Customizing the Structure of a Minimal TIM Barrel to Craft a De Novo Enzyme

Julian Beck,[1,†] Benjamin J. Smith,[2,3,†] Niayesh Zarifi,[2,3] Emily Freund,[1] Roberto A. Chica,[2,3,*] Birte Höcker[1,*]

[1] Department of Biochemistry, University of Bayreuth, 95447 Bayreuth, Germany.

[2] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

[3] Center for Catalysis Research and Innovation, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

[†] These authors contributed equally.

*** Corresponding authors**

Birte Höcker, E-mail: birte.hoecker@uni-bayreuth.de

Roberto A. Chica, E-mail: rchica@uottawa.ca

Keywords

De novo enzyme design, computational protein design, enzymes, TIM barrel, Kemp elimination

Abstract

The TIM barrel is the most prevalent fold in natural enzymes, supporting efficient catalysis of diverse chemical reactions. While *de novo* TIM barrels have been successfully designed, their minimalistic architectures lack the structural elements necessary for substrate binding and catalysis. Here, we present CANVAS, a computational workflow that introduces a structural lid into a minimal *de novo* TIM barrel to anchor catalytic residues and form an active-site pocket for enzymatic function. Starting from two *de novo* TIM barrels, we designed nine variants with distinct lids to form active sites for the Kemp elimination. Experimental testing identified one active variant with catalytic efficiency comparable to previously reported Kemp eliminases. Mutational analyses confirmed that the designed lid is essential for catalysis, while molecular dynamics simulations revealed the importance of lid flexibility for substrate binding. This study reports the first enzymatically active *de novo* TIM barrel and establishes a platform for designing enzymes from minimal protein scaffolds.

Main text

The TIM barrel is one of the most versatile enzyme folds. It is present in six of the seven enzyme classes and supports catalysis at diffusion-limited rates[1-3]. While idealized *de novo* TIM barrels have been successfully designed[4-7], none have been converted into functional enzymes due to their minimalistic architecture, which lacks key structural elements required for catalysis, such as cavities, pockets, and extended loops. Efforts to address these deficiencies have led to the introduction of secondary structural elements to create rudimentary pockets[8-10]. However, these pockets are too solvent exposed and undefined to support the microenvironments required for catalysis. Furthermore, many *de novo* TIM barrels are circularly permuted relative to their natural counterparts, positioning termini on the catalytic face, an arrangement that limits the number of designable βα loops available for scaffolding structural elements needed to form active sites. Therefore, novel strategies are needed to craft custom active sites and endow *de novo* TIM barrels with enzymatic function.

Here, we introduce CANVAS (Customizing Amino-acid Networks for Virtual Active-site Scaffolding), a computational workflow for transforming minimal proteins such as *de novo* TIM barrels into functional enzymes (Figure 1a, Figure S1). The process begins with a minimal TIM barrel template and a theozyme, which is a computational model of an idealized enzyme active site with catalytic groups arranged to stabilize a reaction's transition state[11]. Using the protein design software Triad[12], a catalytic amino acid from the theozyme is placed onto the TIM barrel template, and the transition state is built from its side chain to maintain the catalytic geometry[13]. A second catalytic residue is then built from the transition state, ensuring proper geometry for catalytic interaction. The alpha carbon of this second residue is thus positioned in the empty space above the TIM barrel catalytic face. This allows it to serve as an anchor for constructing a custom lid to form the active-site pocket, which is created using the generative AI tool RFdiffusion[14] followed by sequence design with ProteinMPNN[15] and evaluation with AlphaFold2[16]. The active-

59 site is further optimized with Triad to build a pocket complementary to the transition state[17].
60 Designs are then filtered using key enzyme design criteria, including solvent-accessible surface
61 area of the transition state, active-site preorganization, energy, and catalytic contact geometry[18].

