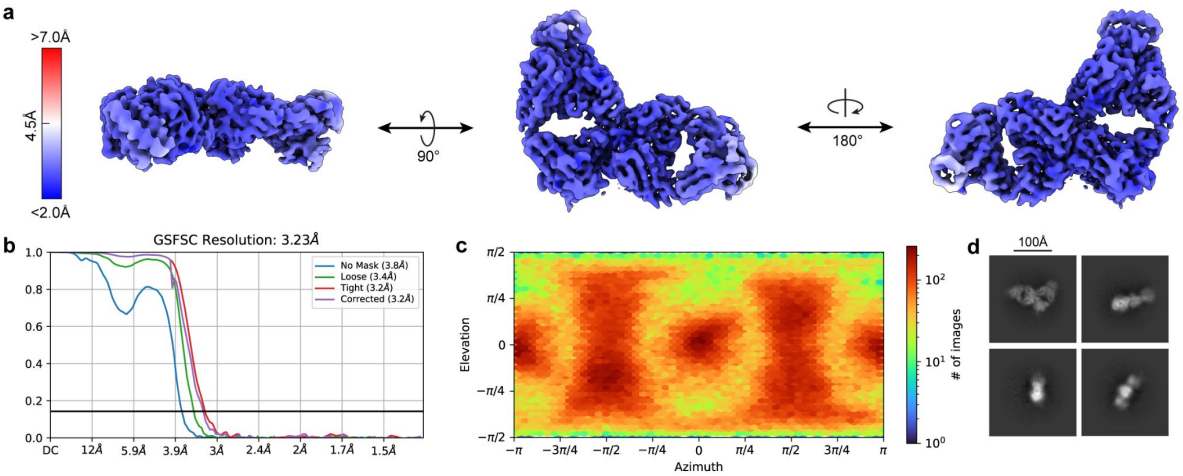
Supplementary Figure 9: Experimental optimization of DBVen1619. a. Computational model of DBVen1619_1 (green) in complex with Bcl2 (gray) and Venetoclax (Magenta). Potential beneficial mutations obtained from site-saturation mutagenesis (SSM) data and subsequent degenerate codons are found in black boxes for each mutated position. **b.** Sequence logo plot of the combinatorial library sorted twice with yeast display. Mutated positions are highlighted with a red asterisk. Mutations selected to constitute DBVen1619_2 are highlighted with a red square. **c.** Comparison of unlabeled yeast (gray), or yeasts displaying DBVen1619_1 (blue), DBVen1619_2 (green, K1Q+M3L+I13K) or the top 3 most enriched sequences of the combinatorial library (orange). Yeast cells were labeled with 3 nM Bcl2:Venetoclax complex.



