nature portfolio

Corresponding author(s): Bruno E. Correia

Last updated by author(s): 27/10/24

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	XDS v Jan 10, 2022 (BUILT=20220220) and autoPROC (GlobalPhasing, version 20230222) package for processing of crystallographic data; Automation program EPU (ThermoFischer Sci., v2.12.1) for cryoEM data collection; Biacore 8K control software (Cytiva, v4.0.8.19879) for measuring surface plasmon resonance; LE-SH800SZFCPL Cell Sorter software (Sony, v2.1.5) for sorting and collecting FACS data; Kaluza for Galios (Beckman Coulter, v1.1.20388.18228) for collecting flowcytometry data; Chromeleon (ThermoFischer Sci, v7.2.10) and Astra (Wyatt Tech., v8.0.2.5) for performing and collecting data of SEC-MALS; GatorOne (Gator Bio, v.2.7.3.0728) for biolayer interferometry; Python (v3.6.5) as well as APBS (v1.5), MSMS (v2.6.1), RDKit (v2020.09.1.0) to run our MaSIF-neosurf pipeline.
Data analysis	Coot (v0.9.5) for structure building; Prism (GraphPad, v10) for graphs generation; FlowJo (BD Bioscience, v10.8.1) for flowcytometry analysis; PyMol (Schrödinger, v2.4) and ChimeraX (UCSF, v1.3) for protein vizualization and structural graphic generation; MolProbity (v4.5.1) for structure evaluation; Phenix (v1.20.1-4487) for molecular replacement phasing and structure refinement; cryoSPARC (v4.4.1) for cryoEM structure determination, Biacore Insight Evaluation Software (Cytiva, v4.0.8.19879) for evaluating surface plasmon resonance measurements, Rosetta modelling suite (v3.13) for protein design and analysis; Custom scripts to analyze protein modeling data are available on Github (https://github.com/LPDI-EPFL/masif_seed)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Crystal structure of DBAct553_2 in complex with Actinonin-bound PDF1 has been deposited at the PDB under the accession code 8S1X (DOI: https://doi.org/10.2210/pdb8S1X/pdb). The refined atomic models and corresponding cryoEM maps of DBPro1156_2 in complex with progesterone-bound DB3 Fab were deposited under PDB: 9FKD (https://doi.org/10.2210/pdb9FKD/pdb) and EMDB: 50522. The scaffold database generated for grafting the seed provided by MaSIF-neosurf is partly available at Zenodo (https://zenodo.org/records/7643697#.Y-z533ZKhaQ) and partly on Github (https://github.com/strauchlab/DBP and https://github.com/strauchlab/scaffold_design/). Data used to generate the figures and supplementary materials, as well as the relevant plasmid maps, were deposited on Zenodo (http://doi.org/10.5281/zenodo.13737922).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the protein designs, ~2000 protein sequences for each target (Bcl2:venetoclax, DB3:Progesterone & PDF1:Actinonin) were filtered computationally and selected by yeast display. Based on experimental success rate observed in previous findings (doi: 10.1038/s41586-023-05993-x), 2000 designs were considered as appropriate sample size.
Data exclusions	There is no data exclusion in this study.
Replication	Each binding candidate detected by yeast display and/or deep-sequencing showed reproducible results and were also tested with negative controls (unrelated protein ligand or unlabeled yeast) and point mutants. Result reproducibility was also confirmed with alternative methods (e.g. SPR)
Randomization	Randomization does not apply in our study as oligos encoding the one-shot designs were pooled together for a high-troughput screening (one oligopool per target protein) and then sorted by yeast display using flow cytometry. The goal was to test binding capacity of each individual design.
Blinding	Researchers were not blinded as this does not apply to our study (no bias expected from cell sorting)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

1) Anti-HA tag, FITC - Ref : A190-138F - Manufacturer : Bethyl - Clone : Unknown - Dilution: 1:100
2) Anti-human IgG, PE - Ref : 12-4998-82 - Manufacturer : Invitrogen - Clone : Polyclonal - Dilution: 1:100
3) Biotinylated Human Her-2 - Ref: HE2-H822R - AcroBiosystems - Clone : N/A - Dilution: 1:100
4) Streptavidin, PE - Ref: 12-4317-87 - Invitrogen - Clone: N/A - Dilution: 1:100
5) Anti-V5 tag, FITC - Ref: GTX21209 - GeneTex - Clone: Polyclonal - Dilution: 1:100
6) Anti-FLAG tag, AlexaFluor647 - Ref: MA1-142-A647 - Thermofisher - Clone: L5 - Dilution: 1:100

Validation

For the commercially available antibodies, no validation reports were provided other than publications citing the products or examples on manufacturer's webpage :
1) 10.21037/atm.2020.03.74
2) https://www.thermofisher.com/antibody/product/Goat-anti-Human-IgG-Fc-Secondary-Antibody-Polyclonal/12-4998-82
3) 10.2967/jnumed.118.223636
4) 10.1016/j.xpro.2020.100129
5) https://www.genetex.com/Product/Detail/V5-tag-antibody-FITC/GTX21209
6) 10.1038/s41589-021-00832-4

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

Expi293 were purchased from ThermoFischer Sci. (Ref: A14635)
HEK293T were purchased from Invitrogen (Ref: R70007)
Phoenix-ECO were purchased from ATCC (Ref: CRL-3214)
MC38-HER2 were generated as described previously (doi: 10.1038/s41590-021-00940-2) and provided by Prof. Li Tang (EPFL)
Primary murine T cell were isolated from C57BL/6 mice

Authentication

Expi293 were authenticated by the provider using short tandem repeat (STR) genotyping
HEK293T were authenticated by the provider using short tandem repeat (STR) genotyping
Phoenix-ECO were authenticated by the provider using short tandem repeat (STR) genotyping
MC38-HER2 was not authenticated in this study.
Primary murine T cells were specifically isolated using a pan T cell kit (Miltenyi Biotec; Ref: 130-095-130)

Mycoplasma contamination

Expi293: Cells were tested negative for mycoplasma contamination by the manufacturer (qPCR) and no additional test was performed as the cells were used for protein expression.
HEK293T : Cells were tested negative for mycoplasma (qPCR)
Phoenix-ECO : Cells were tested negative for mycoplasma (MycoAlert® Mycoplasma Detection Kit - Ref: LT07-318)
MC38-HER2 : Cells were tested negative for mycoplasma (MycoAlert® Mycoplasma Detection Kit - Ref: LT07-318)
Primary murine T cell : Cells were tested negative for mycoplasma (MycoAlert® Mycoplasma Detection Kit - Ref: LT07-318)

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified lines were used in this study

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Yeasts cells (EBY-100) were labeled with the protein target and then with an anti-HA (FITC) for the display signal and an anti-Fc (PE) for binding to the protein target. Cells were washed with PBS supplemented with 0.1% BSA. See methods for more details
Instrument	Sony SH800 (Sorting) and Beckman Coulter Gallios (Analysis)
Software	Sony LE-SH800SZFCPL Cell Sorter software (v2.1.5), Kaluza for Gallios (v1.1.20388.18228) and FlowJo (v10.8.1)
Cell population abundance	Transformed yeasts underwent two rounds of sorting for target binding and amplifications in culture before being sequenced for the isolation of single binding clones. Single clone candidates (re-transformation of naive yeasts with pure DNA) were then individually tested with controls.
Gating strategy	Yeast cells without target labeling served as a negative control and yeasts showing binding signal above this negative threshold were collected (See supplementary fig. 6).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.