62 We applied CANVAS to the Kemp elimination (Figure 1b), a model organic transformation
63 commonly used as a benchmark for *de novo* enzyme design[17-19]. The reaction's theozyme
64 features an Asp as catalytic base and an Asn as H-bond donor to stabilize negative charge buildup
65 on the oxygen at the transition state. We used the crystal structures of two *de novo* TIM barrels
66 (PDB IDs: 7MCD[20] and 7OSV[5]) as templates. Since 7OSV has its N- and C-termini on the catalytic
67 face of the TIM barrel, we applied an *in silico* circular permutation based on the inpainting
68 method[21] (Methods) to generate two modified structures termed NT6-CP1 and NT6-CP2 that
69 contain termini on the stability face (Figure S2). Starting from these structures, we used CANVAS
70 to generate a distinct lid for each template tailored to the Kemp elimination (Figure 1c). The lid
71 diffused onto 7MCD consists of a single 26-residue fragment introduced in βα loop 7 (Figure S3).
72 By contrast, the designed lids in NT6-CP1 and NT6-CP2 comprise a major fragment of 33 or 39
73 residues introduced into βα loops 2 or 6, respectively, to scaffold the catalytic H-bond donor, as
74 well as three or four elongated loops to enhance interactions with the primary fragment. The
75 resulting designs, termed KempTIMs, include variants 1–3 obtained from NT6-CP2, variants 4-5
76 from NT6-CP1 and variants 6–9 from 7MCD, each with a unique active-site configuration (Figure
77 S4) and sequence (Table S1), which yielded high AlphaFold2 pLDDT scores (Figure S5).

78 Of the nine designs, seven expressed in soluble form in *E. coli* but only five could be
79 purified to homogeneity, yielding 3–20 mg per liter of culture (Table S2). Notably, the NT6-CP1
80 template could not be expressed, in contrast to NT6-CP2 and 7MCD. Enzymatic assays revealed
81 a single active design, KempTIM4, which followed Michaelis-Menten kinetics with substrate
82 saturation (Figure 2a), yielding k_{cat}, K_M, and k_{cat}/K_M values of 0.0066 ± 0.0006 s^{-1}, 0.6 ± 0.1 mM,
83 and 11 ± 3 M^{-1} s^{-1}, respectively. This catalytic efficiency is comparable to those of other *de novo*

84 Kemp eliminases[17,19,22], such as KE07 (k_{cat} = 0.018 s^{-1}, K_M = 1.4 mM, and k_{cat}/K_M = 12.2 M^{-1} s^{-1})[19]. KempTIM4 expressed predominantly as a monomeric protein (Figure 3a,b, Table S3) and displayed a mixed αβ secondary structure characteristic of TIM barrels (Figure 3c). Compared to the minimal TIM barrel NT6-CP2, KempTIM4 exhibited increased random coil content, indicated by a stronger circular dichroism (CD) signal at 208 nm. This structural feature aligns with the design of its lid, which includes long loops (Figure 3a, Figure S3). KempTIM4 showed a 35°C reduction in melting temperature relative to NT6-CP2 (Figure 3d), reflecting the destabilizing effects of incorporating a lid and active site onto the minimal TIM barrel scaffold. Mutation of the designed catalytic base and H-bond donor to Ala abolished or substantially reduced activity (Figure 2b) despite size-exclusion chromatograms and CD spectra comparable to KempTIM4 (Figure S6, Table S3). These results confirm that the loss of activity in these knockout variants is due to mutation, not structural disruption. Moreover, the absence of enzymatic activity in NT6-CP2 demonstrates that catalysis in KempTIM4 is driven by the designed lid and catalytic residues.