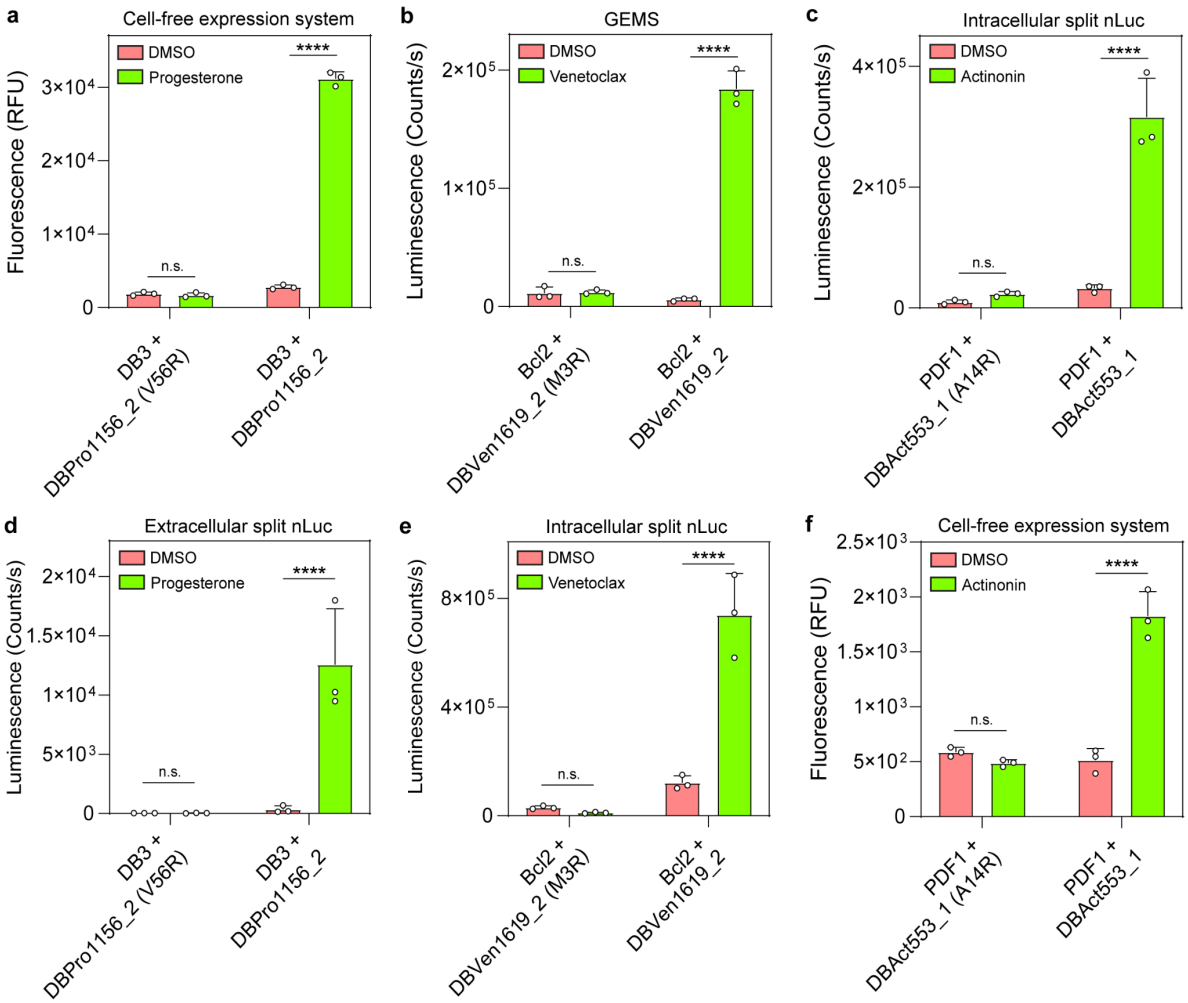
Supplementary Figure 10: Experimental optimization of DBPro1156. a. Computational model of DBPro1156_1 (blue) in complex with DB3 IgG (gray) and Progesterone (yellow). Potential beneficial mutations obtained from site-saturation mutagenesis (SSM) data are found in black boxes for each mutated position**. b.** Binding signal measured by flow cytometry for yeasts displaying DBPro1156_1 (blue) or the corresponding mutant (orange). All yeasts were labeled with 50 nM DB3:Progesterone complex. **c.** Binding signal measured by flow cytometry measured for yeasts displaying DBPro1156_1 (blue) compared to DBPro1156_2 (green, Y12W+S16G). Yeasts were labeled with 3 nM DB3:Progesterone complex. Negative control performed on yeast displaying DBPro1156_2 labeled with DB3 only.



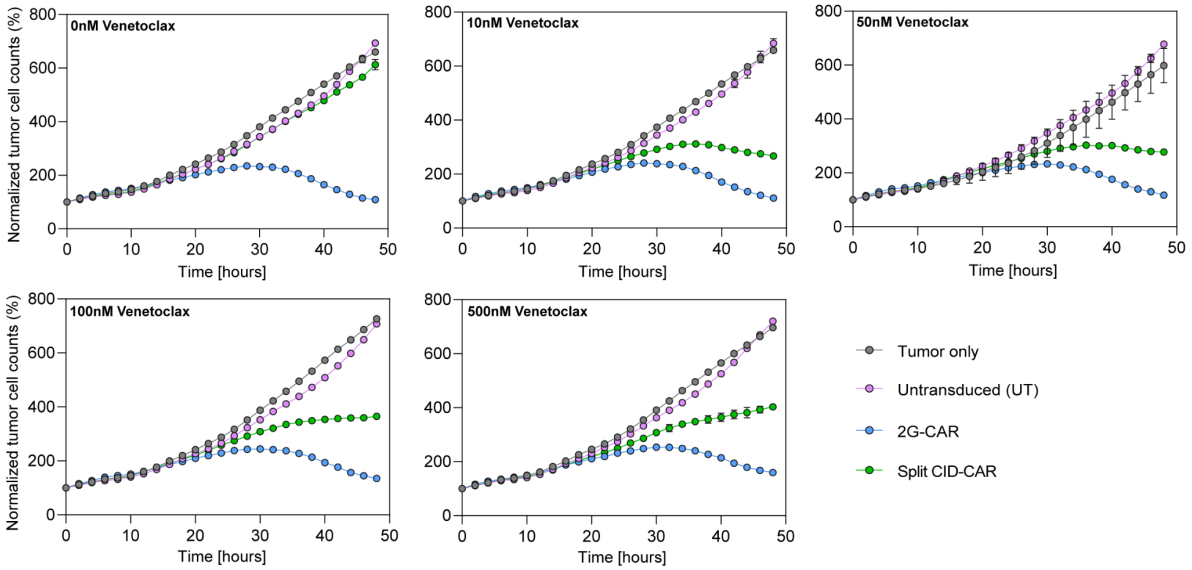
Supplementary Figure 11: Experimental optimization of DBAct553. a. Computational model of DBAct553_1 (pink) in complex with PDF1 (gray) and Actinonin (orange). Potential beneficial mutations obtained from site-saturation mutagenesis (SSM) data are found in black boxes for each mutated position**. b.** Binding signal measured by flow cytometry for yeasts displaying DBAct553_1 (blue) or the corresponding mutant (orange). All yeasts were labeled with 10 nM PDF1:Actinonin complex. **c.** Binding signal measured by flow cytometry measured for yeasts displaying DBAct553_1 (blue) compared to DBAct553_2 (green, R7N+A8R). Yeasts were labeled with 10 nM PDF1:Actinonin complex. Negative control performed on yeast displaying DBAct553_2 labeled with DB3 only. Combination of mutants I3E+R7N was deleterious for binding as compared to unlabeled yeasts.



