

# Enhancing Protein Stability with Genetically Encoded Noncanonical Amino Acids

Jack C. Li, Tao Liu,<sup>Ⓛ</sup> Yan Wang, Angad P. Mehta, and Peter G. Schultz\*<sup>Ⓛ</sup>

Department of Chemistry and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, United States

## Supporting Information

**ABSTRACT:** The ability to add noncanonical amino acids to the genetic code may allow one to evolve proteins with new or enhanced properties using a larger set of building blocks. To this end, we have been able to select mutant proteins with enhanced thermal properties from a library of *E. coli* homoserine O-succinyltransferase (*metA*) mutants containing randomly incorporated noncanonical amino acids. Here, we show that substitution of Phe 21 with (*p*-benzoylphenyl)alanine (pBzF), increases the melting temperature of *E. coli metA* by 21 °C. This dramatic increase in thermal stability, arising from a single mutation, likely results from a covalent adduct between Cys 90 and the keto group of pBzF that stabilizes the dimeric form of the enzyme. These experiments show that an expanded genetic code can provide unique solutions to the evolution of proteins with enhanced properties.

All known organisms use the same 20 canonical amino acids with the only known exceptions being selenocysteine and pyrrolysine, which are found in a small number of proteins. This fact raises the question as to whether 20 is the optimal number, or whether additional amino acids might provide an evolutionary advantage to an organism. We and others have begun to explore this notion by randomly incorporating noncanonical amino acids (ncAAs) into proteins and carrying out screens or selections to identify proteins with enhanced properties. To this end, cyclic peptides, antibodies, enzymes, and phage have been identified in which a noncanonical amino acid provides improved affinity, catalytic activity, or fitness.<sup>1–6</sup> We now ask whether ncAAs can enhance the physical properties of proteins, specifically, their thermal stability.

Protein stability generally depends on a large number of van der Waals, hydrogen bonding, and electrostatic interactions and a limited number of disulfide bonds. With the exception of the latter, these interactions are relatively weak in energy but collectively can lead to a significant favorable free energy of folding. As a result, efforts to increase the thermal stability of proteins often require the additive effects of multiple mutations.<sup>7,8</sup> Due to their covalent nature, engineered cysteine cross-links have been extensively explored as a means of increasing thermal stability, although they are subject to stringent distance and conformational constraints that limit their application.<sup>9–11</sup> With the introduction of additional functional groups in the side chains of ncAAs that can form

new noncovalent or covalent interactions not accessible to the common 20 amino acids, we asked whether additional energetically stabilizing interactions can be realized.<sup>12–15</sup> Here we show that introduction of a single benzophenone containing ncAA into the *Escherichia coli* protein homoserine O-succinyltransferase (*metA*) leads to a 21 °C increase in thermal stability.

