Assessing and enhancing foldability in designed proteins

Dina Listov[1], Rosalie Lipsh-Sokolik[1], Che Yang[2], Bruno E Correia[2], Sarel J Fleishman[1]

[1] Department of Biomolecular Sciences, Weizmann Institute of Science, 7610001, Rehovot, Israel

[2] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Abstract

Recent advances in protein-design methodology have led to a dramatic increase in its reliability and scale. With these advances, dozens and even thousands of designed proteins are automatically generated and screened. Nevertheless, the success rate, particularly in design of functional proteins, is low and fundamental goals such as reliable *de novo* design of efficient enzymes remain beyond reach. Experimental analyses have consistently indicated that a major cause of design failure is inaccuracy and misfolding relative to the design model. To address this challenge, we describe complementary methods to diagnose and ameliorate suboptimal regions in designed proteins: first, we develop a Rosetta atomistic computational mutation scanning approach to detect energetically suboptimal positions in designs; second, we demonstrate that the AlphaFold2 *ab initio* structure prediction method flags regions that may misfold in designed enzymes and binders; and third, we focus FuncLib design calculations on suboptimal positions in a previously designed low-efficiency enzyme, thereby improving its catalytic efficiency by 330 fold. Thus, foldability analysis and enhancement may dramatically increase the success rate in design of functional proteins.

Keywords: protein folding, protein stability, computational design, Rosetta, AlphaFold, FuncLib

Introduction

Over the past decade, protein design methodology has made remarkable progress. New methods enable the routine design of new folds [1–4], assemblies [5,6], and new or improved functions [7–13]. Despite these achievements, however, only a small fraction of experimentally tested designs are functional [8,14,15], and fundamental protein design objectives still lie beyond our reach. For example, reliable *de novo* enzyme design remains an unsolved challenge despite decades of research, and all enzymes designed to date exhibit low efficiency [15–17]. Previous analyses of successful and failed enzyme designs indicated that inaccuracies in the design of catalytic constellations are partly to blame for failures [18–21]. Furthermore, observations across many different folds, binders, and enzyme design studies indicate misfolding relative to the design model as likely to be the most general and critical

problem in failed designs [2,22–25]. Indeed, the significance of misfolding has served as the primary motivation for studies of the design principles of protein folds [2,26,27] and for developing general methods to improve designed proteins' foldability and stability [11,13,27,28].

The current study is motivated by our ongoing efforts to develop a reliable strategy for *de novo* enzyme design. This strategy is based on recent methods developed in our lab to design new backbones through the modular assembly of large (60-150 amino acid) backbone fragments of natural enzymes [25,29,30]. Modular assembly and design was successful in assembling elements from existing protein families, generating accurate antibodies [27] and ultrahigh specificity binders [24] and enzymes [25] with as many as 100 mutations from any natural protein. Nevertheless, we have so far only been able to generate *de novo* enzymes that exhibit activities in the same low range as those designed more than a decade ago. Experimental analyses of 124 of these designs showed that the vast majority (80%) of the failed designs did not exhibit cooperative folding as indicated by thermal denaturation analysis. By contrast, all active designs exhibited clear temperature-dependent denaturation with melting temperatures exceeding 75℃. The very high correspondence between foldability and activity in our *de novo* design study prompts us to examine the sources of low foldability in current atomistic design methods by developing strategies for detecting and ameliorating suboptimality in designed proteins.

Results

Energy-based suboptimality detection

We develop an automated computational approach to identify energetically suboptimal positions in protein designs. Our approach, which we call pSUFER for protein Strain, Unsatisfactoriness, and Frustration findER, starts by relaxing the input structure using Rosetta atomistic modeling. It then models all single-point mutations at every position, iterating sidechain packing and whole-protein minimization, and computes the change in system energy ($\Delta\Delta G$) relative to parental protein. These calculations use the Rosetta all-atom energy function 2015 (ref2015) which is dominated by van der Waals packing, hydrogen bonding, electrostatics and implicit solvation [31]. Finally, pSUFER flags positions that exhibit more than a certain number of mutations (typically 5) that lower the native-state energy ($\Delta\Delta G < 0$) as potentially suboptimal. Customizable scripts for automatically relaxing a protein structure,

computing suboptimal positions and visualizing them in PyMOL are available in https://github.com/Fleishman-Lab/pSUFER. .

Our approach is similar in principle to strategies for computing local frustration in proteins [32,33]. Methods for analyzing local frustration detect positions or pairs of positions that exhibit high energies relative to a computed ensemble of structures or decoys. Such frustrated positions are often associated with the protein's activity since active-site positions, regions that are involved in allosteric communication and conformational change, are evolutionarily selected for their role in activity rather than for improving stability [34]. Our approach is different from conventional frustration analyses in quantifying the number of possible identities that may lower the protein energy by single-point mutations. This strategy reflects the protein-engineering goal that all positions outside the protein's functional site(s) should be relaxed to increase the likelihood of accurate and stable folding into the native state.

Orders of magnitude improvement in catalytic efficiency of a failed design

To test pSUFER, we analyzed a set of enzymes and binders designed and experimentally characterized over the past three years. To be included in this set, we required that the same design method produced both proteins that accurately folded into the design conception according to X-ray crystallography and ones that misfolded. Despite the availability of experimental structures in these cases, we applied pSUFER to the design models to verify that the method could uncover flaws in the design conception without recourse to experimental data. The designs included enzymes [25] and binders [24] generated through modular backbone assembly and design as well as binders in which an immunogenic epitope was incorporated in a *de novo* designed scaffold protein [35]. Therefore, these designs encompass a range of contemporary applications of protein design methodology.