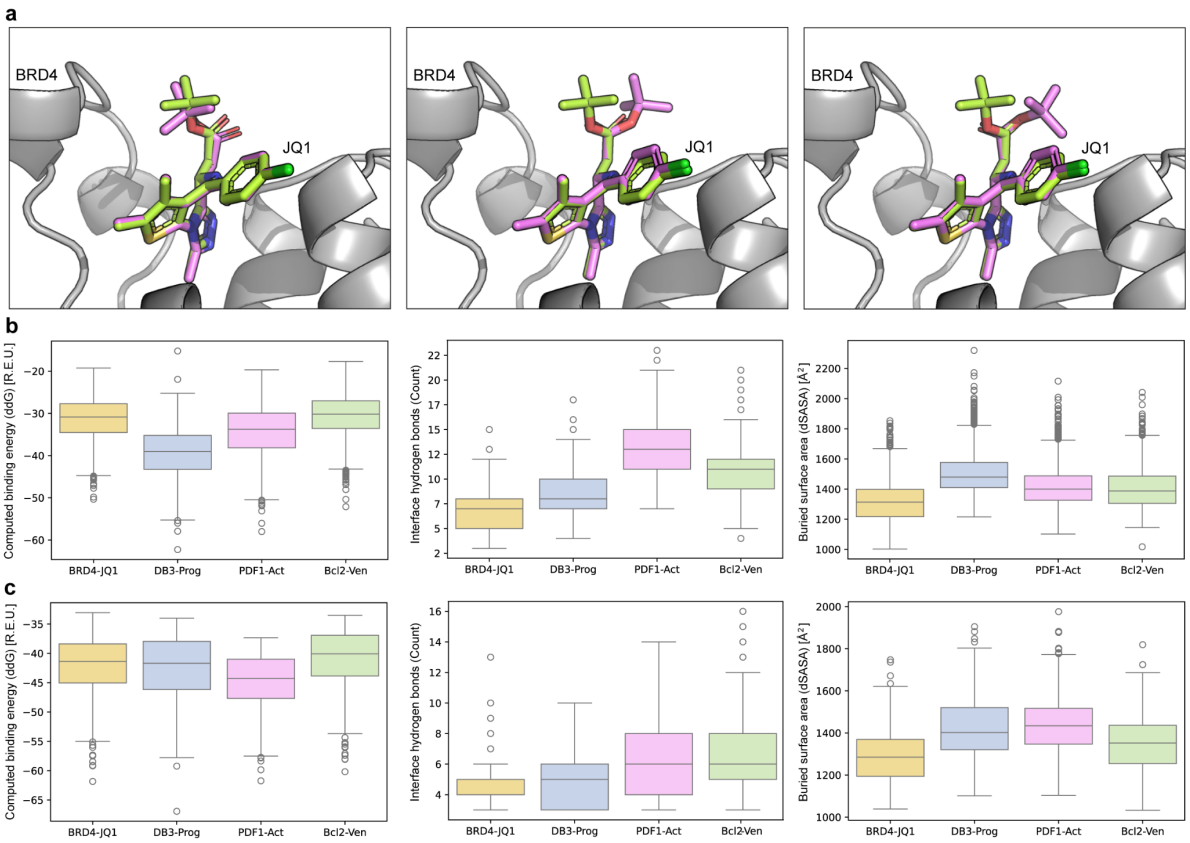
Supplementary Figure 12: Details of Cryo-EM data processing for DBPro1153_2 in complex with DB3. a. The cryo-EM map of anti-Kappa IgG:DB3:DBPro1156_2 used for model building. Views of the unsharpened cryo-EM density maps colored by local resolution. **b.** Gold-standard FSC curve with resolution cutoff indicated at 0.143. **c.** Particle distribution heatmap of the final reconstruction. **d.** Representative 2D classes of the anti-Kappa IgG:DB3:DBPro1156_2 complex.



