Increased Diels-Alderase activity through backbone remodeling guided by Foldit players

Christopher B Eiben[1,6], Justin B Siegel[1,6], Jacob B Bale[1,2], Seth Cooper[3], Firas Khatib[1], Betty W Shen[4], Foldit Players, Barry L Stoddard[4], Zoran Popovic[3] & David Baker[1,5]

Computational enzyme design holds promise for the production of renewable fuels, drugs and chemicals. *De novo* enzyme design has generated catalysts for several reactions, but with lower catalytic efficiencies than naturally occurring enzymes[1–4]. Here we report the use of game-driven crowdsourcing to enhance the activity of a computationally designed enzyme through the functional remodeling of its structure. Players of the online game Foldit[5,6] were challenged to remodel the backbone of a computationally designed bimolecular Diels-Alderase[3] to enable additional interactions with substrates. Several iterations of design and characterization generated a 24-residue helix-turn-helix motif, including a 13-residue insertion, that increased enzyme activity >18-fold. X-ray crystallography showed that the large insertion adopts a helix-turn-helix structure positioned as in the Foldit model. These results demonstrate that human creativity can extend beyond the macroscopic challenges encountered in everyday life to molecular-scale design problems.

Previous computational methods for enzyme design have kept the backbone unchanged, but it is clear from natural evolution that the optimization of new functions generally involves some backbone rearrangements[7]. *De novo* design methods have been used to create new loops and protein structures[8], and motif-directed design methods can introduce new functional loops when specific interactions are used to direct the modeling[9,10]. However, undirected remodeling of a protein backbone structure to improve function has not yet been achieved. A primary challenge is that the number of possibilities for undirected remodeling, once large insertions and sequence variability are allowed, is too large to be systematically searched by automated methods.

Recent work has demonstrated that crowdsourcing protein modeling problems to an online community through the game Foldit, which develops players' structure prediction skills, is an effective way to solve difficult protein structure modeling problems[5,6]. However, it was unclear whether players' modeling expertise, which relies on human creativity and spatial intuition to direct search through alternative protein structures, could be extended to protein design, which involves a much more open-ended search through protein sequence and structure space.

To explore whether human creativity can guide the search in this substantially larger space, we incorporated new tools allowing insertions, deletions and sequence substitutions into Foldit to supplement the existing tools available for manipulating protein conformation. To integrate players into the experimental design process, we presented them with a series of puzzles. To connect Foldit player iterative exploration with experimental testing, we established an advanced Foldit player as an intermediary between the Foldit community and the experimental laboratory, who presented players with puzzles at each stage of the design process. Using Foldit, the advanced player analyzed the top-ranking community designs and built sequence libraries around the structures to stabilize favorable interactions. The designs were then experimentally tested, and the best were used as input for the next puzzle posted to the online community (**Supplementary Fig. 1**).

We challenged Foldit players to remodel the active-site loops of a computationally designed enzyme, DA_20_10 (ref. 3), that catalyzes the Diels-Alder reaction. This reaction, a cornerstone of organic synthesis, creates two new carbon-carbon bonds and up to four stereocenters in one step. DA_20_10 catalyzes the well-studied reaction between 4-carboxybenzyl trans-1,3-butadiene-1-carbamate (diene) and N,N-dimethylacrylamide (dienophile, **Supplementary Fig. 2**). Despite significant catalytic activity, the DA_20_10 active site is open on one side, leaving the substrates quite exposed to solvent (**Fig. 1a**). We reasoned that redesigning active site loops to make additional contacts with the substrates could improve catalytic activity and specificity. However, the previously developed mass spectrometry–based assay for detecting Diels-Alderase activity allows screening of only ~200 variants at one time, and hence screening large libraries[11] is not feasible. Instead, we chose to enlist Foldit players to guide the search for remodeled loops producing higher activities.