The first design pair we examine was generated by modular assembly and design of glycoside hydrolase 10 (GH10) xylanases [25]. In this design set, 21 out of 43 designs exhibited detectable xylanase activity and only two were as active as natural enzymes from the GH10 family. From that study, we chose for pSUFER analysis two designs that exhibited the highest and lowest levels of detectable activity and for which we have crystallographic data, xyl3.1 and xyl8.3 (k_{cat}/K_M 9,417 and 0.61 $M^{-1}s^{-1}$ respectively). Both designs were highly mutated relative to natural GH10 enzymes, exhibiting 105 and 130 mutations, respectively (out of approximately 350 amino acids) and exhibited apparent thermal denaturation temperatures > 55 ℃. The crystal structure of xyl3.1 showed remarkable accuracy relative to the design conception with 0.7 Å root mean square deviation (rmsd) across the entire protein and <1 Å

all-atom rmsd of active-site residues. By contrast, although xyl8.3 was globally accurate (rmsd = 0.9 Å) and core catalytic groups aligned well between the design model and the experimental structure, the experimental structure exhibited significant missing density in two adjacent loops near the active-site pocket, indicative of local misfolding. Visual inspection of the two designs following their structure determination failed to suggest significant flaws in xyl8.3 that might explain design inaccuracy.

Applied to the two xylanase design models, pSUFER flags a similar number of positions. As expected, most of the flagged positions are either in the active-site pocket or in solvent-accessible positions (Fig. 1A,B). Conspicuously, however, in xyl8.3, pSUFER flags position Lys306, which is buried at the stem of one of the loops that does not exhibit electron density. Furthermore, this charged position does not form stabilizing interactions with counter charges, providing a potential mechanistic explanation for the observed misfolding.

We hypothesized that eliminating the strain observed around position Lys306 (and the other positions flagged by pSUFER) might improve the enzyme's catalytic efficiency. To test this hypothesis, we applied FuncLib design [12] to the five positions pSUFER flagged outside the active-site pocket in xyl8.3. FuncLib starts by using phylogenetic information and atomistic design calculations to rule out mutations that may be destabilizing to the native-state structure. It then enumerates all combinations of allowed mutations at the designed positions, ranks them by energy and suggests low-energy combinations of mutations. An advantage of the FuncLib methodology over typical stochastic combinatorial design algorithms is that FuncLib relaxes each combination of mutations by whole-protein minimization. Therefore, FuncLib may find stabilizing mutations, including radical small-to-large mutations, in the core of the protein that may elude other atomistic design methods. Applied to the five suboptimal positions flagged by pSUFER, one of the top FuncLib designs, which we called xyl8.3fix, improved system energy by 10 Rosetta energy units (R.e.u) by mutating four positions (Thr1Asn, Asn4Glu, Gly139Asp, and Lys306Leu). Among these four, only the Lys306Leu is radical, whereas the others mutate solvent-exposed positions to increase surface polarity.

We expressed the xyl8.3fix design fused N-terminally to maltose-binding protein (MBP) in *E. coli* BL21 cells and purified the protein using an amylose column. We then tested the design's catalytic efficiency using the 4-nitrophenyl β-xylobioside (OPNPX2) chromogenic substrate. Remarkably, Michaelis-Menten analysis revealed that xyl8.3fix exhibits catalytic efficiency of 226 $M^{-1}s^{-1}$ (Fig. 1C), 330-fold greater than that of the parental xyl8.3. We also analyzed the thermal denaturation of the two designs, finding that both had similar apparent denaturation temperatures 57 and 59 ℃ for xyl8.3 and xyl8.3fix, respectively (Fig. S1). Therefore, we

concluded that the FuncLib design did not appreciably stabilize the enzyme relative to xyl8.3 but likely improved its ability to fold into the active form.

Significantly, in the modular assembly and design study that included xyl3.1 and xyl8.3 [25], we noted that assembling backbone fragments from more than a few natural templates led to low activity (xyl3.1 and 8.3 were based on 3 and 8 backbone fragments, respectively). The activity we observe for xyl8.3fix, however, would place it among the top three designs in that set despite comprising fragments from eight proteins. Thus, the pSUFER analysis may significantly improve the reliability and diversity of protein design methods.

To test whether a similar approach may improve computed system energies in other designs, we applied FuncLib to positions flagged by pSUFER in 62 *de novo* designed enzymes generated through modular backbone assembly and design. FuncLib introduced an average of five mutations and improved the energies by an average of 13 R.e.u in this set (Fig. 1D), similar to the values we obtained by applying this approach to xyl8.3. Out of these 62 fixed designs, 11 showed cooperative folding (representative data shown in Fig S2).

Suboptimality in designed binders

We next ask whether pSUFER may shed light on problems in designed binders. In the following, we examine binders generated through modular assembly and design and *de novo* designed binders. Using modular assembly and design, we previously started from a high-affinity binding pair to design a set of binders, some of which exhibited ultrahigh specificity relative to the parental pair and other designed pairs [24]. Our approach focused on the design of a new interfacial loop backbone in one of the binding partners by grafting loops from unrelated proteins and designing the sequence to form specific interactions with the binding partner. X-ray crystallographic analysis demonstrated that ultraspecific designs exhibited atomic accuracy in the designed loop relative to the model (for instance, des3). By contrast, one of the multispecific designs exhibited missing density in the designed loop (des4).

In each case, we apply pSUFER to the protein in which the loop backbone was designed in the absence of the binding partner. In each design, pSUFER flags a position within the designed binding loop (Fig. 2A and 2B). The flagged position in des4, Gly26, cannot be redesigned as Gly26's backbone comes into close contact with the binding partner. In des3, by contrast, the flagged position Asn27 can be designed to other identities which are less strained according to our models. In addition, in both designs pSUFER flags several solvent-accessible positions (Fig 2B). The results on des4 demonstrate that the pSUFER analysis may, in some

cases, indicate suboptimality that cannot be relieved without compromising the objectives of design of function. In such cases, the design may be triaged from experimental testing.

A final example is from *de novo* design of biosensors that detect and quantify epitope-specific antibodies. The design strategy uses a bottom-up approach: first defining the function-rendering motifs and then designing a protein fold to support the motifs [35]. One of the designs, named 4H.01 folds accurately at the binding site as determined by X-ray crystallography; however, this design shows substantial backbone deviations (2.9 Å) relative to the design and exhibits missing density in a loop (Fig 2C). pSUFER flags eleven positions in this design; for instance, Gln84, which is positioned in a cavity next to the loop that exhibits missing density, and Leu78. According to the design model, Gln84 is partly desolvated and does not form intimate polar contacts with the loop. Leu78 is forced into a strained sidechain conformation exhibiting Rosetta sidechain conformation energy approximately +3 energy units. This sidechain cannot pack into a favorable conformation due to steric overlaps in relaxed sidechain conformations. By contrast to this design, in the accurately crystallized design, only three positions are flagged, all of which are solvent-exposed (His1, Thr19, and Thr35, Fig2D).