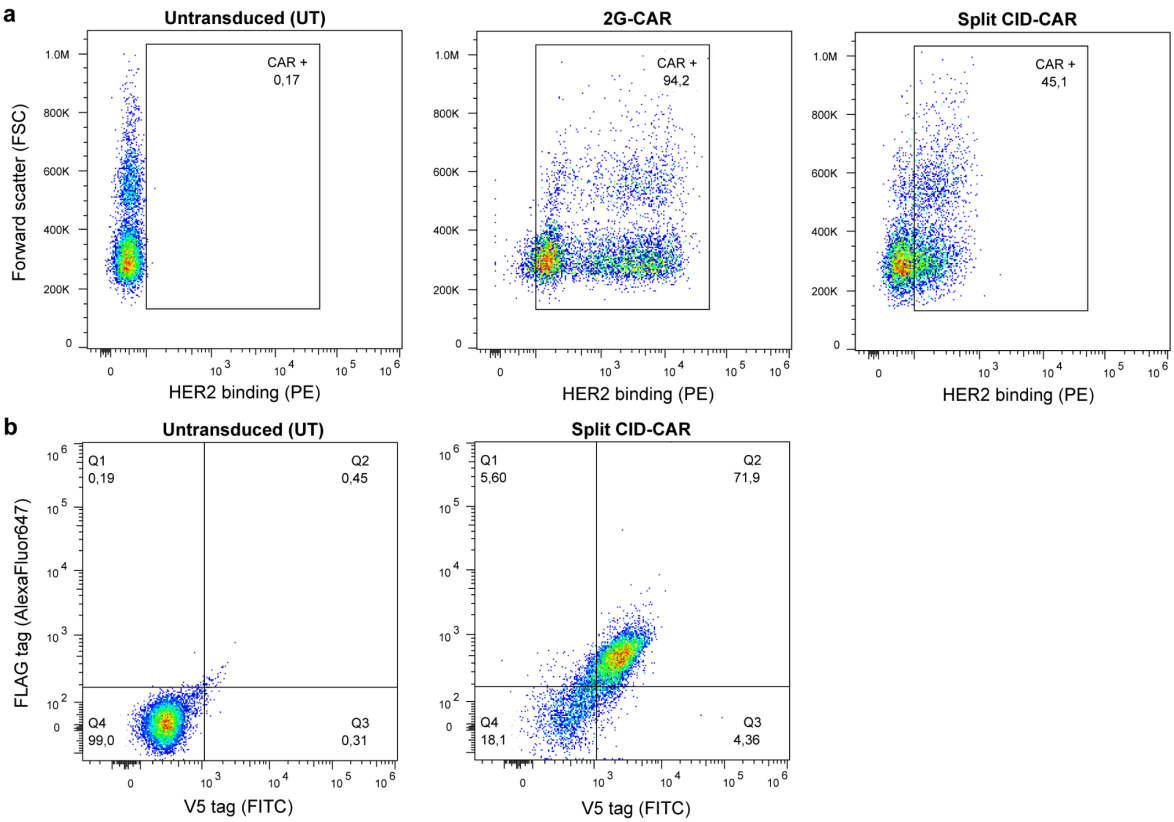
Supplementary Figure 13: Negative controls of cell-based drug-inducible systems. a-f. All CID systems tested in Fig. 5 and Extended Data Fig. 4 were compared to a negative control consisting of a point mutation at the center of the interface made by the computational binder. Point mutations were made for DB3:DPro1156_2 in cell-free expression system (a), Bcl2:DBVen1619_2 in the GEMS system (b), PDF1:DBAct551_1 in the intracellular split nanoLuc system (c), DB3:DBPro1156_2 in the extracellular split NanoLuc system (d), Bcl2:DBVen1619_2 in the intracellular split NanoLuc system (e) and PDF1:DBAct553_1 in the cell-free expression system (f). All measurements were performed in the absence (red) or presence (green) of the respective small molecule. Two-way ANOVA with Tukey's multiple comparison test, $p < 0.0001$ (****), non-significant (ns). Barplots are presented as mean ± standard deviations. Data points are derived from three technical replicates.