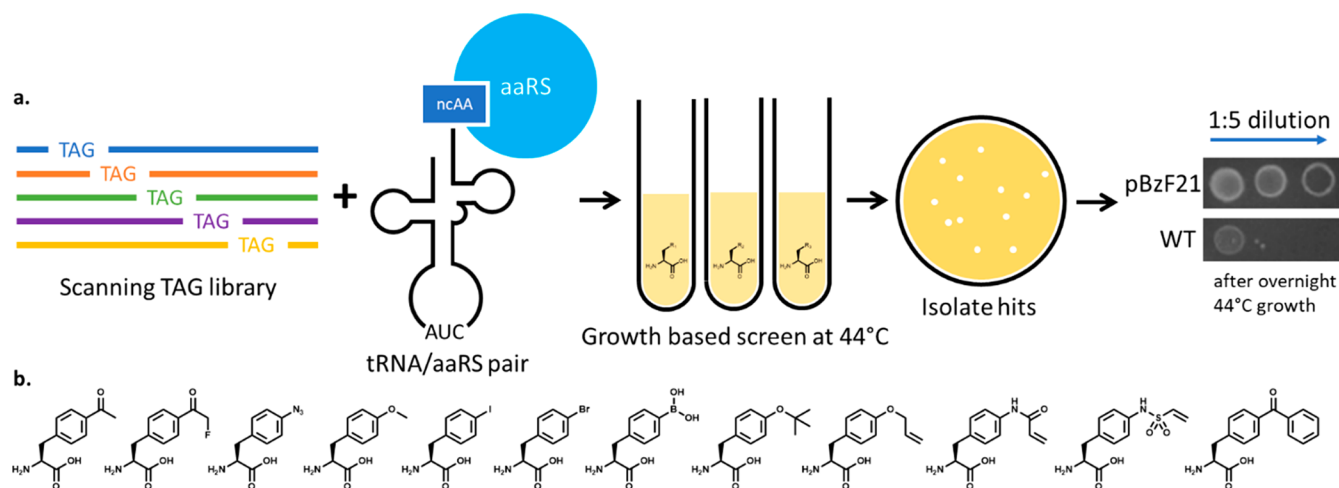
Previously, it has been shown that the growth of *E. coli* is greatly reduced at temperatures of 44 °C and above in minimal medium, and that the addition of methionine to the media is able to partially rescue growth.<sup>16,17</sup> This effect is largely attributed to the thermal instability of a single enzyme, homoserine O-succinyltransferase, an essential enzyme in methionine biosynthesis that begins to unfold and aggregate above 40 °C.<sup>18</sup> Mutations that moderately increase *metA* stability and reduce aggregation have been found to confer enhanced growth rates at 44 °C.<sup>19,20</sup> This property of *metA* provides a rational basis for selection, as *metA* mutants containing a noncanonical amino acid which increases the enzyme's thermal stability should have a significant growth advantage.

To test this notion, a comprehensive amber nonsense TAG scanning library (containing 261 mutants out of 308 possible positions) was generated in which almost every position in *metA* not directly adjacent to the active site was mutated, as determined using a homology model of *E. coli metA* (generated on the basis of the previously solved crystal structure of *metA* in *B. cereus*).<sup>21</sup> The resulting library was then transformed into *E. coli* strain JW3973-1 ( $\Delta metA$ ) along with a plasmid bearing a polyspecific amber nonsense suppressor tRNA/aminoacyl-tRNA synthetase (aaRS) pair that is orthogonal to the endogenous *E. coli* tRNA/aaRS pairs (i.e., does not cross-react with host tRNA/aaRS pairs) and is able to selectively and efficiently incorporate a variety of ncAAs, including (*p*-acetylphenyl)alanine, [*p*-(fluoroacetyl)phenyl]alanine, (*p*-azidophenyl)alanine, O-methyltyrosine, (*p*-iodophenyl)alanine, (*p*-bromophenyl)alanine, O-*tert*-butyltyrosine, O-allyltyrosine, *p*-biphenylalanine, and [*p*-(acrylamido)phenyl]alanine. Three additional tRNA/synthetase pairs that specifically incorporate [*p*-(vinylsulfonamido)phenyl]alanine, (*p*-boronophenyl)alanine, and (*p*-benzoylphenyl)alanine were also cotransformed separately. Transformants were grown at 44 °C in liquid minimal media containing each of the ncAAs for 20–40 h. We expected that *E. coli* containing *metA* mutants

