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. 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. Moreover, grafting a P-loop cassette onto an artificial protein scaffold is sufficient to facilitate ATP binding and hydrolysis. 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. And though other ATP-hydrolyzing motifs occur occasionally in nature (e.g., among tyrosine kinases), all natural ATPases contain β-sheet structure. 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. 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. 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 $10^6$ sequences.

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 (pNP) esters (Figure 1). Second, promising candidates were purified and assayed for ATPase activity.

We demonstrated previously that proteins capable of hydrolyzing pNP esters are not uncommon among binary patterned libraries. Here we built on this finding by screening lysates expressing >1100 novel sequences for the ability to hydrolyze either pNP-palmitate or pNP-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 pNP-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
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 MgII, 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 $k_{cat}/K_M$ of 1.06 M$^{-1}$ s$^{-1}$, nearly 100-fold above background ($k_{cat}/K_M$).

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.
$k_{\text{cat}} = 98$. 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-Tasser17,18 produced a 4-helix bundle similar to that expected from our designed binary pattern (Figures 3b and S7).

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 ROSIE19−22 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 $\alpha$-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/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 surrounding solution conditions. Natural enzymes that work on nucleic acids are often more active when MgII is present in solution because nucleic acids and nucleotides are typically complexed with MgII in vivo. For example, MgII is present in E. coli at concentrations of 20−100 mM.23 However, to our surprise, AltTPase was inhibited by MgII, 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

### Table 1. Kinetic Parameters of AltTPase Mutants with Standard Errors

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$k_{\text{cat}}$ (h$^{-1}$)</th>
<th>$K_{\text{M}}$ (µM)</th>
<th>$k_{\text{cat}}/k_{\text{uncat}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AltTPase</td>
<td>0.035 ± 0.003</td>
<td>10.2 ± 1.3</td>
<td>98</td>
</tr>
<tr>
<td>F47L</td>
<td>0.0075 ± 0.0006</td>
<td>6.8 ± 0.9</td>
<td>21</td>
</tr>
<tr>
<td>N51S</td>
<td>0.005 ± 0.001</td>
<td>4.3 ± 1.2</td>
<td>14</td>
</tr>
<tr>
<td>D46S</td>
<td>0.0036 ± 0.0004</td>
<td>5 ± 1</td>
<td>10</td>
</tr>
<tr>
<td>K55E</td>
<td>undetectable</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

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.

![Figure 3](https://dx.doi.org/10.1021/jacs.0c02954)

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 MgCl$_2$ (gray), 50 µM MgCl$_2$ (dark blue), and without added MgCl$_2$ (blue) show that magnesium is a competitive inhibitor.

![Figure 4](https://dx.doi.org/10.1021/jacs.0c02954)
was observed for Ca$^{2+}$ and Zn$^{2+}$, and to a lesser extent for Na$^{+}$ (Figure S12). It is possible that these metals compete with K$^+$ for ionic interactions with the ATP phosphate tail. Yet inhibition by Mg$^{2+}$ 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, AltTAPse could not have evolved as a functional enzyme in nature.

Because AltTAPse 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. 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. Finally, Monasterio et al. reported a novel sequence that aggregates into an amyloid-like structure and has a very low level of Zn$^{2+}$-dependent ATPase activity. We compare AltTAPse 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, AltTAPse 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 arise 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 AltTAPse, 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$^{2+}$. 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, 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:0000-0002-5538-9813; Email: hecht@princeton.edu

Author
Michael S. Wang – Department of Chemistry, Princeton University, New Jersey 08540, United States; orcid: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

(4) Milner-White, E. J. Protein three-dimensional structures at the origin of life. Interface Focus 2019, 9, 20190057.
(7) Hanson, P. I.; Whiteheart, S. W. AAA proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. 2005, 6, 519.