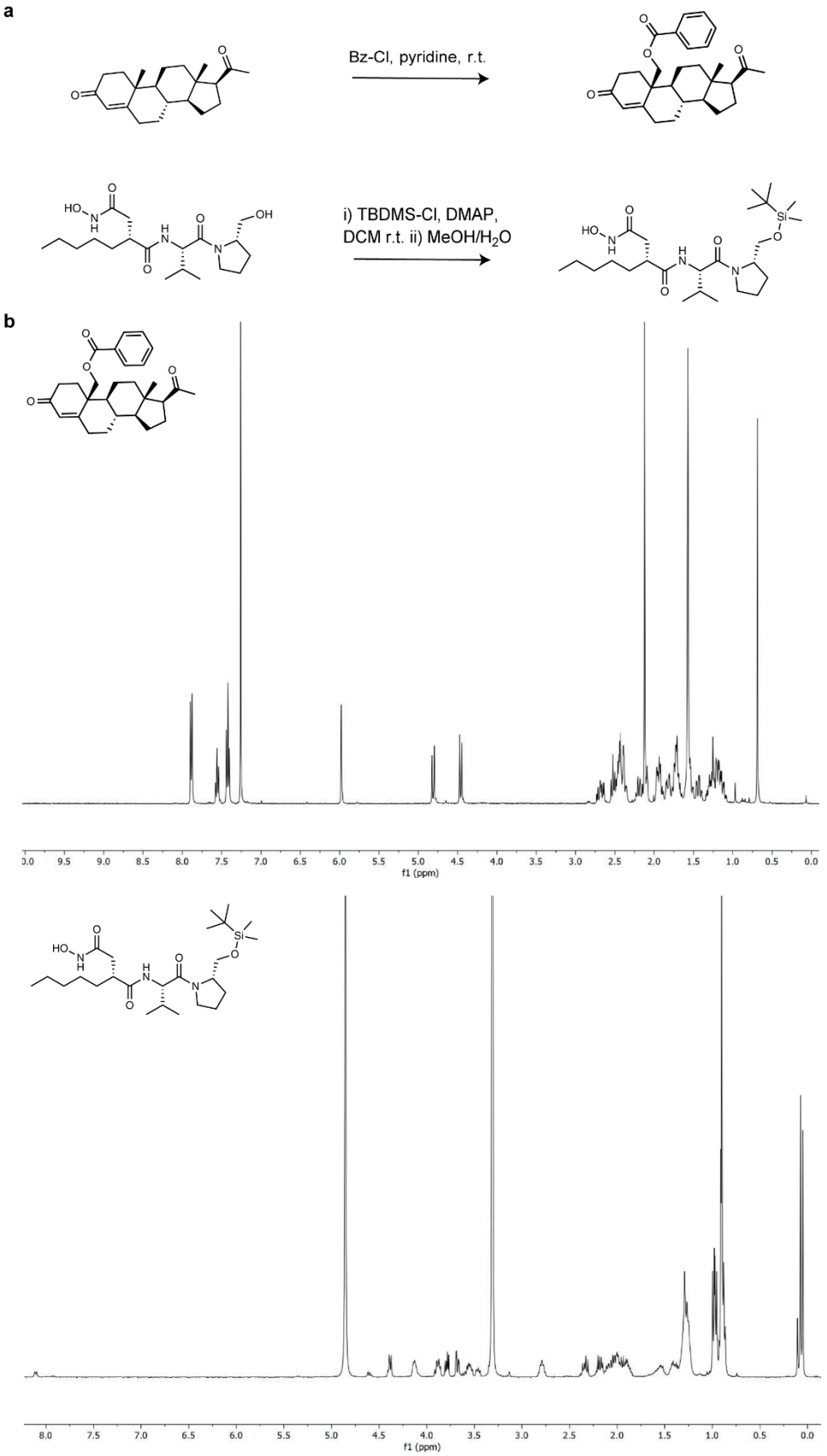
Supplementary Figure 14: CAR-T cell killing efficiency at different inducer concentrations. Cell counts at different concentrations of Venetoclax (0, 10, 50, 100 and 500 nM). Tumor cell counts at different time points are normalized to the number of live target cells in each well at t = 0 h. HER2-MC38 target cells were cultured alone (gray) or with untransduced primary murine T cell (UT, pink) or primary murine T cell transduced with a classical second generation αHER2 CAR (2G-CAR, blue) or inducible αHER2 CAR (Split CID-CAR, green) for 48 h. Tumor cell counts were measured using the IncuCyte® live-cell imaging system. Data represent biological replicates (n = 3) with mean and standard deviations.



Supplementary Figure 15: Analysis of the target interface and computational designs of the unsuccessful BRD4:JQ1 example. a. Molecular docking of JQ1 to BRD4 protein (gray) performed with Gnina. Original pose of JQ1 (PDB ID: 3MXF) used for design is coloured in green and molecular docking in pink. **b-c.** Computational metrics of the selected designs for each ligand-protein complex in the original pool (b) and after optimization with LigandMPNN (c). Computed binding energy (ddG, left), number of interface hydrogen bonds (middle) and buried surface area (SASA, right) were measured. Boxes represent the quartiles and whiskers show data points within 1.5x the interquartile range. Data outside this range are shown as flier points.



Supplementary Figure 16: CAR transduction efficiency in primary murine T cells. a. Transduction efficiency of the classical αHER2-CAR (2G-CAR) and the drug-inducible αHER2-CAR (split CID-CAR). Transduction efficiency was assessed by flow cytometry by measuring binding of a biotinylated HER2 protein labeled with PE-conjugated streptavidin. **b.** Flow cytometry plots showing labeling of both chains of the split CID-CAR. The chain containing FLAG-tagged Bcl2 and αHER2 were labeled with an A647-conjugated anti-FLAG antibody and the chain containing V5-tagged DBVen1619 was labeled with a FITC-conjugated anti-V5 antibody.