As it was not clear which loop to engineer, the first Foldit puzzle, "Cover the Ligand," asked players to remodel any of four active site loops in DA_20_10 to make additional molecular contacts to the ligand. Players were allowed to add or delete up to five amino acids in addition to mutating residues in the active site (**Supplementary Fig. 1a** and **Supplementary Data**). After a week of game play, the 69,773 designs made by the players were ranked by energy, the lowest 50 were visually assessed, and four designs that made particularly favorable

---

Figure 1 Crystal structure of Foldit design is similar to player model. (**a**) Crystal structure of starting design Diels Alderase. The structure of DA_20_10 has not been solved; shown is the structure of a precursor (3I1C, yellow) which differs by only three amino acid substitutions. The modeled transition state of the Diels-Alder reaction is depicted in cyan. (**b**) Structure of starting Diels Alderase (yellow) overlaid on CE6 player predicted structure (purple). CE6 has a 24-amino-acid-loop remodel, including a 13-residue insertion, converting an unstructured loop into a helix-turn-helix motif. (**c**) Crystal structure of CE6 (3UOS, light blue) overlaid with player predicted structure of CE6 (purple). The helix-turn-helix motif is clearly present and has a position close to that in the predicted structure. (**d**) Correlation between CE6-remodeled helices and corresponding experimental electron density. The electron density map is contoured to 1.0 σ. Side-chain alanine 45 is near the center of the picture on the left, with helix 1 above and helix 2 below. (**e**) The player-designed side chains in CE6 that form the interface between helix 1 (purple) and the transition state model (cyan) are found in the same rotameric conformation as in the crystal structure of CE6 (3UOS, light blue). **Figure 1d** generated in Coot[13]. All other figures generated in PyMol[14].
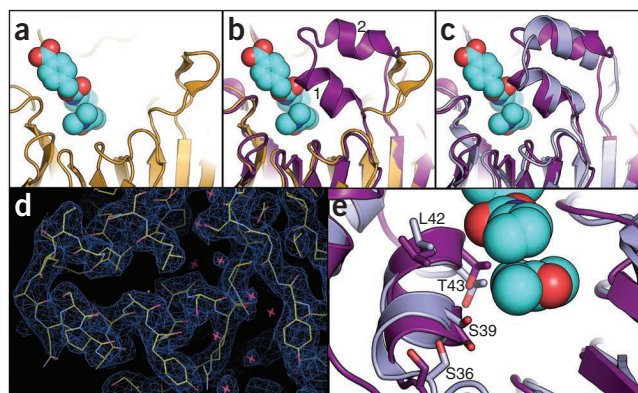
interactions with the ligands were chosen to undergo additional rounds of refinement (**Supplementary Figs. 1b** and **3**). Starting with these four loops, the advanced player designed a library of 36 sequences predicted to interact favorably with the substrates and/or stabilize the designed structure (**Supplementary Library 1** and **Supplementary Fig. 1c**). Whereas most variants exhibited no significant levels of activity, one (CE0) had a catalytic efficiency of $0.5\ s^{-1}M^{-1}M^{-1}$, roughly a tenfold decrease relative to DA_20_10, which has a catalytic efficiency of $4.7\ s^{-1}M^{-1}M^{-1}$ (**Table 1**). We hypothesized that the designed insertion may have the desired structure, but the current amino acids interacting with the substrates or transition state were suboptimal. We explored this design further by making and testing an additional 500 sequence variants predicted to make favorable interactions with the modeled ligands (**Supplementary Library 2**). The most active of these designs, CE4, consisted of a helix buttressing the ligands, followed by an unstructured loop, and is ninefold more active than DA_20_10 with a catalytic efficiency of $42.4\ s^{-1}M^{-1}M^{-1}$ (**Table 1**).