97 Although active, KempTIM4 was prone to aggregation at high concentrations. This led us to redesign its sequence for enhanced expression and solubility using ProteinMPNN, which has been shown to produce stable and soluble proteins[23]. In this procedure, identities of active-site residues designed with Triad were preserved to maintain catalytic function, while the rest of the protein was redesigned. The redesigned protein, KempTIM4-2, demonstrated approximately 10-fold higher expression yields compared to KempTIM4 (Table S2) and existed exclusively as a monomer with no tendency to form dimers (Figure 3b, Table S3). CD spectroscopy showed increased α-helicity compared to KempTIM4 and higher stability, comparable to NT6-CP2 (Figure 3c,d). Despite these improvements, KempTIM4-2 exhibited lower activity than KempTIM4 (k_{cat}/K_M = 7.0 ± 0.1 M^{-1} s^{-1}) and could not be saturated within the substrate's solubility limit (Figure 2a), even though active-site residues remained unchanged. As with KempTIM4, mutation of the catalytic base in KempTIM4-2 abolished activity (Figure 2b) without disrupting structure (Figure

109 S6, Table S2). However, mutation of the H-bond donor caused only a small reduction in catalytic
110 activity compared to KempTIM4, suggesting that this designed Asn residue adopts a suboptimal
111 orientation hindering its catalytic function. Nevertheless, inhibition assays using the 6-
112 nitrobenzotriazole (6NBT) transition-state analogue demonstrated competitive binding in the
113 active site (Figure 2c), confirming the integrity of the designed active site pocket.

114 To explore why only KempTIM4 and KempTIM4-2 were catalytically active, microsecond-
115 timescale molecular dynamics simulations were performed on these designs and the inactive
116 KempTIM9, both in the presence and absence of the transition-state analogue. KempTIM9 was
117 selected as a comparison due to its good expression (Table S2), well-folded structure and high
118 melting temperature (67°C, Figure S7), which rule out misfolding as the cause of inactivity. The
119 simulations revealed that the helical component of the lids, which anchor the catalytic Asn, were
120 rigid (Figure 4a). This rigidity stabilizes the transition-state analogue within the binding pocket,
121 maintaining catalytic interactions with the catalytic base for much of the simulation, with active
122 variants (KempTIM4 and KempTIM4-2) outperforming the inactive design KempTIM9 (Figure 4b).
123 However, only KempTIM4 consistently maintained catalytic interactions with the H-bond donor.
124 Both KempTIM4-2 and KempTIM9 showed weaker contacts with this Asn residue, with
125 KempTIM4-2 being unable to form this interaction throughout the simulation. This result aligns
126 with the small impact on activity observed when this residue was mutated to Ala in KempTIM4-2
127 (Figure 2d).

128 A key difference between the active designs and the inactive KempTIM9 was the
129 significantly narrower entrance to the active site in KempTIM9 (Figure 4b). Previous studies have
130 demonstrated that widening the active-site entrance enhances Kemp eliminase activity by
131 lowering the energy barriers associated with substrate entry and product release[24]. Thus, we
132 postulate that the inactivity of KempTIM9 is due to an insufficiently open active site, which likely
133 impedes substrate binding. By contrast, the lids of KempTIM4 and KempTIM4-2, which

134 incorporate five fragments rather than the single fragment found in KempTIM9 (Figure S3), may
135 facilitate substrate entry. The additional inserted loops in these designs likely contribute to active-
136 site accessibility by providing greater flexibility to the lid in the unbound state (Figure 4a), allowing
137 the active site to open more readily. While other factors, such as incorrect positioning of catalytic
138 residues and/or lid misorientation, cannot be excluded, these findings suggest that reducing
139 barriers to substrate entry during lid generation could improve catalytic performance in future
140 designs.

141 Natural TIM barrels use lids to bind substrates and form an active site conducive to
142 catalysis[25]. By contrast, minimal *de novo* TIM barrels lack such lids and are inherently
143 functionless. Using our CANVAS approach, we successfully designed a custom lid to form an
144 active site for a target reaction, resulting in the first enzyme constructed from a *de novo* TIM
145 barrel. While KempTIM4 exhibits modest activity, it matches the performance of many first-round
146 designs generated by traditional methods[17,19], despite the added complexity of engineering a
147 tailored lid comprising multiple loops to create an active site cavity. Unlike traditional enzyme
148 design methods that rely on pre-existing scaffolds, CANVAS offers the flexibility to customize lids
149 for specific reaction requirements. It can also be readily adapted to scaffold more than two
150 catalytic residues, a task that becomes increasingly challenging with pre-existing scaffolds as the
151 number of catalytic groups grows. With CANVAS, we anticipate that any minimal TIM barrel,
152 including AI-generated virtual ones, could serve as a blank canvas for the creation of a large
153 diversity of *de novo* enzymes, unlocking the full potential of this versatile fold for catalysis of a
154 wide range of reactions.