Supplementary Figure 17: Chemical synthesis and ^{1}H NMR spectra validation. a. Chemical synthesis reaction of 19-O-Benzoyl-Progesterone (OBz-Progesterone, top) and Tertbutyldimethylsilyl-Actinonin (TBDMS-Actinonin, bottom). **b.** ^{1}H NMR spectra of OBz-Progesterone (top) and TBDMS-Actinonin (bottom).

Supplementary Table 1: Metrics and cutoffs for binder design with MaSIF-seed.

Target	Motif	Site	Interface cutoff	NN score cutoff	Descriptor distance cutoff	#seeds	#selected seeds	#designs (#grafted seeds)	#select designs (#seeds)	Total design
Bcl2: Ven	S	1	0.65	0.9	2.0	1743	78	28396 (69)	1456 (67)	1995
		2	0.65	0.87	2.2	1048	33	7485 (29)	464 (27)	
		3	0.6	0.85	2.3	1012	11	1073 (8)	75 (8)	
DB3: Pro	H	1	0.75	0.9	1.8	995	49	160488 (49)	975 (44)	1998
		2	0.65	0.9	2.1	1046	36	147940 (36)	548 (34)	
	S	1	0.8	0.9	1.7	1775	98	10097 (39)	475 (37)	
PDF1: Act	H	1	0.65	0.87	2.2	1272	74	56813 (67)	1447 (66)	1997
	S	1	0.65	0.85	2.3	1373	98	3711 (56)	550 (55)	

Supplementary Table 2: Deep sequencing analysis of FACS-enriched populations. For each design, the target protein name and the topology and origin of the recipient scaffold is indicated. The log-enrichment consists of the logarithm of the number of reads in the binding population normalized to the number of reads in the non-binding (i.e. a log-enrichment of 1 means 10-fold enrichment in the binding vs non-binding populations). 3H: Three helical bundles. PDB: Protein Data Bank.

Round 1 (Original pipeline)							
Target	**Design**	**Scaffold type**	**Topology**	**Seed**	**Binding counts**	**Non-binding counts**	**Log-Enrichment**
Bcl2	DBVen1619	Beta grasp	EEHEE	S52	459449	15561	1.47
DB3	DBPro1156	3H (Set 1)	HHH	H26	134116	652	2.31
PDF1	DBAct553	3H (Set 1)	HHH	H49	35195	557	1.80
Round 2 (Scaffold optimization with LigandMPNN)							
Target	**Design**	**Scaffold type**	**Topology**	**Seed**	**Binding counts**	**Non-binding counts**	**Log-Enrichment**
Bcl2	DBVen420	Beta grasp	EEHEE	S52	404680	23634	1.23
PDF1	DBAct3	PDB	HH	H49	1039	6	2.24
	DBAct108	3H (Set 1)	HHH	H66	66148	431	2.19
	DBAct129	3H (Set 1)	HHH	H59	619	1	2.79
	DBAct169	3H (Set 1)	HHH	H49	15824	210	1.88
	DBAct175	3H (Set 1)	HHH	H59	21067	233	1.96
	DBAct253	Coil	HHHH	H59	11451	14	2.91
	DBAct281	3H (Set 2)	HHH	H53	6438	33	2.29
	DBAct327	3H (Set 2)	HHH	H56	17617	226	1.89
	DBAct355	3H (Set 2)	HHH	H57	5983	65	1.96
	DBAct361	3H (Set 2)	HHH	H74	13821	523	1.42
	DBAct379	3H (Set 2)	HHH	H57	67012	450	2.17
	DBAct389	3H (Set 2)	HHH	H7	79929	454	2.25

Supplementary Table 3: Computational analysis of ligand contributions. Summary of different metrics in absence or presence of ligands. The buried solvent-accessible surface area (SASA) of the protein-ligand target complex, the percentage of ligand contribution to this buried SASA, computed binding energy (ddG) in Rosetta Energy Unit (R.E.U.) and the number of atoms in contact with the target complex have been measured. Atom contacts were calculated based on the Van der Waals radii (r_{VdW} + 0.2 Å tolerance) of each pair of atoms. N/A: Not applicable.

Design	Ligand	Target buried SASA [Å^2]	Ligand contribution	ddG [R.E.U]	ddG shift	Atom contact [Counts]	Atom cont. shift
DBVen1619_1	-	N/A	N/A	-25.94	-17.04%	94	+7
	+	672	9.98%	-30.36		101	
DBPro1156_1	-	N/A	N/A	-28.59	-18.46%	117	+7
	+	702	10.57%	-33.87		124	
DBAct553_1	-	N/A	N/A	-32.18	-27.72%	123	+26
	+	886	12.13%	-41.10		149	

Supplementary Table 4: Docking benchmark complexes. List of 14 protein-ligand complexes and 200 decoys used in the binding partner recovery experiment. Search parameters: interface cutoff = 0.0; NN score cutoff = 0.8; descriptor distance cutoff = 3.0; #sites = 3; selection radius = 10Å.