A second puzzle, "Back Me Up," was then posted to the Foldit community asking players to stabilize the initially designed helix by transforming an unstructured loop into an additional neighboring structured helix (**Supplementary Data**). They were allowed to change the structure, sequence and length as before, but only for the unstructured loop. After another week and 109,421 designs, the top designs had converged on a helix-turn-helix motif, as requested (**Supplementary Fig. 1d**). Again, the advanced player constructed two libraries based on the community-designed helix-turn-helix motif, each consisting of roughly 200 sequences (**Supplementary Library 3**). The most active design from these libraries was identified as CE6, with a catalytic efficiency of $87.3\ s^{-1}M^{-1}M^{-1}$ (**Table 1**). This corresponds to a >150-fold increase in activity relative to the initial player-designed model (CE0) and a >18-fold improvement relative to the original enzyme, DA_20_10. The third and final puzzle challenged players to predict the structure of the large insertion in CE6 starting from the crystal structure of the original design (**Supplementary Fig. 1e**). After a week, players generated 335,697

solutions, and the lowest energy of these was selected as the player-predicted structure of CE6 (**Fig. 1b** and **Supplementary Fig. 1f**).

To validate the accuracy of the top-scoring CE6 model, we determined the structure of CE6 by X-ray crystallography (**Supplementary Table 1**). The designed helices are well resolved in the electron density (**Fig. 1d**), and the player-designed helix-turn-helix model is remarkably close to the actual structure (**Fig. 1c** and **Supplementary Fig. 1g**). Helix 1 has the correct secondary structure, register, placement and orientation, resulting in a Cα-r.m.s. deviation of 1.21 Å across the length of the designed helical element (spanning residues 36 to 44 in the design). All three designed residues in the interface between helix 1 and the modeled transition state (Ser 39, Leu 42 and Thr 43) are in the same rotameric conformation as predicted in the final designed model (**Fig. 1e**). In addition, serine 36, which was designed to cap the N terminus of helix 1, is also modeled correctly (**Fig. 1e**).

Helix 2 (which was designed to interact with helix 1, and corresponds to residues 48 to 56 in the designed enzyme) is well ordered in the crystal structure and has the same overall placement as in the Foldit model, but its packing angle and orientation relative to helix 1 differ somewhat from the model. Helix 2 was predicted to have a packing angle of ~30 degrees relative to helix 1, whereas in the crystal structure the two helices are parallel. The difference in the position of the designed helix versus the crystal structure results from a small rotation around the center of helix 2 (near alanine 51).

The backbone r.m.s. deviation over the full 24-residue designed helix-turn-helix motif is 3.13 Å, but most of this increased r.m.s. deviation is a result of the shifted orientation of helix 2. Some of the differences between the final design and the experimentally determined crystal structure are located near crystal contacts with the C terminus of a second molecule of CE6 in the asymmetric unit (**Supplementary Fig. 4**). To evaluate whether the observed crystal contacts occur in solution, we mutated the most buried residues in the crystal interface (F324R, I326G and F327K). The activity of this mutant was indistinguishable from that of CE6, suggesting that the interface does not form in solution (**Supplementary Fig. 5**).

The Michaelis constants, $K_{M\text{-diene}}$ and $K_{M\text{-dienophile}}$ (**Table 1** and **Fig. 2**), of CE6 are improved sixfold and threefold, respectively, compared to the starting design, but the turnover number (k_{cat}) is unchanged.

Table 1 Kinetic constants derived from the nonlinear regression fits in Figure 2

Protein	$K_{M\text{-diene}}$ (mM)	$K_{M\text{-dienophile}}$ (mM)	k_{cat} (h^{-1})	$k_{cat}/(K_{M\text{-diene}} * K_{M\text{-dienophile}})$ (s^{-1}M^{-1}M^{-1})
DA2010	1.2 ± 0.2	101 ± 21	2.1 ± 0.3	4.7 ± 1.5
CE0	n.d.	n.d.	n.d.	0.5 ± 0.05
CE4	0.5 ± 0.03	31 ± 3.0	2.4 ± 0.1	42.4 ± 5.7
CE6	0.2 ± 0.03	35 ± 1.4	2.2 ± 0.1	87.3 ± 13.9