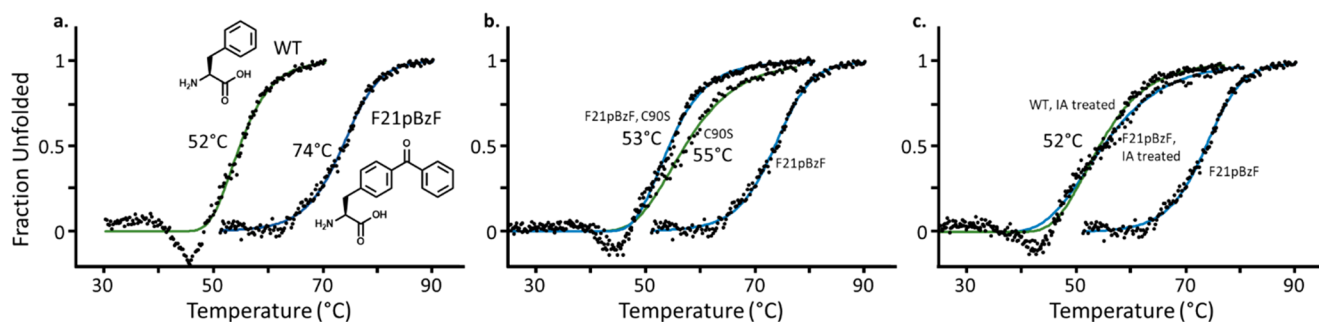
Received: July 11, 2018

Published: November 15, 2018





**Figure 1.** Screening for high thermal stability mutants of *metA*. (a) An amber stop codon scanning library of *metA* was cotransformed with an orthogonal amber suppressor tRNA/aaRS pair into JW3973-1 *E. coli*. Transformants were incubated in minimal media liquid culture supplemented with 1 mM of the cognate ncAA at 44 °C for 20–40 h. Cells were then plated and colonies were sequenced. The consensus clone F21pBzF *metA* was retransformed into JW3973-1 and shows an increased growth rate compared to that of JW3973-1 transformed with WT *metA*. (b) ncAAs used in this screen.



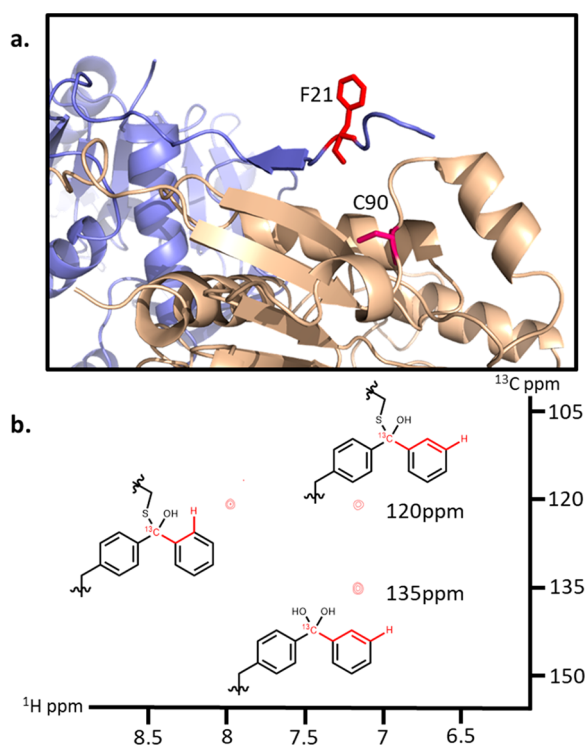
**Figure 2.** Thermal denaturation of *metA*. (a) Purified recombinant WT *metA* (green) and F21pBzF *metA* (blue) were analyzed by CD spectroscopy at 223 nm in 200 mM sodium phosphate, pH 7.4. The fraction unfolded is represented as normalized ellipticity. F21pBzF *metA* shows a 21 °C increase in melting temperature compared to that of its WT counterpart. (b, c) Mutation of Cys 90 to Ser (b) or Cys 90 alkylation by IA (c) reverts the  $T_m$  increase of the F21pBzF mutant to WT *metA* or near WT *metA* temperatures.

that confer increased thermal stability to outgrow other bacteria. Cultures were subsequently plated and colonies were picked and sequenced. Two consensus mutations, F21TAG and N86TAG, were observed when cells were grown in the presence of 1 mM (*p*-benzoylphenyl)alanine (pBzF) and 1 mM *O*-*tert*-butyltyrosine (OtBuY), respectively. The F21TAG *metA* mutant showed markedly higher end point growth than WT in the presence of pBzF when retransformed into  $\Delta metA$  cells (Figure 1).