Protein-protein-ligand complex			
PDB ID	**Protein 1**	**Protein 2**	**Ligand**
1A7X	FKBP12 (chain A)	FKBP12 (chain B)	BENZYL-CARBAMIC ACID [8-DEETHYL-ASCOMYCIN-8-YL]ETHYL ESTER (FKA)
1S9D	ADP-Ribosylation Factor 1 (chain A)	Arno (chain E)	1,6,7,8,9,11A,12,13,14,14A-DECAHYDRO-1,13-DIHYDROXY-6-METHYL-4H-CYCLOPENT[F]OXACYCLOTRIDECIN-4-ONE (AFB)
1TCO	SERINE/THREONINE PHOSPHATASE B2 (chains A and B)	FK506-BINDING PROTEIN (chain C)	8-DEETHYL-8-[BUT-3-ENYL]-ASCOMYCIN (FK5)
3QEL	NMDA glutamate receptor subunit (chain A)	Glutamate [NMDA] receptor subunit epsilon-2 (chain B)	4-[(1R,2S)-2-(4-benzylpiperidin-1-yl)-1-hydroxypropyl]phenol (QEL)
4DRI	Peptidyl-prolyl cis-trans isomerase FKBP5 (chain A)	Serine/threonine-protein kinase mTOR (chain B)	RAPAMYCIN IMMUNOSUPPRESSANT DRUG (RAP)
4MDK	Ubiquitin-conjugating enzyme E2 R1 (chain A)	Ubiquitin (chain E)	4,5-dideoxy-5-(3',5'-dichlorobiphenyl-4-yl)-4-[(methoxyacetyl)amino]-L-arabinonic acid (U94)
6ENG	DNA gyrase subunit B (chain A)	DNA gyrase subunit B (chain B)	Coumermycin A1 (BHW)
6H0F	Protein cereblon (chain B)	DNA-binding protein Ikaros (chain C)	S-Pomalidomide (Y70)
6N4N	NS3 protease (chain A)	Rosetta-designed danoprevir/NS3a complex reader 2 (chain F)	(2R,6S,12Z,13aS,14aR,16aS)-6-[(tert-butoxycarbonyl)amino]-14a-[(cyclopropylsulfonyl)carbamoyl]-5,16-dioxo-1,2,3,5,6,7,8,9,10,11,13a,14,14a,15,16,16a-hexadecahydrocyclopropa[e]pyrrolo[1,2-a][1,4]diazacyclopentadecin-2-yl 4-fluoro-2H-isoindole-2-carboxylate (TSV)
6OB5	Maltodextrin-binding	Ankyrin Repeat Domain	FARNESYL

	protein (chain B)	(AR), S3-2D variant (chain D)	DIPHOSPHATE (FPP)
6QTL	VHH (chain A)	VHH (chain C)	CAFFEINE (CFF)
6SJ7	DDB1- and CUL4-associated factor 15 (chain A)	RNA binding protein 39 (chain C)	N~1~-(3-chloro-1H-indol-7-yl)benzene-1,4-disulfonamide (EF6)
7DC8	Switch Ab Fab light & heavy chain (chains A and B)	Interleukin-6 receptor subunit alpha (chains C and F)	ADENOSINE-5'-TRIPHOSPHATE (ATP)
7TE8	DB21 (chain A)	CA14 (chain C)	cannabidiol (P0T)

Supplementary Table 5: Target protein and binder sequences.