k_{uncat} under these conditions is 2.2×10^{-2} M^{-1} h^{-1}. n.d., not detectable.
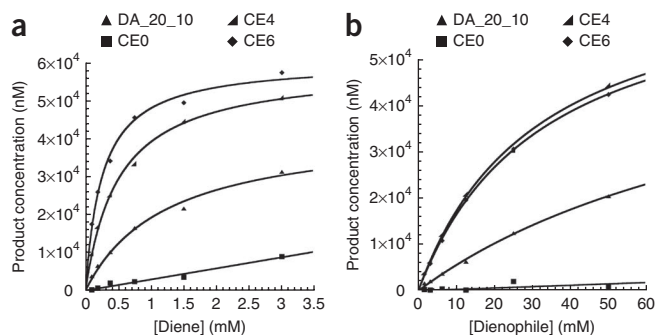
Figure 2 Activity of designed enzymes DA_20_10, CE0, CE4 and CE6. (**a**) Product observed after 3 h with the dieneophile held constant at 50 mM and the diene at 3, 1.5, 0.8, 0.4, 0.2 and 0.1 mM. (**b**) Product observed after 3 h with the diene held constant at 3 mM and the dienophile at 50, 25, 12, 6, 3 and 1.5 mM. The kinetic constants were determined using 10 μM purified protein (**Supplementary Fig. 8**) and a standard Michaelis-Menten assay as described in the **Supplementary Methods**.

The improvement in K_M but lack of change in k_{cat} are consistent with the design model and crystal structure. The designed loop interacts with and likely increases affinity for both substrates consistent with the decrease in K_M. The catalytic residues that stabilize the transition state are on the opposite side of the active site, and these residues have almost identical locations in CE6 and the original design (**Supplementary Fig. 6**) despite the large-scale remodeling of the loop. Given this similarity, and the fact that the conformation of the substrates and transition state are very similar in the region of the designed loop, it is not surprising that, as suggested by the lack of change in k_{cat}, the loop does not selectively stabilize the transition state.

In addition to increases in activity, we hypothesized that the increased buried surface area and hydrophobicity of the binding pocket would increase dienophile specificity for other hydrophobic substrates. We tested this hypothesis by assaying enzyme activity of CE6 with a series of modified dienophiles previously described[3]. Relative to DA_20_10, the player-designed CE6 exhibited a threefold increase in binding specificity for the hydrophobic dienophile 2A over the hydrophilic dienophile 2E (**Supplementary Fig. 7**), consistent with our prediction. However, CE6 shows no significant preference for 2A compared to the similar-sized hydrophobic 2B and 2C substrates. The increase in specificity for hydrophilic substrates and loss of specificity for hydrophobic substrates suggests that although the desired hydrophobic pocket is formed, further improvements to the shape complementarity between the substrate and the engineered enzyme remain possible. Given the new backbone structure of the active site, future studies will explore additional backbone remodeling to modulate substrate specificity.

Insertion of helix-turn-helix motifs may be broadly useful in computational protein design. An advantage of helical hairpins is that they are to a large extent self-stabilizing and do not require additional tertiary interactions to form[12]. This allows most of the experimental sampling to be focused on introducing new functional interactions with ligands. The highly ordered and predictable helical register enables sampling to be focused on a small subset of positions predicted to be pointing directly toward the ligand of interest.

We have demonstrated that crowdsourcing complex computational protein-design problems with a community of game-developed experts can be an effective way of creatively sampling the potential sequence space for the design of active-site loops that modulate enzyme activity. To our knowledge, this is the most extensive remodeling by design of a

functional protein structure to date, and was accomplished by screening fewer than 1,000 sequences. The Foldit community's ability to successfully guide large-scale protein design problems demonstrates that biophysics expertise developed through game playing transfers effectively from structure prediction to more open-ended design challenges.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Accession code. The X-ray crystallographic coordinates have been deposited in the Protein Data Bank with accession ID 3U0S.

Note: Supplementary information is available on the Nature Biotechnology website.