We next expressed and purified the N-terminal 6xHis tagged WT, N86OtBuY, and F21pBzF variants of *metA*, referred to hereafter as WT *metA*, N86OtBuY *metA*, and F21pBzF *metA*. The yields of the WT protein were typically 8–9 mg/L and those of the mutants were 5–6 mg/L. Temperature dependent protein unfolding was then measured using circular dichroism (CD) spectroscopy. The F21pBzF mutant showed a remarkable 21 °C increase in melting temperature ( $T_m$ ) compared to that of the WT protein (53 °C for WT vs 74 °C for the F21pBzF mutant) (Figure 2), while N86OtBuY *metA* showed a modest 6 °C increase (Figure S1). Given the substantially enhanced stability of F21pBzF *metA*, we focused on this mutant for further study. The WT *metA* and F21pBzF *metA* were then coexpressed bicistronically to generate a mixture of homo and heterodimers. CD melting curves for this

mixture had an expected bimodality, with a lower  $T_m$  curve matching the WT protein at ~51 °C and a higher  $T_m$  curve at ~65 °C, likely corresponding to a WT/mutant heterodimer (and mutant homodimer in lower abundance, consistent with the typically lower expression yields of proteins containing a ncAA) (Figure S2).

On the basis of our homology model, *metA* is predicted to be a homodimer, which we verified by size exclusion elution chromatography of the purified recombinant protein (Figure S3) and dynamic light scattering. Phe 21 is predicted to lie in a highly flexible N-terminal domain region that interacts with the other monomer of the homodimer in a domain swapping mode (Figure 3). This observation suggests that pBzF 21 may stabilize the dimeric state of *metA*. Interestingly, many of the other colonies that grew at elevated temperatures but did not converge on a consensus sequence had mutations in the same N-terminal domain, indicating it may be a generally important area for *metA* stability. When  $\Delta metA$  *E. coli* was retransformed with the F21TAG *metA* mutant and the cognate tRNA/aaRS pair (as pEVOL-pBzF) in media containing 1 mM pBzF, bacteria containing the mutant *metA* appear to be able to grow faster than those containing the wildtype (WT) *metA* (Figure S4).



**Figure 3.** Expected ncAA interaction. (a) A homology model of *E. coli metA* was generated on the basis of *metA* from *B. cereus* (pdb 2VDJ). The protein is predicted to form a homodimer. The mutant residue, Phe 21 (red), is hypothesized to cross-link with Cys 90 (magenta) of the other monomer. (b) HMBC spectra of  $^{13}\text{C}$  labeled F21pBzF *metA* shows noncarbonyl cross peaks consistent with hemithioketal formation. The 195 ppm carbonyl signal of pBzF was not observed.

The benzophenone side chain of pBzF can form van der Waals interactions, hydrogen bonds, or covalent interactions with neighboring residues. The latter can be formed through electrophilic addition of nucleophilic amino acid side chains to the ketone. Given the large increase in  $T_m$  associated with this mutation, we explored the possibility that the aryl keto group is forming a covalent bond with a Lys, Cys, or Ser residue. As denaturing SDS-PAGE gels of our protein and mass spectrometry consistently show only the monomeric mass (Figure S8), this bond is likely reversible. On the basis of the predicted location of Phe 21 in *E. coli metA*, a small proximity-based list of candidate residues was generated that could form a covalent adduct with the benzophenone side chain. Among these, mutagenesis of Cys 90 to Ser showed a complete reversion of the melting temperature from 74 to 53 °C (while the same mutation is actually slightly stabilizing in the case of the WT protein), consistent with the reduced nucleophilicity of serine compared to that of cysteine. Covalent modification of Cys 90 with iodoacetamide (IA) similarly reduced the  $T_m$  for F21pBzF *metA* to 53 °C while not significantly affecting the  $T_m$  of the WT protein. These results suggest that the increased thermostability afforded by the benzophenone side chain is due to its interaction with Cys 90, likely in the form of a hemithioketal (Figure 2).

Hemithioketals are usually intermediates in thioketal formation and the formation of hemithioketals is readily reversible, with previous studies having shown that ketone to hemithioketal equilibria can vary by 4 orders of magnitude or more, depending on conditions.<sup>22,23</sup> Here, it is possible that a hemithioketal is stabilized by the surrounding protein

architecture. Mutation of Phe 21 to [*p*-(fluoroacetyl)phenyl]-alanine or (*p*-acetylphenyl)alanine afforded  $T_m$ 's of 60 and 53 °C, respectively, which may reflect the reduced stability of the thiol adducts or additional stabilizing interactions with the additional aryl ring of pBzF (Figure S5). Because hemithioketals may be further reacted with an additional thiol to form a stable thioketal, it may be possible to trap the F21pBzF *metA* covalent dimer irreversibly by addition of  $\beta$ -mercaptoethanol (BME). Indeed, addition of BME trapped a fraction of the mutant in its homodimeric form, as evidenced by SDS-PAGE, while mass spectrometry of the trapped complex confirmed the covalent addition of a single BME (Figure S6).