Figures

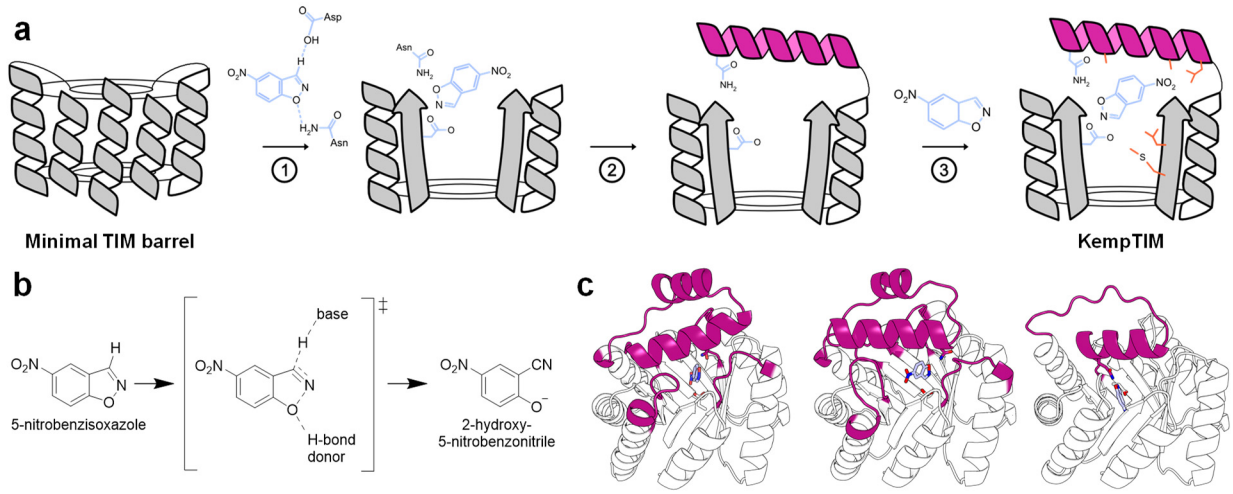

Figure 1. De novo enzyme design using CANVAS. (a) CANVAS follows three general steps: (1) placement of the theozyme (blue) onto the minimal TIM-barrel scaffold using a single catalytic residue as an anchor; (2) construction of a custom lid (magenta) to form the active-site pocket and anchor the second catalytic residue; and (3) active-site repacking to further stabilize the transition state. Details are provided in Main Text. (b) Kemp elimination reaction. (c) Structure of three distinct lids (magenta) tailored to the Kemp elimination, designed using CANVAS. Transition state and designed catalytic residues are shown as sticks.

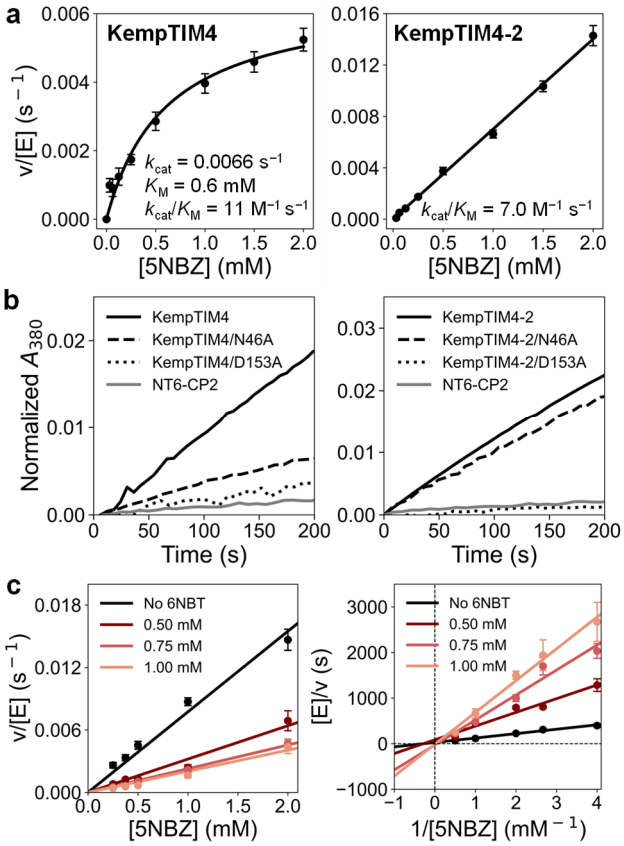

Figure 2. Kinetic characterization of KempTIMs. (a) Michaelis–Menten plots of normalized initial rates as a function of 5-nitrobenzisoxazole (5-NBZ) concentrations. Data represent the average of six replicate measurements from two independent protein batches, with error bars indicating the SEM. (b) Reaction progress curves ([5NBZ] = 2 mM) demonstrate the contributions of designed catalytic residues D153 and N46 to catalysis. (c) Michaelis-Menten (left) and Lineweaver-Burk (right) plots reveal that the transition-state analogue 6NBT acts as a competitive inhibitor of KempTIM4-2. Data represented as in (a).

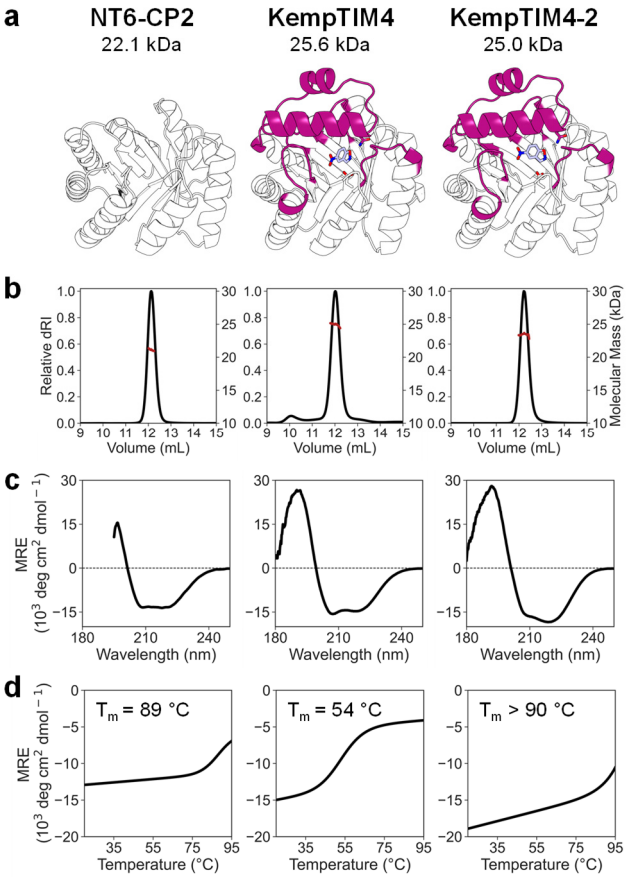

Figure 3. Structural characterization. (a) Computational models of minimal TIM-barrel NT6-CP2 and active KempTIMs with the designed lid shown in magenta and the theozyme as sticks. (b) SEC-MALS indicates that the proteins are predominantly monomeric, with molecular weights matching their expected values. (c) CD spectra reveal a mixed αβ signal characteristic of TIM barrels. MRE: mean residue ellipticity. (d) Melting curves demonstrate that KempTIM4 is substantially destabilized compared to NT6-CP2. However, redesign to yield KempTIM4-2 restores stability.