AUTHOR CONTRIBUTIONS
C.B.E. analyzed community models, in addition to designing, building and testing the enzyme libraries. J.B.S. designed the experimental and computational methods, and built the initial computational models. F.K. set up the Foldit puzzles and curated the player results for analysis by C.B.E. S.C. led design and development of Foldit. B.L.S., J.B.B. and B.W.S. grew the crystals and collected X-ray diffraction data, processed X-ray data and analyzed the structure. Foldit Players designed new protein backbones and sequences. Z.P. and D.B. contributed to the writing of the manuscript.

1. Rothlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
2. Jiang, L. *et al. De novo* computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008).
3. Siegel, J.B. *et al.* Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313 (2010).
4. Bar-Even, A. *et al.* The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
5. Khatib, F. *et al.* Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat. Struct. Mol. Biol.* **18**, 1175–1177 (2011).
6. Cooper, S. *et al.* Predicting protein structures with a multiplayer online game. *Nature* **466**, 756–760 (2010).
7. Sterner, R. & Hocker, B. Catalytic versatility, stability, and evolution of the (βα)8-barrel enzyme fold. *Chem. Rev.* **105**, 4038–4055 (2005).
8. Hu, X., Wang, H., Ke, H. & Kuhlman, B. High-resolution design of a protein loop. *Proc. Natl. Acad. Sci. USA* **104**, 17668–17673 (2007).
9. Murphy, P.M., Bolduc, J.M., Gallaher, J.L., Stoddard, B.L. & Baker, D. Alteration of enzyme specificity by computational loop remodeling and design. *Proc. Natl. Acad. Sci. USA* **106**, 9215–9220 (2009).
10. Havranek, J.J. & Baker, D. Motif-directed flexible backbone design of functional interactions. *Protein Sci.* **18**, 1293–1305 (2009).
11. Park, H.S. *et al.* Design and evolution of new catalytic activity with an existing protein scaffold. *Science* **311**, 535–538 (2006).
12. Religa, T.L. *et al.* The helix-turn-helix motif as an ultrafast independently folding domain: the pathway of folding of Engrailed homeodomain. *Proc. Natl. Acad. Sci. USA* **104**, 9272–9277 (2007).
13. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
14. The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.

ONLINE METHODS

Library whole cell lysate assay. Libraries were screened using a whole cell lysate assay. Single colonies were isolated from Petri dishes and used to start cultures in 0.5 ml lysogeny broth (LB) in 2 ml 96-deep-well plates with 50 µg/ml kanamycin. Plates were grown with shaking at 1,250 r.m.p. in a plate shaker at 37 °C for 16–24 h. We used 20 µl of overnight to inoculate 1 ml of terrific broth (TB) with 50 µg/ml kanamycin and allowed to grow for 2.5 h at 37 °C shaking at 1,250 r.m.p. After 2.5 h, cells were induced with 50 µl of 10 mM isopropyl β-D-1-thiogalactopyranoside and transferred to an 18 °C shaker for 24 h. Cells were harvested by centrifugation at 2,900g for 20 min and the supernatant decanted. Cells were resuspended in 250 µl of lysis buffer consisting of 2 mg/ml lysozyme isolated from egg white, 0.2 mg/ml DNAse, 1 mM phenylmethanesulfonylfluoride and 1× PBS. Lysis was performed either by three cycles of freeze thaw (from −80 °C to 25 °C) or by means of bath sonication (Ultrasonic Processor Gex750) at 70% power for 1 min, pulsing 20 s on, 40 s off. Whole cell lysate was then spun down at 2,900g for 20 min, and supernatant was used for product formation assay.

Diels-Alder product formation assay. The Diels-Alder reactions were done in PBS, at 25 °C, with 4% DMSO. The DMSO came from the diene stock, for which the stock solution (100 mM) was made in DMSO and always diluted such that the final reaction would contain 4% DMSO. After incubation of the desired amount of diene and dienophile in the presence or absence of protein, 5 µl aliquots were taken at various time points and quenched in a solution of 80:20 acetonitrile/water with 0.1% formic acid and 1 mM benzoic acid (95 µl). This solution was incubated for 5 min and then the precipitated protein was filtered using a Millipore multiscreen solvinert filter plate. We subsequently injected 20 µl of the filtered quench solution and analyzed it using a liquid chromatography–tandem mass spectrometry (LC-MSMS) assay. The column used for the chromatography run was a Hypersil Gold C18 (dimensions 100 × 2.1, 1.9 µm particle size, Thermo). The following gradient was preformed for the chromatography run at a flow of 500 µl/min: 95:5 water:acetonitrile (0.1% formic acid) for 30 s, followed by a gradient over 4.5 min ending at 5:95 water/acetonitrile (0.1% formic acid), ending by switching back to 95:5 water/acetonitrile (0.1% formic acid) and letting the column re-equilibrate for 1 min before the next injection. The Diels-Alder reaction product and benzoic acid (used as an internal standard for injection differences and sample evaporation) was detected using a TSQ Quantum Access (Thermo Fisher Scientific) with an ESI probe to ionize the parent ion followed by fragmentation using Argon gas with the following parameters (tuned on the diene product standard, synthesized as previously described[3]):

    Spray voltage: −4,000
    Sheath gas pressure = 20
    Aux gas pressure = 5
    Capillary temperature = 300
    Chrom filter = 10
    Collision pressure = 1.5
    Skimmer offset = 10
    Scan width = 0.01
    Scanned in negative centroid mode
    Benzoic acid specific parameters
    Parent mass = 120.994
    Product mass = 120.994
    Scan time = 0.08
    Collision energy = 5
    Tube lens = 112
    Diels-Alder product specific parameters
    Parent mass = 345.150
    Product mass = 211.033
    Scan time = 0.2
    Collision energy = 21
    Tube lens = 140

Mutagenesis. All mutagenesis was done following the Kunkel protocol[15], using oligonucleotides designed with Agilent Technologies's primer design online tool (https://www.genomics.agilent.com/CollectionSubpage. aspx?PageType=Tool&SubPageType=ToolQCPD&PageID=15). To design

degenerative primers, we used the AA-Calculator (http://guinevere.otago. ac.nz/cgi-bin/aef/AA-Calculator.pl).

Kunkel mutagenesis reactions were transformed directly into Stratagene BL21DE3*, and colonies were picked and screened for Diels-Alderase activity using the whole cell lysate assay described above. For each library we screened enough colonies to cover the theoretical library by fourfold.

Kinetic characterization. The most promising designs (CE0, CE4, CE6 and DA_20_10) were grown up at 0.5 L scale and purified as previously described[16]. Protein concentrations were determined by measuring their absorbance at 280 nm, and using their calculated extinction coefficients to convert the absorbance into molarity. After normalization the proteins concentrations were verified to be equivalent by running them on an SDS-PAGE gel (**Supplementary Figure 7**). Kinetics were measured by holding either the diene or dienophile constant and varying the other in the presence or absence of enzyme. Reactions were incubated at 298K in PBS, 4% DMSO, pH 7.4, 298K, 10 µM protein. Triplicate samples were taken after 3-h points from which the amount of product formation was determined using the LC-MSMS assay described earlier. For obtaining kinetic constants we used the standard Michaelis-Menten model to fit each curve, with the exception of CE0, which was fit to a linear slope as no saturation was observed. (**Fig. 2**). The k_{cat} was determined from the calculated maximum product production from the dienophile curve, and by dividing by the amount of enzyme and hours the sample was incubated before quenching. A linear product production rate was assumed as the substrate depletion was <5% total substrate. Errors for $k_{cat}/(K_{M\text{-diene}} * K_{M\text{-dienophile}})$ were calculated using the standard error propagation formula: $\delta k_{cat}/(K_{M\text{-diene}} * K_{M\text{-dienophile}}) = |k_{cat}/(K_{M\text{-diene}}*K_{M\text{-dienophile}})| * \sqrt{((\delta k_{cat}/k_{cat})^2 + (\delta K_{M\text{-diene}}/K_{M\text{-diene}})^2 + (\delta K_{M\text{-dienophile}}/K_{M\text{-dienophile}})^2)}$.

Substrate specificity. To experimentally determine substrate specificities, each dienophile from **Supplementary Figure 6** was either purchased or synthesized, as previously described[16]. The dienophiles were assayed in the presence or absence of 5 µM CE6 or DA_20_10 with 0.1 mM diene and 5 mM dienophile at 298 K in a PBS solution with 4% DMSO. Samples were taken in triplicate at 3 h, quenched and filtered as previously described. To detect product, a liquid chromatography–mass spectrometry assay was used. The mass spectrometer was used in a selective-ion-monitoring mode and set to specifically detect the expected product from each reaction. The same tuning parameters were used as previously described. The average area of each detected peak formed per hour was used to determine the relative amount of product formed with and without enzyme. The relative peak areas formed per hour are reported as **Supplementary Figure 6**.

Crystallographic methods: data collection and refinement. CE6 was expressed in *Escherichia coli* BL21DE3* cells (Stratagene) and purified as previously described[16]. Purified protein was dialyzed into crystallization buffer containing 25 mM HEPES pH 7.25, 100 mM NaCl and 5% glycerol, and concentrated to ~30 mg/ml.

CE6 crystals were grown at 18 °C by the hanging drop vapor diffusion method. A 15 mg/ml solution was prepared of CE6 in crystallization buffer with Trypsin (Sigma T1426) added in a 1:10,000 molar ratio as compared to CE6. Approximately 20 min later, drops were set containing 1 µl of this protein solution and 1 µl of mother liquor (100 mM HEPES pH 7.5, 2% vol/vol PEG-400, and 2.0 M $(NH_4)_2SO_4$) over a reservoir volume of 0.3 ml. Crystal growth was observed within 1 week. Crystals were cyroprotected in a mixture of 75% mother liquor with 25% ethylene glycol and flash frozen in liquid nitrogen.

A data set was collected at the Advanced Light Source on BL5.0.1 (at a temperature of 100 K and a wavelength of 0.97740 Å) and processed using HKL2000 (ref. 17) (**Supplementary Table 1** for statistics). Phases were determined by molecular replacement (MR) using the program PHASER[18] provided in the Collaborative Computational Project No. 4 software suite (CCP4)[19]. The MR search model was derived from the wild-type 1E1A structure[20] by removing all water molecules and ligands, 6 residues of the unstructured C terminus, and residues 36-44 of the redesigned active site loop. Following MR, the missing residues and ligands of CE6 were built into the model using the molecular modeling software Coot[21,22], and refinement was performed

using the CCP4 provided version of REFMAC5 (refs. 23–25). The C-terminal HIS-tag and linker region were not visible in the electron density and thus were not modeled. Alternate conformations were modeled for some surface exposed residues, including LYS44 and ARG56 of the redesigned loop. Also of note, a HEPES molecule was modeled into the active site in both chains of the asymmetric unit. Ramachandran statistics (**Supplementary Table 1**) were generated using the online version of PROCHECK[26] provided by the JCSG server found here: http://www.jcsg.org/scripts/prod/validation/sv_final.cgi.

Coordinate and data deposition. Structure factors and coordinates have been deposited in the RCSB database under PDB ID code 3U0S.

15. Kunkel, T.A. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**, 488–492 (1985).
16. Siegel, J.B. *et al.* Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313 (2010).
17. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods In Enzymology: Macromolecular Crystallography. Part A* **276**, 307–326 (1997).
18. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
19. Collaborative Computational Project, Number 4. The *CCP4* Suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
20. Scharff, E.I., Koepke, J., Fritzsch, G., Luecke, C. & Rueterjans, H. Crystal structure of diisopropylfluorophosphatase from Loligo vulgaris. *Structure* **9**, 493–502 (2001).
21. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
22. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).
23. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
24. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 122–133 (2001).
25. Vagin, A.A. *et al.* Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–2195 (2004).
26. Laskowski, R.A., Macarthur, M.W., Moss, D.S. & Thornton, J.M. *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).