Analyzing design foldability by AlphaFold2

Recently, methods that use deep learning have been very successful in *ab initio* structure prediction [36–39]. The most successful of these, AlphaFold2, has shown atomic-accuracy prediction in community blind-prediction experiments and other structure-prediction challenges [40]. Significantly for the purposes of assessing design foldability, AlphaFold2 provides a confidence score (per-residue local distance difference test; plDDT) for each position in the predicted model structure. We hypothesized that the plDDT scores and the rmsd between the AlphaFold2 predicted structure and the design model might predict design accuracy and foldability. AlphaFold2 depends on multiple-sequence alignments of homologs for accurate structure prediction so we did not analyze the *de novo* designed binders.

The AlphaFold2 analysis of the two accurately designed proteins (xyl3.1 and des3) corresponded very well to the design models (rmsd < 0.7 Å). Furthermore, except in the N and C-terminal tails, the plDDT scores were high in xyl3.1 and the majority of des3 (>90%). For des3, the backbone of the binding-surface loop was grafted from another protein, explaining why the plDDT scores are not as high in this region (>80%). Significantly, in the case of des4, the AlphaFold2 model deviates from the design in the region corresponding to

the missing density. Furthermore, the plDDT scores clearly depress around these regions in both des3 and xyl8.3 relative to the two other designs. We next analyzed the AlphaFold2 results for xyl8.3fix. The AlphaFold2 model structure recapitulated the design model with rmsd < 0.8Å, similar to xyl8.3. Remarkably, the plDDT scores in the region surrounding position 307, as well as in the loops that exhibit missing density according to the crystallographic analysis are significantly higher and equivalent to those observed for the remainder of the protein (>92%). This result suggests that the AlphaFold2 confidence scores are sensitive even to single-point mutations in designed proteins.

Finally, we generated AlphaFold2 models for 124 *de novo* designed enzymes for which we obtained experimental data. Remarkably, the models for the 26 designs that folded cooperatively exhibited low rmsd (the majority below 1.5 Å) and high average plDDT scores (>91%) relative to the remainder of the set (Fig. 3), (p value < 5% according to double sided Fisher's exact test). We concluded that the AlphaFold2 analysis can provide critical information on the likely accuracy and foldability of newly designed backbone structures, indicate regions that may misfold, and assess mutations that are designed to mitigate misfolding.

Discussion

Recent advances in protein design methodology extend its scope to the design of large proteins that are highly mutated relative to natural ones [25] and to sets comprising thousands of designs [3,14]. Nevertheless, the success rate is typically very low and some critical goals remain elusive. Thus, automated methods to assess design reliability may have a profound impact. Assessing a design's accuracy and foldability, however, remains challenging. As a general rule, design studies, particularly ones devoted to backbone design, often reveal significant deviations between experimental structures and design conceptions. When the designs are small (<100 amino acids), they may be subjected to atomistic forward-folding *ab initio* simulations to verify that the sequences favor the designed conformation over others [1–3,41]. In large proteins, however, accurate forward-folding simulations were until recently impossible. Our results demonstrate that the deep-learning based method AlphaFold2 provides critical information on design foldability even in large proteins that exhibit more than 100 mutations from any natural protein.

Furthermore, the pSUFER energy-based method can pinpoint specific positions that may be poorly designed. The functional consequences of poor design choices are strikingly demonstrated in our study: by redesigning four positions in xyl8.3 that were flagged by

pSUFER (out of 350 positions), the enzyme's catalytic efficiency rises by three orders of magnitude. These flagged positions are located outside the active-site pocket, demonstrating the significance of accurate and strain-free design throughout the protein. These observations, together with the improvement in the AlphaFold2 confidence scores for this variant, implicate foldability as the cause of improvement in activity. It is also striking that a handful of poor design choices in a large protein led almost to complete dysfunction. The ability to automatically detect poor design choices and fix them may free protein designers to introduce significantly more radical changes than previously and increase the success of backbone design in large enzymes and binders.

Acknowledgements

We thank Olga Khersonsky for help in xylanase activity screens and Ravit Netzer for advising in pSUFFER binder results analysis. Research in the Fleishman lab was supported by a Consolidator Grant from the European Research Council (815379) and by an Individual Grant from the Israel Science Foundation (1844/19). R.L-S is supported by a fellowship from the Arianne de Rothschild Women Doctoral Program. The Correia lab was supported by a Starting Grant from the European Research Council (716058) and a project grant from the Swiss National Science Foundation (310030_163139).

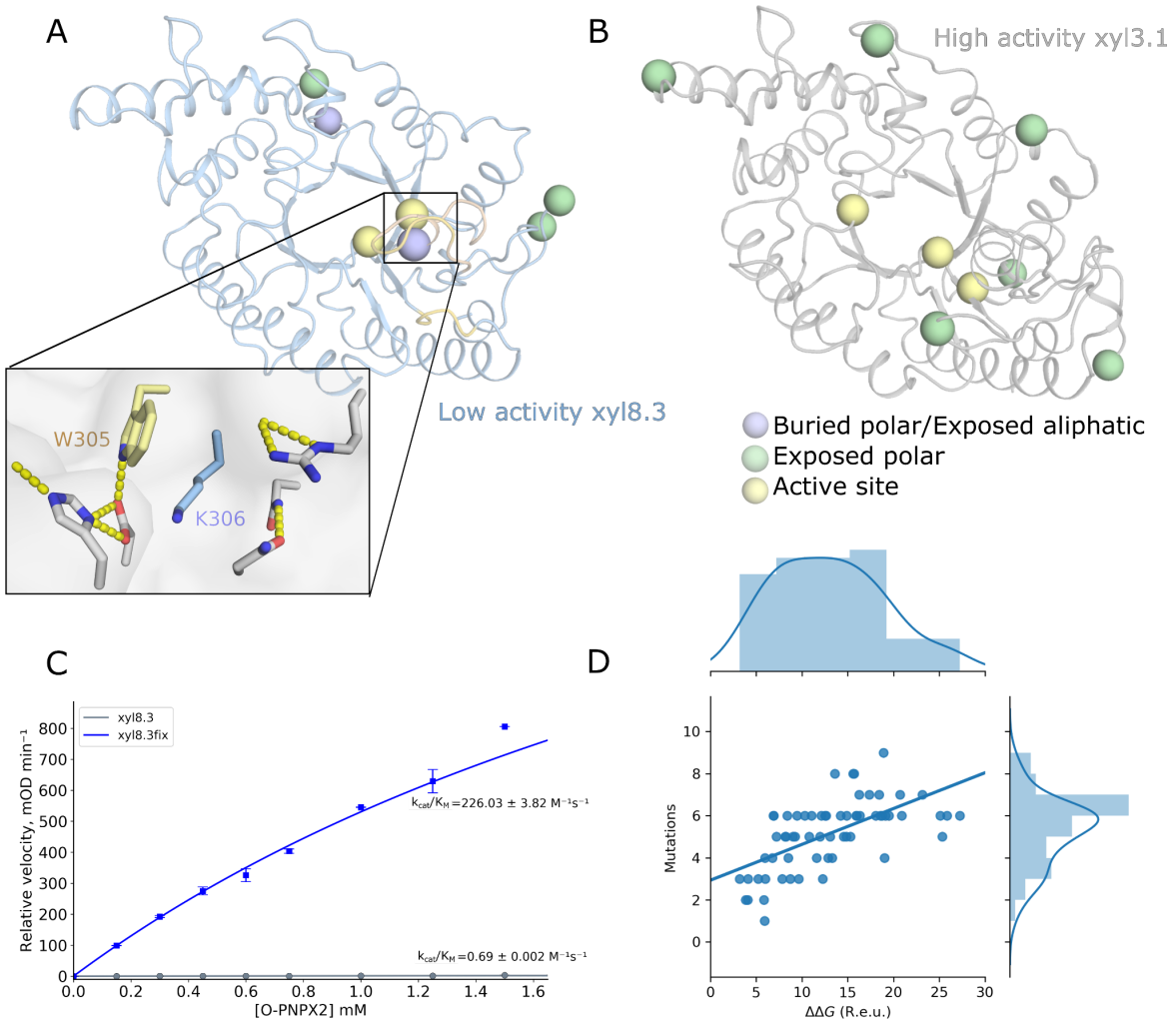


Figure 1. pSUFER analysis of two GH10 xylanases generated using modular assembly and design. **(A)** xyl8.3 backbone shown in cartoon with regions that failed to exhibit electron density in a crystallographic analysis (PDB entry: 6FHE) shown in wheat. (inset) Lys306 is flagged by pSUFER since the Lys is buried in a hydrophobic region without countercharge stabilization. Lys306 is located directly under the two loops that failed to exhibit electron density and is in close contact with active site Trp305. **(B)** For comparison, in the case of the high-efficiency and accurately designed xyl3.1 pSUFER only flags surface-exposed polar positions and active-site residues. In both cases, a position was flagged if mutational scanning suggested at least six amino acid identities with $\Delta\Delta G < 0$ at the position. **(C)** Xyl8.3fix shows an improvement of 330 fold in activity compared to xyl8.3, k_{cat}/K_M=226 and 0.69 $M^{-1}s^{-1}$, respectively. Results are an average of at least two repetitions. **(D)** Improvement in system energy following FuncLib design of positions flagged by pSUFER in 62 models of *de novo* designed enzymes generated by modular assembly and design. On average, 5-6 mutations are introduced, yielding an average improvement of 13 Rosetta energy units (R.e.u.).

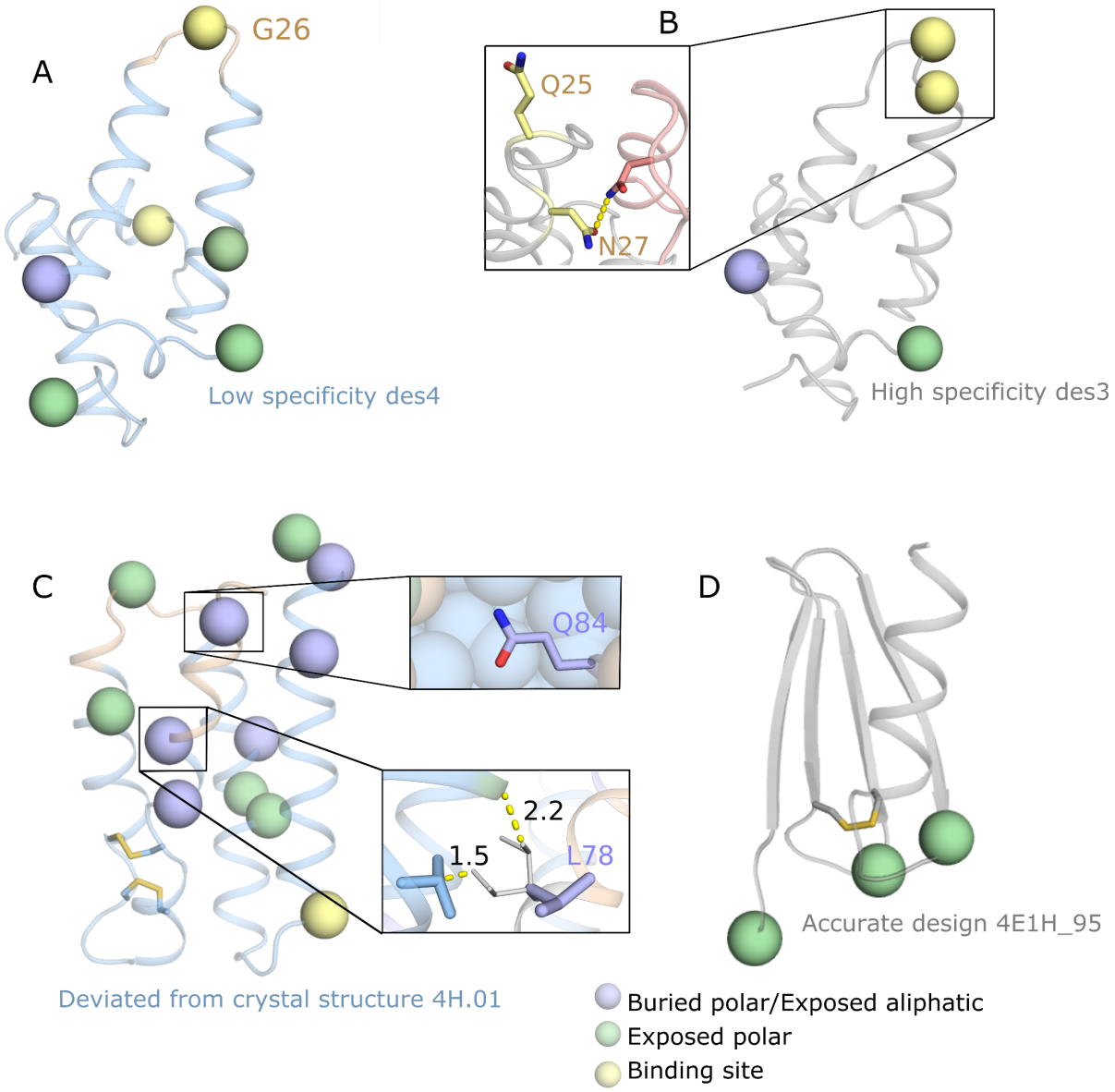


Figure 2. pSUFER analysis of designed binders. **(A+B)** binders generated through modular assembly and design by grafting an interfacial loop (top) from an unrelated protein and designing the two binding partners [24]. Positions were flagged if mutational scanning suggested at least six amino acid identities with $\Delta\Delta G < 0$ at the position. Wheat backbone marks missing density on the design model. **(A)** The interfacial loop failed to exhibit electron density. pSUFER analysis flags Gly26 in the designed loop. **(B)** pSUFER analysis of the accurately designed des3 flags exposed residues and two interfacial positions. Asn27 is one of the interface loop positions that forms polar contacts with the binding protein and thus may be crucial for function (inset). **(C+D)** pSUFER analysis of *de novo* designed binders. Positions were flagged if mutational scanning suggested at least four amino acid identities with $\Delta\Delta G < 0$. Wheat backbone marks missing density on the design model. **(C)** Crystallographic analysis of 4H.01 (PDB entry: 6YWD) revealed regions of missing density. The pSUFER analysis flags several strained positions surrounding the missing density. (insets) Gln53 is partly

desolvated but does not form stabilizing hydrogen bonds. The sidechain conformation of Leu78 is strained in the design model (purple). The most likely sidechain conformation for this position (white) is disallowed due to steric overlaps with neighboring sidechains. **(D)** Design 4E1H_95 was atomically accurate as verified by x-ray crystallography (PDB entry: 6YWC) and pSUFER analysis flags only a few exposed polar sidechains.

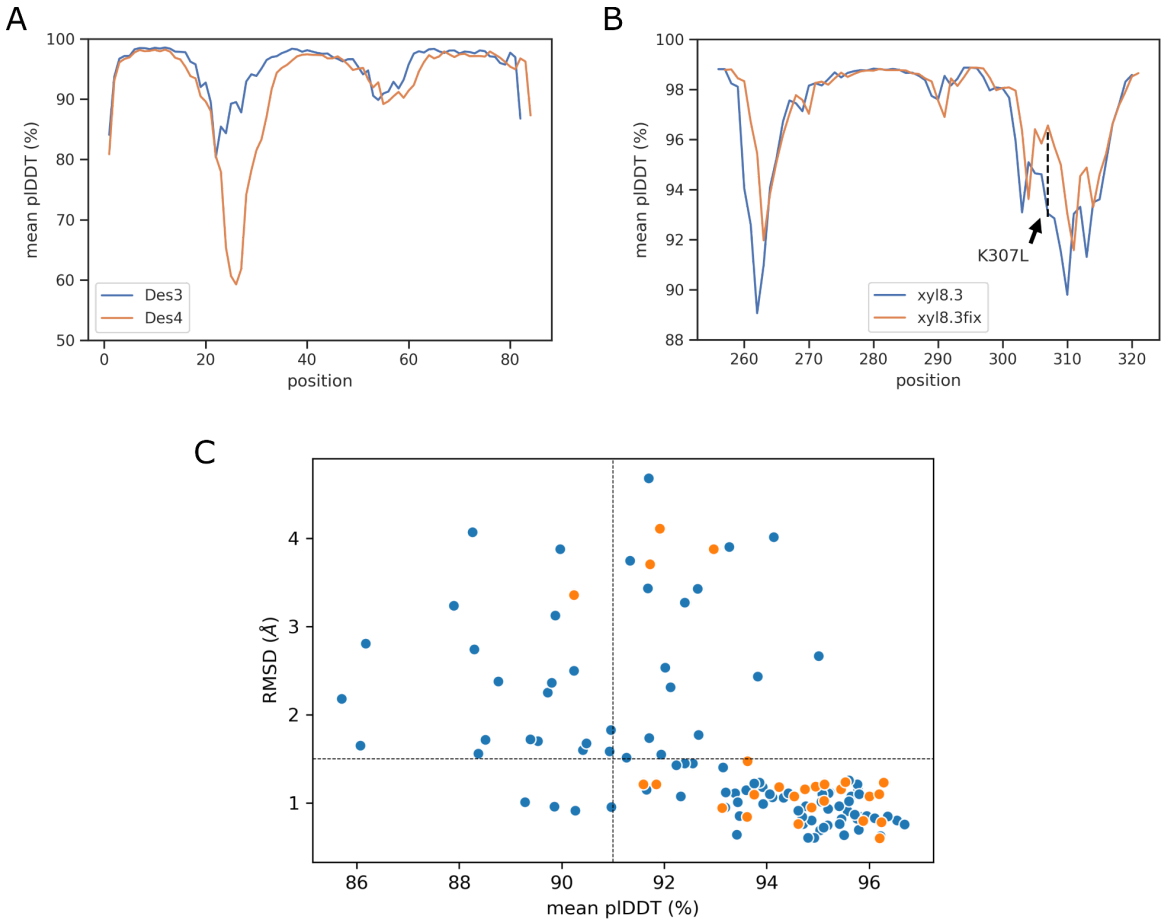


Figure 3. AlphaFold2 analysis of designs. **(A)** plDDT scores for designed binders des3 and des4. The plDDT scores of des4 are depressed relative to those of des3 in the region in which des4 exhibits missing density. **(B)** Comparison of the plDDT scores of xyl8.3 and design xyl8.3fix in the region surrounding the positions that exhibit missing density in crystallographic analysis of xyl8.3 (PDB entry: 6FHE; density is missing surrounding positions 259-265 and 301-312). In this region only a single mutation was designed in xyl8.3fix, Lys306Leu; yet, the plDDT scores of both loops improve significantly. **(C)** AlphaFold2 analysis of 124 proteins designed for *de novo* enzyme activity. Designs that exhibit cooperative folding in thermal-melt analyses (orange dots) tend to exhibit higher mean plDDT scores (>91%) and lower rmsd (<1.5 Å) compared to those that failed to exhibit cooperative folding (blue dots).

Materials and methods

pSUFER algorithm

RosettaScripts [42] and commandlines are available through github https://github.com/Fleishman-Lab/pSUFER. For all Rosetta calculations the ref15 energy function is used for scoring [43]. We start the procedure with four iterations of refinement of the input structure comprising sidechain packing and harmonically constrained whole-protein minimization. Next, computational mutation scanning is performed using the FilterScan mover in Rosetta [10]: for each position all 20 amino acids are modeled against the refined structure, and sidechains within 8 Å are repacked including constrained whole-protein minimization in order to accommodate the mutation. The energy difference between the refined structure and the single-point mutant is calculated. Positions that exhibit several mutations with $\Delta\Delta G \leq 0$ R.e.u. are flagged as suboptimal with the mutation threshold in this work set to four for the *de novo* designed binders and six for all the other designs. Thresholds can be set by the user.

FuncLib design calculations

FuncLib was performed on the pSUFER flagged positions in xyl8.3 as described in [12] with the following thresholds PSSM ≥ -2 and $\Delta\Delta G \leq 1$ R.e.u. in all flagged positions outside the active site. All combinations of mutations were scored using Rosetta. Within the top 20 designs ranked by system energy, the designs converged on the same solution for two of the positions and we chose the design that exhibited the most polar solvent facing residues (Thr1Asn, Asn4Glu) for experimental characterization.

AlphaFold2 *ab initio* structure prediction calculations

All AlphaFold2 [40] calculations were implemented by adapting and locally running the code written by ColabFold [44]. All runs were performed using the five model parameters presented in CASP14, without templates and with Amber relaxation.

Materials

4-nitrophenyl β-xylobioside (OPNPX2) was purchased from Megazyme.

Cloning, protein expression and purification

All experimental procedures were performed as described in [25]. Briefly, the xyl8.3fix design was ordered as a synthetic gene fragment from Twist Bioscience and cloned into pETMBPH vector which contains N-terminal 6-His-tag and MBP. EcoRI and PstI sites were used. The xyl8.3 and xyl8.3fix designs were transformed into BL21 DE3 cells and the DNA was extracted and validated by Sanger sequencing.

50ml of 2YT with 50μg ml^{-1} kanamycin was inoculated with 500μl overnight culture and grown at 37 ℃ to OD of 0.4-0.8. Overexpression was induced by 0.2mM IPTG and the cultures were grown for ~20h at 20℃. Bacteria were pelleted by centrifugation and frozen for at least 20 min before purification.

Pellets were resuspended in lysis buffer (50mM Tris-Cl pH 6.5, 150mM NaCl, benzonase and 0.1 mg ml^{-1} lysozyme) and lysed by sonication. The supernatant was loaded on a column packed with amylose resin (NEB), washed twice with 50 mM Tris pH 6.5 and 150 mM NaCl, and eluted with wash buffer containing 10 mM maltose. Protein purity was evaluated by SDS-PAGE gel and protein concentration was estimated by OD_{280}. In cases where purity was not satisfactory, the elution was loaded on an Ni-NTA, washed and eluted (50mM Tris-Cl pH 6.5, 150mM NaCl, 20mM Imidazole and 250mM Imidazole respectively). The proteins were then dialyzed against 50mM Tris-Cl pH 6.5, 150mM NaCl buffer.

Kinetic measurements

Kinetic measurements were performed with purified proteins (fused to MBP) in 96-well plates (optical length –0.5 cm) by monitoring the absorbance of the leaving group of O-PNPX2 at 405 nm (activity buffer 50 mM Tris pH 6.5 and 150, 25 °C). No background hydrolysis was observed with O-PNPX2. Final protein concentrations in the reaction varied between 2.56 μM and 6.23 μM. The data were fitted to the linear regime of the Michaelis-Menten model ($v_0 = [S]_0[E]_0 k_{cat}/K_M$) and kcat/KM values were deduced from the slope. The reported values represent the means ± S.D. of at least two independent measurements.

Apparent melting-temperature measurements

Tm measurements were done using nanoDSF experiments [45] performed on Prometheus™ NT.Plex instrument (NanoTemper Technologies). Melting temperatures were between 20°C-85°C with 1.0°C/min slope. For xyl8.3 and xyl8.3fix the MBP tag was removed by TEV cleavage

References

1. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Design of a novel globular protein fold with atomic-level accuracy. Science. 2003;302: 1364–1368.

2. Koga N, Tatsumi-Koga R, Liu G, Xiao R, Acton TB, Montelione GT, et al. Principles for designing ideal protein structures. Nature. 2012;491: 222–227.

3. Rocklin GJ, Chidyausiku TM, Goreshnik I, Ford A, Houliston S, Lemak A, et al. Global analysis of protein folding using massively parallel design, synthesis, and testing. Science. 2017;357: 168–175.

4. Jacobs TM, Williams B, Williams T, Xu X, Eletsky A, Federizon JF, et al. Design of structurally distinct proteins using strategies inspired by evolution. Science. 2016;352: 687–690.

5. King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, Andre I, et al. Computational design of self-assembling protein nanomaterials with atomic level accuracy. Science. 2012;336: 1171–1174.

6. Hsia Y, Bale JB, Gonen S, Shi D, Sheffler W, Fong KK, et al. Design of a hyperstable 60-subunit protein icosahedron. Nature. 2016;535: 136–139.

7. Mou Y, Yu J-Y, Wannier TM, Guo C-L, Mayo SL. Computational design of co-assembling protein-DNA nanowires. Nature. 2015;525: 230–233.

8. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch E-M, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science. 2011;332: 816–821.

9. Chen K-YM, Keri D, Barth P. Computational design of G Protein-Coupled Receptor allosteric signal transductions. Nat Chem Biol. 2020;16: 77–86.

10. Whitehead TA, Chevalier A, Song Y, Dreyfus C, Fleishman SJ, De Mattos C, et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. Nat Biotechnol. 2012;30: 543–548.

11. Goldenzweig A, Goldsmith M, Hill SE, Gertman O, Laurino P, Ashani Y, et al. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. Mol Cell. 2018;70: 380.

12. Khersonsky O, Lipsh R, Avizemer Z, Ashani Y, Goldsmith M, Leader H, et al. Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. Mol Cell. 2018;72: 178–186.e5.

13. Bednar D, Beerens K, Sebestova E, Bendl J, Khare S, Chaloupkova R, et al. FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants. PLoS Comput Biol. 2015;11: e1004556.

14. Chevalier A, Silva D-A, Rocklin GJ, Hicks DR, Vergara R, Murapa P, et al. Massively parallel de novo protein design for targeted therapeutics. Nature. 2017. doi:10.1038/nature23912

15. Röthlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, et al.

Kemp elimination catalysts by computational enzyme design. Nature. 2008;453: 190–195.

16. Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, et al. De novo computational design of retro-aldol enzymes. Science. 2008;319: 1387–1391.

17. Privett HK, Kiss G, Lee TM, Blomberg R, Chica RA, Thomas LM, et al. Iterative approach to computational enzyme design. Proc Natl Acad Sci U S A. 2012;109: 3790–3795.

18. Baker D. An exciting but challenging road ahead for computational enzyme design. Protein Sci. 2010;19: 1817–1819.

19. Sterner R. Faculty Opinions recommendation of Precision is essential for efficient catalysis in an evolved Kemp eliminase. Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature. 2013. doi:10.3410/f.718146647.793485608

20. Khare SD, Fleishman SJ. Emerging themes in the computational design of novel enzymes and protein-protein interfaces. FEBS Lett. 2013;587: 1147–1154.

21. Korendovych IV, DeGrado WF. Catalytic efficiency of designed catalytic proteins. Curr Opin Struct Biol. 2014;27: 113–121.

22. Khersonsky O, Kiss G, Röthlisberger D, Dym O, Albeck S, Houk KN, et al. Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59. Proc Natl Acad Sci U S A. 2012;109: 10358–10363.

23. Fleishman SJ, Whitehead TA, Strauch EM, Corn JE, Qin S, Zhou HX, et al. Community-wide assessment of protein-interface modeling suggests improvements to design methodology. J Mol Biol. 2011;414: 289–302.

24. Netzer R, Listov D, Lipsh R, Dym O, Albeck S, Knop O, et al. Ultrahigh specificity in a network of computationally designed protein-interaction pairs. Nature Communications. 2018. doi:10.1038/s41467-018-07722-9

25. Lapidoth G, Khersonsky O, Lipsh R, Dym O, Albeck S, Rogotner S, et al. Highly active enzymes by automated combinatorial backbone assembly and sequence design. Nat Commun. 2018;9: 2780.

26. Dou J, Vorobieva AA, Sheffler W, Doyle LA, Park H, Bick MJ, et al. De novo design of a fluorescence-activating β-barrel. Nature. 2018;561: 485–491.

27. Baran D, Pszolla MG, Lapidoth GD, Norn C, Dym O, Unger T, et al. Principles for computational design of binding antibodies. Proc Natl Acad Sci U S A. 2017;114: 10900–10905.

28. Goldenzweig A, Fleishman SJ. Principles of Protein Stability and Their Application in Computational Design. Annu Rev Biochem. 2018;87: 105–129.

29. Lapidoth GD, Baran D, Pszolla GM, Norn C, Alon A, Tyka MD, et al. AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. Proteins. 2015;83: 1385–1406.

30. Lipsh-Sokolik R, Listov D, Fleishman SJ. The AbDesign computational pipeline for

modular backbone assembly and design of binders and enzymes. Protein Sci. 2021;30: 151–159.

31. Park H, Bradley P, Greisen P Jr, Liu Y, Mulligan VK, Kim DE, et al. Simultaneous Optimization of Biomolecular Energy Functions on Features from Small Molecules and Macromolecules. J Chem Theory Comput. 2016;12: 6201–6212.

32. Ferreiro DU, Komives EA, Wolynes PG. Frustration in biomolecules. Q Rev Biophys. 2014;47: 285–363.

33. Ferreiro DU, Komives EA, Wolynes PG. Frustration, function and folding. Curr Opin Struct Biol. 2018;48: 68–73.

34. Freiberger MI, Guzovsky AB, Wolynes PG, Parra RG, Ferreiro DU. Local frustration around enzyme active sites. Proc Natl Acad Sci U S A. 2019;116: 4037–4043.

35. Yang C, Sesterhenn F, Bonet J, van Aalen EA, Scheller L, Abriata LA, et al. Bottom-up de novo design of functional proteins with complex structural features. Nat Chem Biol. 2021;17: 492–500.

36. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021. doi:10.1038/s41586-021-03819-2

37. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373: 871–876.

38. Wang S, Sun S, Li Z, Zhang R, Xu J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS Comput Biol. 2017;13: e1005324.

39. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv. 2021. p. 2021.10.04.463034. doi:10.1101/2021.10.04.463034

40. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Applying and improving AlphaFold at CASP14. Proteins. 2021. doi:10.1002/prot.26257

41. Elazar A, Chandler NJ, Davey AS, Weinstein JY, Nguyen JV, Trenker R, et al. De novo designed receptor transmembrane domains enhance CAR-T cell cytotoxicity and attenuate cytokine release. bioRxiv. 2020. p. 2020.07.26.221598. doi:10.1101/2020.07.26.221598

42. Fleishman SJ, Leaver-Fay A, Corn JE, Strauch E-M, Khare SD, Koga N, et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. PLoS One. 2011;6: e20161.

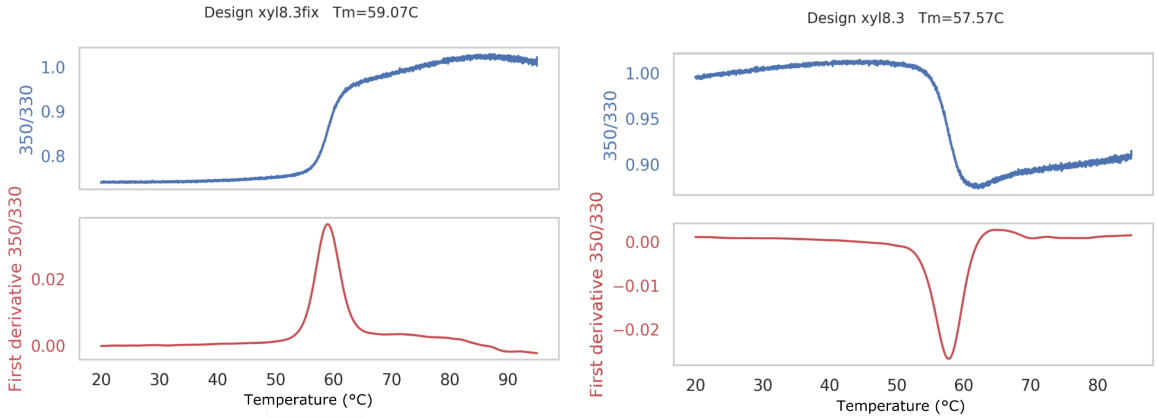
43. Alford RF, Leaver-Fay A, Jeliazkov JR, O'Meara MJ, DiMaio FP, Park H, et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. J Chem Theory Comput. 2017;13: 3031–3048.

44. Mirdita M, Schuetze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold - Making protein folding accessible to all. doi:10.1101/2021.08.15.456425

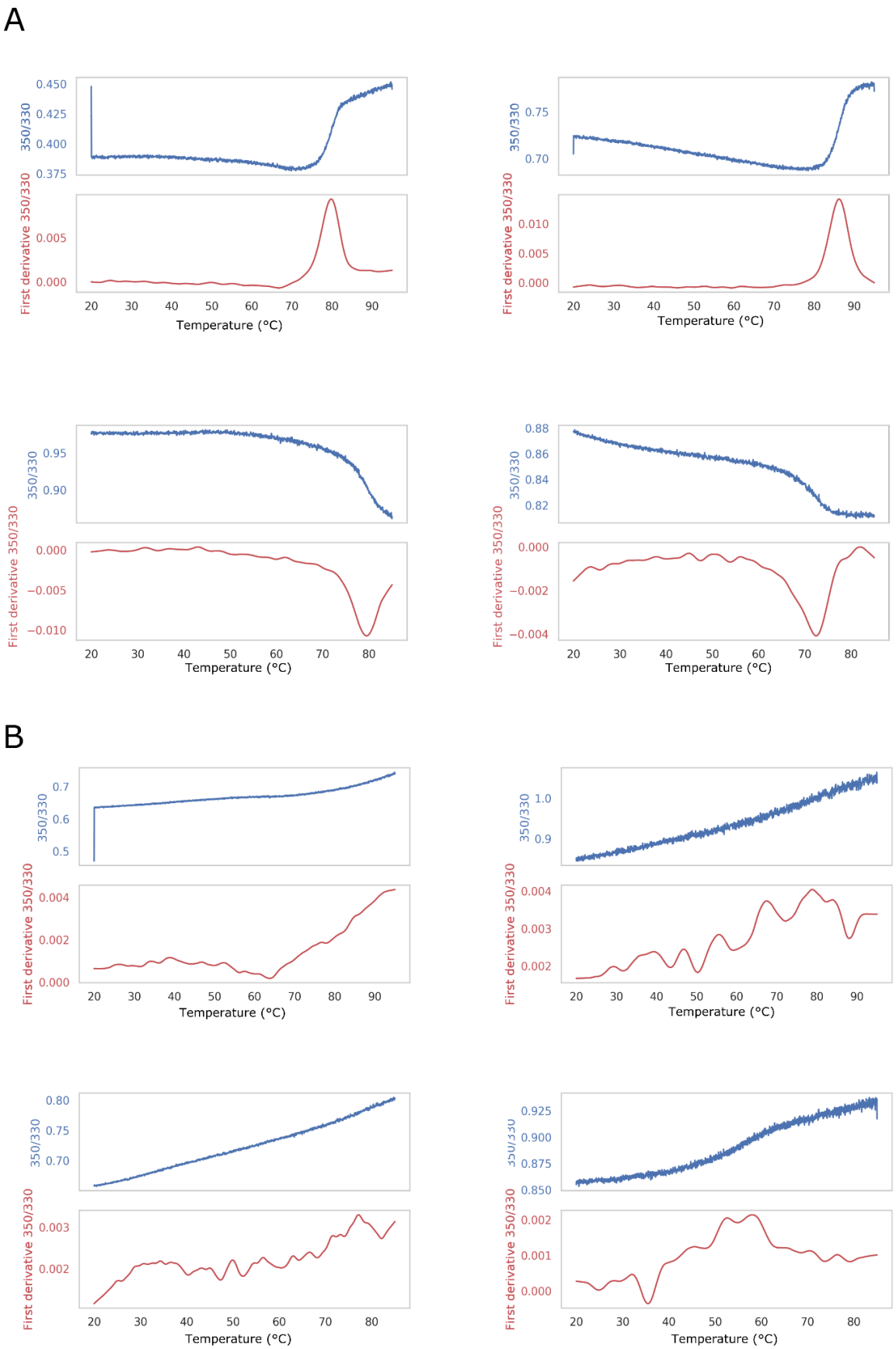
45. Wedde S, Kleusch C, Bakonyi D, Gröger H. High-Throughput Feasible Screening Tool

for Determining Enzyme Stabilities against Organic Solvents Directly from Crude Extracts. Chembiochem. 2017;18: 2399–2403.

Supplementary material



Supplementary Figure 1. Denaturation curves for xyl8.3 and xyl8.3fix (blue - ratio, red - derivative) using nanoDSF.



Supplementary Figure 2. Typical denaturation curves for *de novo* designed enzymes generated through modular assembly and design (blue - ratio, red - derivative). (A) Cooperatively folded proteins with apparent melting temperatures of 76-86 °C. (B) Representative misfolded or unfolded proteins.