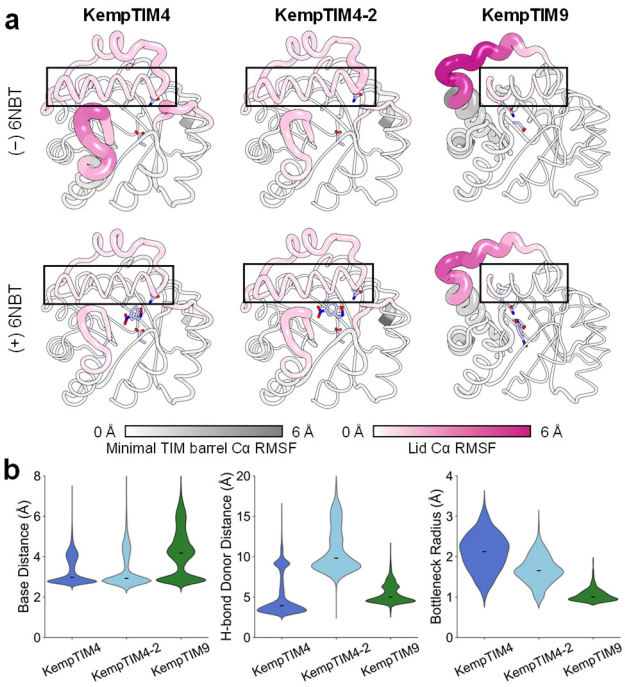

Figure 4. Molecular dynamics in the presence and absence of transition-state analogue 6NBT. (a) Root-mean-square fluctuations (RMSF) of the protein backbone, represented as putty plots, reveal that the lid α-helix anchoring the catalytic H-bond donor (boxed) is rigid, whereas loops exhibit greater flexibility. 6NBT and catalytic residues are depicted as sticks. (b) Distributions of catalytic contact distances between transition-state analogue and catalytic base or H-bond donor are shown alongside the active-site entrance bottleneck radius, presented as violin plots, with the median value indicated by a black line.

Acknowledgments

J.B. and B.H. acknowledge support from the Elite Network of Bavaria and its study program "Biological Physics". B.H. acknowledges support from European Union's Horizon 2020 research and innovation program under grant agreement No 951375 (ArtMotor). R.A.C. acknowledges grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-03484 and RGPAS-2021-00017) and the Canada Foundation for Innovation (26503). This research was enabled in part by support provided by Compute Ontario (www.computeontario.ca) and the Digital Research Alliance of Canada (alliancecan.ca). We thank Sabrina Wischt for technical support and the University of Bayreuth Centre of International Excellence "Alexander von Humboldt" for facilitating this collaboration through a short-term grant to R.A.C.

Competing interests

The authors declare no competing financial interest.

ORCID identifiers

Julian Beck: 0009-0007-3555-9890

Roberto A. Chica: 0000-0003-3789-9841

Emily Freund: 0009-0006-2876-3319

Birte Höcker: 0000-0002-8250-9462

Benjamin J. Smith: 0009-0003-5953-4998

Niayesh Zarifi: 0000-0003-4748-7082

Supporting Information

Supporting Tables S1–S8, Supporting Figures S1–S7, experimental procedures.

References

(1) Nagano, N.; Orengo, C. A.; Thornton, J. M. *Journal of Molecular Biology* **2002**, *321*, 741.

(2) Höcker, B.; Jürgens, C.; Wilmanns, M.; Sterner, R. *Current Opinion in Biotechnology* **2001**, *12*, 376.

(3) Wierenga, R. K.; Kapetaniou, E. G.; Venkatesan, R. *Cellular and Molecular Life Sciences* **2010**, *67*, 3961.

(4) Huang, P.-S.; Feldmeier, K.; Parmeggiani, F.; Fernandez Velasco, D. A.; Höcker, B.; Baker, D. *Nature Chemical Biology* **2016**, *12*, 29.

(5) Kordes, S.; Romero-Romero, S.; Lutz, L.; Höcker, B. *Protein Science* **2022**, *31*, 513.

(6) Romero-Romero, S.; Costas, M.; Silva Manzano, D.-A.; Kordes, S.; Rojas-Ortega, E.; Tapia, C.; Guerra, Y.; Shanmugaratnam, S.; Rodríguez-Romero, A.; Baker, D.; Höcker, B.; Fernández-Velasco, D. A. *Journal of Molecular Biology* **2021**, *433*, 167153.

(7) Chu, A. E.; Fernandez, D.; Liu, J.; Eguchi, R. R.; Huang, P.-S. *BioDesign Research* **2022**, *2022*, 9842315.

(8) Beck, J.; Shanmugaratnam, S.; Höcker, B. *Protein Science* **2024**, *33*, e5001.

(9) Kordes, S.; Beck, J.; Shanmugaratnam, S.; Flecks, M.; Höcker, B. *Protein Engineering, Design and Selection* **2023**, *36*, gzad012.

(10) Wiese, J. G.; Shanmugaratnam, S.; Höcker, B. *Protein Science* **2021**, *30*, 982.

(11) Tantillo, D. J.; Jiangang, C.; Houk, K. N. *Current Opinion in Chemical Biology* **1998**, *2*, 743.

(12) Lee, F. S.; Anderson, A. G.; Olafson, B. D. *Protein Engineering, Design and Selection* **2023**, *36*, gzad013.

(13) Lassila, J. K.; Privett, H. K.; Allen, B. D.; Mayo, S. L. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 16710.

(14) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. *Nature* **2023**, *620*, 1089.

(15) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. *Science* **2022**, *378*, 49.

(16) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. *Nature* **2021**, *596*, 583.

(17) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 3790.

(18) Rakotoharisoa, R. V.; Seifinoferest, B.; Zarifi, N.; Miller, J. D. M.; Rodriguez, J. M.; Thompson, M. C.; Chica, R. A. *J Am Chem Soc* **2024**, *146*, 10001.

(19) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190.

(20) Anand, N.; Eguchi, R.; Mathews, I. I.; Perez, C. P.; Derry, A.; Altman, R. B.; Huang, P.-S. *Nature Communications* **2022**, *13*, 746.

(21) Wang, J.; Lisanza, S.; Juergens, D.; Tischer, D.; Watson, J. L.; Castro, K. M.; Ragotte, R.; Saragovi, A.; Milles, L. F.; Baek, M.; Anishchenko, I.; Yang, W.; Hicks, D. R.; Expòsit, M.; Schlichthaerle, T.; Chun, J.-H.; Dauparas, J.; Bennett, N.; Wicky, B. I. M.; Muenks, A.; DiMaio, F.; Correia, B.; Ovchinnikov, S.; Baker, D. *Science* **2022**, *377*, 387.

(22) Zhang, S.; Zhang, J.; Luo, W.; Wang, P.; Zhu, Y. *Enzyme and Microbial Technology* **2022**, *160*, 110093.

(23) Sumida, K. H.; Núñez-Franco, R.; Kalvet, I.; Pellock, S. J.; Wicky, B. I. M.; Milles, L. F.; Dauparas, J.; Wang, J.; Kipnis, Y.; Jameson, N.; Kang, A.; De La Cruz, J.; Sankaran, B.; Bera, A. K.; Jiménez-Osés, G.; Baker, D. *Journal of the American Chemical Society* **2024**, *146*, 2054.

(24) Broom, A.; Rakotoharisoa, R. V.; Thompson, M. C.; Zarifi, N.; Nguyen, E.; Mukhametzhanov, N.; Liu, L.; Fraser, J. S.; Chica, R. A. *Nature Communications* **2020**, *11*, 4808.

(25) Sterner, R.; Höcker, B. *Chemical Reviews* **2005**, *105*, 4038.