The putative cross-link between pBzF 21 and Cys 90 was further confirmed using  $^{13}\text{C}$  NMR, in which a  $^{13}\text{C}$  labeled version of pBzF was synthesized and incorporated into *metA*. Proton-carbon heteronuclear multiple bond correlation (HMBC) NMR on this sample shows three distinct non-carbonyl aromatic cross peaks that likely correspond to the hemithioketal and hydrated forms of pBzF (Figures 3 and S7) in the region 120–135 ppm on the  $^{13}\text{C}$  axis while the 195 ppm signal due to the  $\text{sp}^2$  carbonyl group was not observed.<sup>24</sup> We further attempted to obtain a crystal structure of *metA* but thus far have been unsuccessful in obtaining crystals for the thermostable form of this protein.

While there have been extensive protein engineering efforts to increase melting temperatures, most result in small improvements on the order of 5–10 °C.<sup>25–28</sup> Two notable previous examples include a 25 °C melting temperature increase in GH11 xylanase, which involved seven mutations that resulted in subtle perturbations to the protein's overall structure, and a 26 °C melting temperature increase in *Leishmania* triosephosphate isomerase, in which a single glutamate was mutated to glutamine, which restored a hydrogen bonding network within the protein that had been naturally lost.<sup>29,30</sup> In our case, a single novel point mutation was selected that was able to increase the melting temperature of *metA* by over 20 °C by stabilizing the protein's native dimer configuration. In doing so, this work underscores how an expanded chemical space gives natural systems unique avenues for evolving enhanced properties.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b07157.

Thermal denaturation graphs, SEC and DLS, growth rates, dimer trapping gels and MS, NMR spectra, SDS-PAGE and MS (Figures S1–S7), gels and mass spectrometry (Figure S8), strains, plasmids, and amino acids used (Tables S1 and S2), experimental methods, and supplementary references (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*P. G. Schultz. Email: schultz@scripps.edu

### ORCID

Tao Liu: 0000-0001-5347-5892

Peter G. Schultz: 0000-0003-3188-1202

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank the Scripps Research Institute Mass Spectrometry Core and Linh T. Hoang; David E. Wemmer, Alex Siegel, Laura Pasternack, and Dee-Hua Huang for their valuable assistance with <sup>13</sup>C NMR; and Raymond C. Stevens and Fariborz Nasertorabi for assistance with crystallography. This work is supported by NIH grant R01 GM062159.

## ■ REFERENCES

- (1) Liu, C. C.; Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **2010**, *79*, 413–44.
- (2) Wang, L.; Schultz, P. G. Expanding the genetic code. *Angew. Chem., Int. Ed.* **2004**, *44* (1), 34–66.
- (3) Wang, L.; Xie, J.; Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–49.
- (4) Xie, J.; Schultz, P. G. A chemical toolkit for proteins — an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775–782.
- (5) Liu, T.; Wang, Y.; Luo, X.; Li, J.; Reed, S. A.; Xiao, H.; Young, T. S.; Schultz, P. G. Enhancing protein stability with extended disulfide bonds. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (21), 5910–5.
- (6) Xiao, H.; Nasertorabi, F.; Choi, S. H.; Han, G. W.; Reed, S. A.; Stevens, R. C.; Schultz, P. G. Exploring the potential impact of an expanded genetic code on protein function. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (22), 6961–6.
- (7) Rha, E.; Kim, S.; Choi, S. L.; Hong, S. P.; Sung, M. H.; Song, J. J.; Lee, S. G. Simultaneous improvement of catalytic activity and thermal stability of tyrosine phenol-lyase by directed evolution. *FEBS J.* **2009**, *276* (21), 6187–94.
- (8) Singh, R. K.; Tiwari, M. K.; Singh, R.; Lee, J. K. From protein engineering to immobilization: promising strategies for the upgrade of industrial enzymes. *Int. J. Mol. Sci.* **2013**, *14* (1), 1232–77.
- (9) Matsumura, M.; Becktel, W. J.; Levitt, M.; Matthews, B. W. Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 6562–6566.
- (10) Wetzel, R.; Perry, L. J.; Baase, W. A.; Becktel, W. J. Disulfide bonds and thermal stability in T4 lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 401–405.
- (11) Dombkowski, A. A.; Sultana, K. Z.; Craig, D. B. Protein disulfide engineering. *FEBS Lett.* **2014**, *588* (2), 206–12.
- (12) Furman, J. L.; Kang, M.; Choi, S.; Cao, Y.; Wold, E. D.; Sun, S. B.; Smider, V. V.; Schultz, P. G.; Kim, C. H. A genetically encoded aza-Michael acceptor for covalent linking of protein-receptor complexes. *J. Am. Chem. Soc.* **2014**, *136* (23), 8411–7.
- (13) Xiang, Z.; Ren, H.; Hu, Y. S.; Coin, I.; Wei, J.; Cang, H.; Wang, L. Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. *Nat. Methods* **2013**, *10* (9), 885–8.
- (14) Xiang, Z.; Lacey, V. K.; Ren, H.; Xu, J.; Burban, D. J.; Jennings, P. A.; Wang, L. Proximity-enabled protein linking through genetically encoding haloalkane unnatural amino acids. *Angew. Chem., Int. Ed.* **2014**, *53* (8), 2190–3.
- (15) Lee, Y. J.; Wu, B.; Raymond, J. E.; Zeng, Y.; Fang, X.; Wooley, K. L.; Liu, W. R. A genetically encoded acrylamide functionality. *ACS Chem. Biol.* **2013**, *8* (8), 1664–70.
- (16) Ron, E. Z.; Davis, B. D. Growth Rate of *Escherichia coli* at Elevated Temperatures: Limitation by Methionine. *J. Bacteriol.* **1971**, *107* (2), 391–396.
- (17) Ron, E. Z.; Shani, M. Growth Rate of *Escherichia coli* at Elevated Temperatures: Reversible Inhibition of Homoserine Trans-Succinylase. *J. Bacteriol.* **1971**, *107* (2), 397–400.
- (18) Gur, E.; Biran, D.; Gazit, E.; Ron, E. Z. In vivo aggregation of a single enzyme limits growth of *Escherichia coli* at elevated temperatures. *Mol. Microbiol.* **2002**, *46* (5), 1391–1397.
- (19) Mordukhova, E. A.; Kim, D.; Pan, J. G. Stabilized homoserine o-succinyltransferases (MetA) or L-methionine partially recovers the growth defect in *Escherichia coli* lacking ATP-dependent proteases or the DnaK chaperone. *BMC Microbiol.* **2013**, *13*, 179.
- (20) Mordukhova, E. A.; Lee, H. S.; Pan, J. G. Improved thermostability and acetic acid tolerance of *Escherichia coli* via directed evolution of homoserine o-succinyltransferase. *Appl. Environ. Microbiol.* **2008**, *74* (24), 7660–8.
- (21) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296–W303.
- (22) Lienhard, G. E.; Jencks, W. P. Thiol addition to the carbonyl group. Equilibria and Kinetics. *J. Am. Chem. Soc.* **1966**, *88* (17), 3982–3995.
- (23) Burkey, T. J.; Fahey, R. C. Equilibrium Studies of Water and Thiol Addition to Ketones: Substituent and Solvent Effects for Methyl Ketones. *J. Am. Chem. Soc.* **1983**, *105*, 868–871.
- (24) Karthikeyan, J.; Parthasarathy, K.; Cheng, C. H. Synthesis of biarylketones and phthalides from organoboronic acids and aldehydes catalyzed by cobalt complexes. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47* (37), 10461–3.
- (25) Hecht, M.; Sturtevant, J. M.; Sauer, R. T. Effect of single amino acid replacements on the thermal stability of the NH2-terminal domain of phage lambda repressor. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 5685–5689.
- (26) Gray, K. A.; Richardson, T. H.; