

A Completely *De Novo* ATPase from Combinatorial Protein Design

Michael S. Wang and Michael H. Hecht*

Cite This: *J. Am. Chem. Soc.* 2020, 142, 15230–15234

Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: Our understanding of biological chemistry is shaped by the observation that all life comes from other life—as Pasteur put it, *omne vivum ex vivo*. A key step in expanding our biochemical vocabulary is to recapitulate biogenic catalysis using non-natural sequences that did not arise from common ancestry. Here we describe an enzyme designed completely *de novo* that hydrolyzes ATP. This protein was designed to lack β -sheet structure and is competitively inhibited by magnesium, two traits that are unlike natural ATPases.

The hydrolysis of adenosine triphosphate (ATP) fuels many life-essential functions. Yet the molecular toolkit used to accomplish this reaction is surprisingly limited. Sequence conservation among ATP hydrolases is dominated by the phosphate-loop (P-loop) motif (consensus: GXXXXGK[TS]), which is found across all kingdoms of life.^{1–4} The structural features constraining this motif are local rather than long range, thereby enabling short peptides modeled after the P-loop to be structurally ordered and capable of binding phosphate.⁵ Moreover, grafting a P-loop cassette onto an artificial protein scaffold is sufficient to facilitate ATP binding and hydrolysis.¹ Structural diversity in this function remains limited in that, although phospho-ligand binding emerged >250 independent times in ancient proteins, the P-loop was the convergent solution.⁶ And though other ATP-hydrolyzing motifs occur occasionally in nature (e.g., among tyrosine kinases), all natural ATPases contain β -sheet structure.⁷ Does this limited diversity indicate that only certain privileged sequences and structures can accomplish ATP hydrolysis? Or alternatively, does the similarity of natural ATPases simply reflect evolution from common ancestry?

The limits of nature's catalytic repertoire can be understood and expanded by designing novel protein sequences.^{8–10} To test whether artificial proteins containing neither P-loops nor β -structure can also hydrolyze ATP, we examined a collection of novel sequences designed to fold into 4-helix bundles. This collection was culled from a library that had been constructed previously using the binary code for protein design.^{11–13} The binary code specifies the sequence positions of polar and nonpolar amino acids to correspond with surface and buried locations, respectively, on the structure of an idealized 4-helix bundle. However, the binary code does not specify side chain identities; instead, it allows surface positions to sample several polar side chains and buried positions to sample several hydrophobic side chains, resulting in enormous sequence diversity. Although this library of binary patterned sequences was designed explicitly for structure, with no consideration of enzymatic function, we have shown previously that novel proteins capable of functional activity—both *in vitro* and *in*

vivo—can readily be isolated from a library containing 10⁶ sequences.^{14,15}

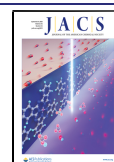
Motivated by the occurrence of other active proteins in these libraries, we asked whether binary patterning might also yield α -helical proteins capable of ATP hydrolysis. Initially, we considered screening lysates from cultures expressing each sequence in the library. However, two factors prevent high-throughput screening for ATP hydrolases in lysates: First, *E. coli* contains natural enzymes that would overwhelm any signal from a novel enzyme with low activity. Second, lysates contain nucleotides and free phosphate, which interfere with assays for ATPase activity. Therefore, we employed a lower throughput two-step screen: First, we screened lysates for the ability to cleave *para*-nitrophenyl (*p*NP) esters (Figure 1). Second, promising candidates were purified and assayed for ATPase activity.

We demonstrated previously that proteins capable of hydrolyzing *p*NP esters are not uncommon among binary patterned libraries.¹⁵ Here we built on this finding by screening lysates expressing >1100 novel sequences for the ability to hydrolyze either *p*NP-palmitate or *p*NP-phosphate. From those displaying hydrolytic activity, we chose five well-expressed proteins for purification and assessment of ATPase activity. One of these proteins showed ATP hydrolysis activity substantially above background (Figures 1 and S1–S3). Because this protein represents a structural and sequential alternative to natural ATPases, we named it Alternative ATPase (AltTPase).

AltTPase was identified based on its ability to hydrolyze *p*NP-palmitate and ATP. Interestingly, it acts on all nucleotide triphosphates and shows no preference for ATP over dATP (Figure 2a). However, this promiscuity does not mean nonspecificity. Separation of ATP products by ion-pair

Received: March 16, 2020

Published: August 24, 2020



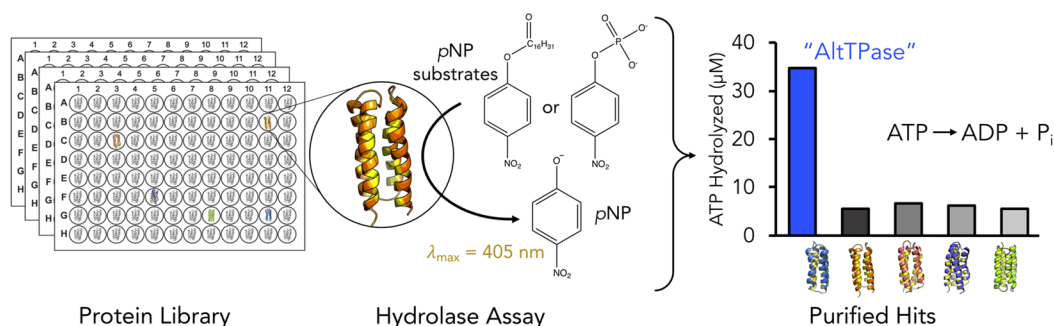


Figure 1. Strategy to search for ATPases. First, a library of *de novo* proteins expressed in *E. coli* was searched for hydrolase activity using chromogenic *para*-nitrophenyl (*pNP*) substrates. Hits from this screen were purified and tested for ATPase activity by incubating 25 μM protein with 100 μM ATP for 24 h and monitoring phosphate appearance by a Malachite green assay.

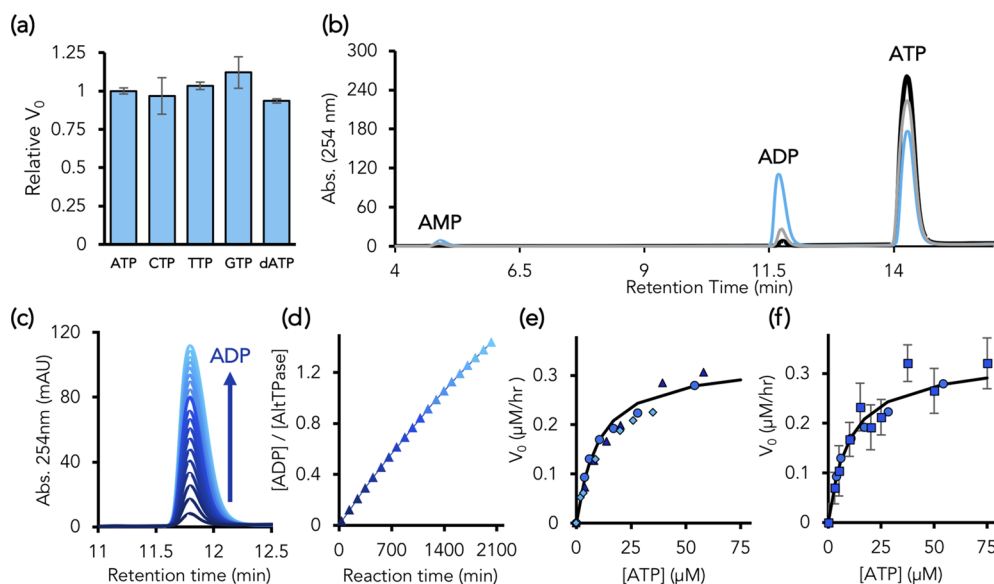


Figure 2. Kinetic characterization of AltTPase. (a) Kinetics of 10 μM AltTPase with 50 μM NTPs showing enzyme promiscuity. (b) Ion pair HPLC chromatogram for 50 μM ATP at 0 h (black), after 34 h without protein (gray), and after 34 h with 10 μM AltTPase (blue). (c) Ion-pair HPLC data showing a time course of ADP appearance when 50 μM ATP reacts with 10 μM AltTPase. (d) ADP generated per protein in (c) with background hydrolysis subtracted showing multiple turnovers. (e) Michaelis–Menten AltTPase kinetics for 10 μM protein in biological triplicate (different shapes) determined by ion-pair HPLC. (f) Michaelis–Menten AltTPase kinetics for 10 μM AltTPase measured by Malachite green assay (squares) or ion-pair HPLC (circles) with the best fit shown as a solid line.

HPLC shows that AltTPase converts ATP to ADP and does not further hydrolyze ADP, as would be seen for a nonspecific phosphatase (Figure 2b). A similar observation holds for the other NTPs (Figures S4 and S5). Additionally, it is more active with *pNP*-palmitate than *pNP*-phosphate (Figure S2b).

Next, the ATPase kinetics were characterized using ion-pair HPLC. This showed multiple turnovers (Figure 2c,d) and reproducible activity in biological triplicate (Figure 2e). Kinetics assessed using a chromogenic Malachite green assay to quantify accumulation of inorganic phosphate, the other enzymatic product, agreed with the ion-pair HPLC characterization (Figure 2f).

Because the observed ATPase activity is relatively low, we considered the possibility that the activity may result from contamination by small amounts of an endogenous *E. coli* ATPase. We addressed this concern in several ways: First, the binary patterned proteins were purified using two orthogonal steps—metal affinity and size exclusion chromatography—and a high level of purity was verified by reverse phase HPLC and mass spectrometry (Figure S6). Second, we used the same two

orthogonal steps to purify several other binary patterned proteins and assayed their activities (Figure 1). None of these other proteins showed significant ATPase activity, indicating that endogenous *E. coli* ATPases are unlikely to co-purify with binary patterned proteins. Third, we purified three independent preparations of AltTPase and showed they had similar ATPase activity (Figure 2e). Fourth, the reaction was specific in converting ATP to ADP (Figure 2b). This precludes a generic phosphatase contaminant that would produce AMP and adenosine. Fifth, as described below, the activity of AltTPase is inhibited by Mg^{II} , which is the opposite of what would be expected for activity arising from a natural ATPase. Finally, to explicitly test whether the observed activity is due to the novel protein sequence, we tested whether changes in sequence produced changes in activity. As described below, mutating single amino acids substantially lowered activity. Altogether, these results provide compelling evidence that the AltTPase protein is responsible for the observed activity.

The purified AltTPase protein hydrolyzes ATP with a $k_{\text{cat}}/K_{\text{M}}$ of $1.06 \text{ M}^{-1} \text{ s}^{-1}$, nearly 100-fold above background ($k_{\text{cat}}/$

$k_{\text{uncat}} = 98$). This activity is significantly slower than those of natural ATPase enzymes, which can reach $k_{\text{cat}}/k_{\text{uncat}} > 10^7$.¹⁶ The greater activity of natural enzymes relative to AltTPase is not surprising: Nature has relied on billions of years of life-or-death selections to evolve fast enzymes. In contrast, AltTPase is an unevolved protein chosen from a collection of >1100 sequences, which were designed to fold into helical bundles without any consideration of catalytic activity.

After establishing the activity of AltTPase, we investigated the relationship between sequence, structure, and activity in this novel protein. When studying natural proteins, such investigations typically begin by searching for homologous proteins. However, by design, the binary code proteins share no ancestry with naturally evolved sequences. Therefore, we relied on computational methods to predict the structure and guide the search for a putative active site. Structure prediction by I-Tasser^{17,18} produced a 4-helix bundle similar to that expected from our designed binary pattern (Figures 3b and S7)

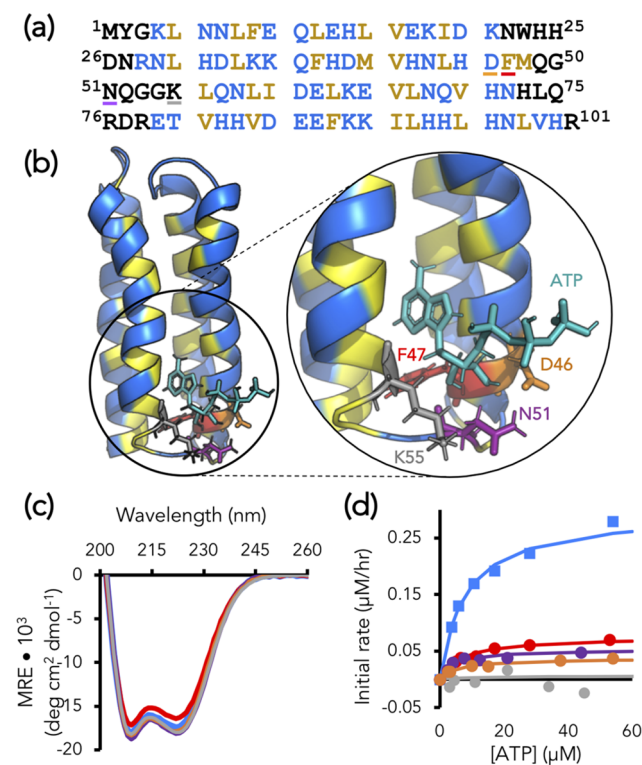


Figure 3. Mutagenesis of the predicted active site of AltTPase. (a) AltTPase sequence, with mutated residues underlined. (b) Predicted docking of ATP into AltTPase. The mutated residues are identified and color-coded in the magnified view. (c) Overlaid CD spectra show nearly identical helical secondary structure for the parent and mutant proteins. (d) Michaelis–Menten curves for 10 μM protein in 10 mM Tris (pH 7.2) based on assays using ion-pair HPLC after 7 h. Strongly disrupting mutations are shown by circles.

Next, we tested the validity of this prediction with a series of biophysical methods. Circular dichroism (CD) showed that the secondary structure of AltTPase matches that of previously determined 4-helix bundle structures from our libraries¹¹ (Figures 3c and S8). Next, we used size-exclusion chromatography and analytical ultracentrifugation to probe the oligomeric state of AltTPase (Figure S9). Both methods showed monomers, and no oligomeric species were detected.

Together, these results support the predicted monomeric 4-helix bundle structure.

With this predicted structure as a template, we searched for an active site using ROSIE^{19–22} to dock ATP into the structure (Figure 3b). Results from these predictions were used to guide site-directed mutagenesis of several putative binding sites. Mutations in some of the predicted sites had no impact on enzyme kinetics, so these sites were rejected. However, a promising ATP binding site was identified at the turn between the second and third α-helices (Figure 3b). Several mutations in this putative active site reduced catalytic activity (Figure S10). Four mutations in particular—D46S, F47L, N51S, and K55E—diminished catalytic activity substantially (Figure 3d, Table 1). Notably, these mutations did not alter the polar/

Table 1. Kinetic Parameters of AltTPase Mutants with Standard Errors

	k_{cat} (h ⁻¹)	K_{M} (μM)	$k_{\text{cat}}/k_{\text{uncat}}$
AltTPase	0.035 ± 0.003	10.2 ± 1.3	98
F47L	0.0075 ± 0.0006	6.8 ± 0.9	21
N51S	0.005 ± 0.001	4.3 ± 1.2	14
D46S	0.0036 ± 0.0004	5 ± 1	10
K55E	undetectable	—	—

nonpolar binary code of the sequence, nor did they alter the secondary structure of the protein as measured by CD (Figure 3c). Therefore, we suggest these mutations diminish activity by disrupting the active site of the novel enzyme. K55 in particular is likely the key positive residue for interacting with the ATP phosphate tail. Yet K55 is constant in the library design, so other sequence considerations likely make the AltTPase more active than other tested proteins.

Having shown that the activity of AltTPase depends on its sequence, we next tested how this activity depends on the surrounding solution conditions. Natural enzymes that work on nucleic acids are often more active when Mg^{II} is present in solution because nucleic acids and nucleotides are typically complexed with Mg^{II} *in vivo*. For example, Mg^{II} is present in *E. coli* at concentrations of 20–100 mM.²³ However, to our surprise, AltTPase was inhibited by Mg^{II}, and analysis of the enzyme kinetics revealed competitive inhibition (Figure 4). Moreover, this inhibition occurred without changing the oligomeric state of the protein (Figure S11). Similar inhibition

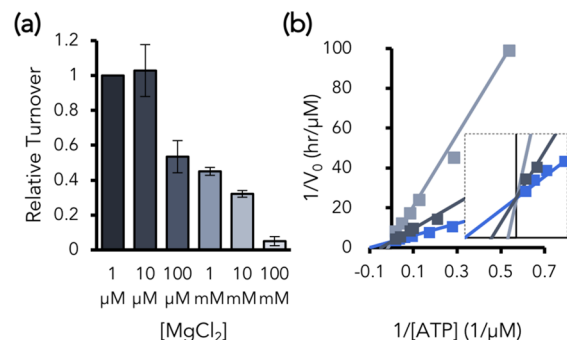


Figure 4. AltTPase inhibition by magnesium as measured by ion-pair HPLC. (a) 50 μM ATP was incubated with 10 μM AltTPase or a protein-free blank in 10 mM Tris (pH 7.2) for 7 h. (b) Lineweaver–Burke plots of enzyme kinetics with 10 mM MgCl₂ (gray), 50 μM MgCl₂ (dark blue), and without added MgCl₂ (blue) show that magnesium is a competitive inhibitor.

was observed for Ca^{II} and Zn^{II} , and to a lesser extent for Na^+ (Figure S12). It is possible that these metals compete with K55 for ionic interactions with the ATP phosphate tail. Yet inhibition by Mg^{II} in particular highlights how natural selection operates in a specific physiologically defined environment to impose functional constraints that need not apply to proteins designed *de novo*. Indeed, given the high concentration of magnesium in biological systems, AltTPase could not have evolved as a functional enzyme in nature.

Because AltTPase is slower than natural enzymes, we also compared its activity to three previous attempts to find ATPase activity in semi-natural or non-natural sequences: Romero et al. constructed a chimeric protein by grafting a consensus P-loop based on natural sequences into the structure of designed beta/alpha repeat proteins.¹ The chimeric protein bound ATP and in some experiments—but not others—hydrolyzed one ATP per 30 min. That work demonstrated that natural P-loops can function as cassettes that can be moved into different structural contexts. An example of a completely non-natural sequence that catalyzes ATP hydrolysis was described by Simmons et al., who characterized a protein selected from a random sequence library for its ability to bind ATP. Surprisingly, the crystal structure of this protein revealed ADP, rather than ATP, suggesting that the crystallized form had some level of ATPase activity.²⁴ Finally, Monasterio et al. reported a novel sequence that aggregates into an amyloid-like structure and has a very low level of Zn^{II} -dependent ATPase activity.²⁵ We compare AltTPase to these previous studies by noting that the first example relied on sequences derived from natural P-loops, while the second and third examples found weak ATP hydrolysis only in the context of insoluble proteins (crystals or amyloid). Thus, AltTPase is the first example of a novel enzyme capable of hydrolyzing ATP *in solution*.

The studies described herein were motivated by the observation that although ATP hydrolysis is central to energy metabolism in all forms of life, it is catalyzed by a relatively limited diversity of local sequences and structures. The lack of diversity among natural ATPases might indicate that appropriate binding and catalysis cannot be accomplished by alternative sequences and structures. Alternatively, diversity might be limited simply because natural sequences all arose from common ancestry. To distinguish between these possibilities, we probed a collection of novel proteins, which (a) fold into fundamentally different structures than natural ATPases and (b) share no ancestry with natural sequences. We screened >1100 binary patterned α -helical proteins for simple hydrolytic activity and interrogated five of the most-active sequences for the ability to hydrolyze ATP. From this small collection, we isolated AltTPase, a novel protein that hydrolyzes ATP at a rate ~100-fold above background and is competitively inhibited by ATP's most common biological counterion, Mg^{II} . Finding this enzyme in such a small collection of novel sequences suggests that some level of ATPase activity is not uncommon in protein sequence space and, moreover, that it can occur in sequences and structures that differ dramatically from natural ATPases. Together with previous results demonstrating that novel proteins can provide a range of life-sustaining regulatory and enzymatic functions,^{14,15,26–29} these results demonstrate the vast potential of protein sequence space to sustain life in ways that are alternative to those explored by nature.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.0c02954>.

Detailed materials and methods and supplemental Table S1 and Figures S1–S15 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Michael H. Hecht — Department of Chemistry, Princeton University, Princeton, New Jersey 08540, United States; orcid.org/0000-0002-5538-9813; Email: hecht@princeton.edu

Author

Michael S. Wang — Department of Chemistry, Princeton University, Princeton, New Jersey 08540, United States; orcid.org/0000-0002-1487-5719

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/jacs.0c02954>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Venu Vandavasi for assistance with AUC, Kaelix Johnson for assistance making some constructs, and Shlomo Zarzhitsky for helpful feedback throughout. This work was supported by NSF grants MCB-1409402 and MCB-1947720.

■ REFERENCES

- (1) Romero, M. L. R.; Yang, F.; Lin, Y.; Toth-Petroczy, A.; Berezhovsky, I. N.; Goncearenco, A.; Yang, W.; Wellner, A.; Kumar-Deshmukh, F.; Sharon, M.; Baker, D.; Varani, G.; Tawfik, D. S. Simple yet functional phosphate-loop proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E11943–E11950.
- (2) Saraste, M.; Sibbald, P. R.; Wittinghofer, A. The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **1990**, *15*, 430–434.
- (3) Davidson, A. L.; Dassa, E.; Orelle, C.; Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 317–364.
- (4) Milner-White, E. J. Protein three-dimensional structures at the origin of life. *Interface Focus* **2019**, *9*, 20190057.
- (5) Bianchi, A.; Giorgi, C.; Ruzza, P.; Toniolo, C.; Milner-White, E. J. A synthetic hexapeptide designed to resemble a proteinaceous p-loop nest is shown to bind inorganic phosphate. *Proteins: Struct., Funct., Genet.* **2012**, *80*, 1418–1424.
- (6) Longo, L. M.; Petrović, D.; Kamerlin, S. C. L.; Tawfik, D. S. Short and simple sequences favored the emergence of N-helix phospho-ligand binding sites in the first enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 5310–5318.
- (7) Hanson, P. I.; Whiteheart, S. W. AAA proteins: have engine, will work. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 519.
- (8) 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. Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190–195.
- (9) Kries, H.; Blomberg, R.; Hilvert, D. De novo enzymes by computational design. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221–228.
- (10) Marshall, L. R.; Zozulia, O.; Lengyel-Zhand, Z.; Korendovych, I. V. Minimalist de Novo Design of Protein Catalysts. *ACS Catal.* **2019**, *9*, 9265–9275.

- (11) Wei, Y.; Kim, S.; Fela, D.; Baum, J.; Hecht, M. H. Solution structure of a de novo protein from a designed combinatorial library. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13270–13273.
- (12) Wei, Y.; Liu, T.; Sazinsky, S. L.; Moffet, D. A.; Pelczer, I.; Hecht, M. H. Stably folded de novo proteins from a designed combinatorial library. *Protein Sci.* **2003**, *12*, 92–102.
- (13) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. Protein design by binary patterning of polar and nonpolar amino acids. *Science* **1993**, *262*, 1680–1685.
- (14) Donnelly, A. E.; Murphy, G. S.; Digianantonio, K. M.; Hecht, M. H. A de novo enzyme catalyzes a life-sustaining reaction in *Escherichia coli*. *Nat. Chem. Biol.* **2018**, *14*, 253.
- (15) Patel, S. C.; Bradley, L. H.; Jinadasa, S. P.; Hecht, M. H. Cofactor binding and enzymatic activity in an unevolved superfamily of de novo designed 4-helix bundle proteins. *Protein Sci.* **2009**, *18*, 1388–1400.
- (16) Wolfenden, R.; Snider, M. J. The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **2001**, *34*, 938–945.
- (17) Xu, D.; Zhang, Y. Toward optimal fragment generations for ab initio protein structure assembly. *Proteins: Struct., Funct., Genet.* **2013**, *81*, 229–239.
- (18) Xu, D.; Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Struct., Funct., Genet.* **2012**, *80*, 1715–1735.
- (19) Lyskov, S.; Chou, F.; Conchúir, S. Ó.; Der, B. S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B. D.; Renfrew, P. D.; Sripakdeevong, P.; et al. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS One* **2013**, *8*, No. e63906.
- (20) Kothiwale, S.; Mendenhall, J. L.; Meiler, J. BCL:: Conf: small molecule conformational sampling using a knowledge based rotamer library. *J. Cheminf.* **2015**, *7*, 47.
- (21) Combs, S. A.; DeLuca, S. L.; DeLuca, S. H.; Lemmon, G. H.; Nannemann, D. P.; Nguyen, E. D.; Willis, J. R.; Sheehan, J. H.; Meiler, J. Small-molecule ligand docking into comparative models with Rosetta. *Nat. Protoc.* **2013**, *8*, 1277.
- (22) DeLuca, S.; Khar, K.; Meiler, J. Fully flexible docking of medium sized ligand libraries with RosettaLigand. *PLoS One* **2015**, *10*, No. e0132508.
- (23) Moncany, M.; Kellenberger, E. High magnesium content of *Escherichia coli* B. *Experientia* **1981**, *37*, 846–847.
- (24) Simmons, C. R.; Stomel, J. M.; McConnell, M. D.; Smith, D. A.; Watkins, J. L.; Allen, J. P.; Chaput, J. C. A synthetic protein selected for ligand binding affinity mediates ATP hydrolysis. *ACS Chem. Biol.* **2009**, *4*, 649–658.
- (25) Monasterio, O.; Nova, E.; Díaz-Espinoza, R. Development of a novel catalytic amyloid displaying a metal-dependent ATPase-like activity. *Biochem. Biophys. Res. Commun.* **2017**, *482*, 1194–1200.
- (26) Hoegler, K. J.; Hecht, M. H. A de novo protein confers copper resistance in *Escherichia coli*. *Protein Sci.* **2016**, *25*, 1249–1259.
- (27) Digianantonio, K. M.; Korolev, M.; Hecht, M. H. A non-natural protein rescues cells deleted for a key enzyme in central metabolism. *ACS Synth. Biol.* **2017**, *6*, 694–700.
- (28) Digianantonio, K. M.; Hecht, M. H. A Protein Constructed De Novo Enables Cell Growth by Altering Gene Regulation. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2400–2405.
- (29) Fisher, M. A.; McKinley, K. L.; Bradley, L. H.; Viola, S. R.; Hecht, M. H. De novo designed proteins from a library of artificial sequences function in *Escherichia coli* and enable cell growth. *PLoS One* **2011**, *6*, No. e15364.