Design	Sequence	Mutations from native
DBVen1619_1	KYMLVVKGPNVTIFRWVDSSSEAETLARKIAKKLGLEVKSVEKKGNAVRVEIG	
DBVen1619_2	QYLLVVKGPNVTKFRWVDSSSEAETLARKIAKKLGLEVKSVEKKGNAVRVEIG	K1Q, M3L, I13K
DBPro1156_1	DEKAKTAETLIYQLFSKAMQQSDPNEAEKLLKKAEELAKKANDPRLEQVVRQHQVVV RFLV	
DBPro1156_2	DEKAKTAETLIWQLFGKAMQQSDPNEAEKLLKKAEELAKKANDPRLEQVVRQHQVVV RFLV	Y12W, S16G
DBAct553_1	DYIRELRAALILLALKKQHAEDPDAQRVADELMKKLFDAAHRNDKDKVKKVVEEAKKV VSTY	
DBAct553_2	DYIRELNRALILLALKKQHAEDPDAQRVADELMKKLFDAAHRNDKDKVKKVVEEAKKV VSTY	R7N, A8R
DB3_H	QIQLVQSGPELKKPGETVKISCKASGYAFTNYGVNWVKEAPGKELKWMGWINIYTGE PTYVDDFKGRFAFSLETSASTAYLEINNLKNEDTATYFCTRGDYVNWYFDVWGAGTT VTVSSAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHT FPAVLQSDLYTLSSSVTVPSSPRPSETVTCNVAHPASSTKVDKKIVPR	
DB3_L	DVVMTQIPLSLPVNLGDQASISCRSSQSLIHSNGNTYLHWYLQKPGQSPKLLMYKVSN RFYGVPDRFSGSGSGTDFTLKISRVEAEDLGIYFCSQSSHVPPTFGGGTKLEIKRADA APTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDS KDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNR	
DB3_H (Chimeric)	QIQLVQSGPELKKPGETVKISCKASGYAFTNYGVNWVKEAPGKELKWMGWINIYTGE PTYVDDFKGRFAFSLETSASTAYLEINNLKNEDTATYFCTRGDYVNWYFDVWGAGTT VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHT FPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHT	
DB3_L (Chimeric)	DVVMTQIPLSLPVSLGEQASISCRSSQSLIHSNGNTYLHWYLQKPGQSPKLLMYKVSN RFYGVPDRFSGSGSGTDFTLKISRVEAEDLGIYFCSQSSHVPPTFGGGTKLEIKRTVA APSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	
Anti-kappa_H	EVKLLESGGGLVQPGRSLRLSCIASGFDFSGYWMTWVRQAPGKGLEWIGDINPDSST INSTPSLKDKVIISRDNAKNTLFLQMSKVRSEDTALYYCAQRGNYVPFPYWGQGTLVT VSAAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFP AVLQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRDCGCK	
Anti-kappa_L	SIVMTQTPKFLFVSAGDRVTITCKASQSVSNDVEWYQQKPGQSPKLMIYFASKRYNG VPDRFTGSGFGTEFTFTISTVQAEDLAVYFCQQDYSSPWTFGGGTKLEIKRADAAPTV SIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDST YSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRGEC	
Bcl2	MAHAGRTGYDNREIVMKYIHYKLSQRGYEWDAGDDAEENRTEAPEGTESEVVHRAL RDAGDDFERRYRRDFAEMSSQLHLTPDTARQRFETVVEELFRDGVNWGRIVAFFEF GGVMCVESVNREMSPLVDNIAEWMTEYLNRHLHTWIQDNGGWDAFVELYGPSMR	
PDF1	AILNILEFPDPRLRTIAKPVEVVDDAVRQLIDDMFETMYEAPGIGLAATQVNVHKRIVVM DLSEDKSEPRVFINPEFEPLTEDMDQYQEGCLSVPGFYENVDRPQKVRIKALDRDGN PFEEVAEGLLAVCIQHECDHLNGKLFVDYLSTLKRDRIRKKLEKQHRQQA	
DBAct3	GEIEEVEEKIEELEKKLKEETKKLEEAKKAVEKNANKENIAKENEARIKVALLEEKLKEL RARLAELKAAA	
DBAct108	SQARRDRLVAEAQTLIDTYADAEIRRAARRVILRAAVAADRGDDAALARVEAELAAVK ARAA	
DBAct129	DSLLYENQVRIRLALLLEKEKDAANREKIRALQEKALAAVKAGDRATVERLVAEARRIA EKSNA	
DBAct169	NEKIKAINRIRIKFALAAEKHDSERMLAVAEEARKLAEEYKDEEIRKLADALAKLAKLLA	

DBAct175	SAAKQEEDEKRVKRLAAEIAHATDEADRERLRAEALEIIESNKDQKHRDRLRILFALTLY KAS	
DBAct253	DDEAILKEVEKIAREIAKLHGYEKAADVIVSKDAEKLVELYKTDKQKLVDTIRVRAALLAY KA	
DBAct281	DERIERLRELAKEAAKALGLSEEEQEEYAERIYQKALTLDEVELRILEAAVQYKLEKAKA	
DBAct327	SELLFWGNRIRILYATRLDRGSLSAAEIVAELHRVLDGRHAAGEISDEVYAAVRQQIQD LEAQA	
DBAct335	DIVERAKEVIELAKKKGDLKAINEARIVIALTEEGKATEEGAQKTIDKLFERLA	
DBAct361	SVVALLDEAIRLAEKKGREDLVKELRALKAALEAGEISEEKAEATARVVKARLEY	
DBAct379	DELARLLAEGRAIAEAKGDLKALNELRIIEALVKEGRADREKATERAREIVERLSG	
DBAct389	SKLEELVERAAELYKAGRLTLSQAKNWFRVRAKHEGVSEEEIEEGKERITELALA	
DBVen420	AVLLVVRGDDTFRLHFFDTAEEAIAAAHEIAAELGYEVREVELRGNAVVVRIG	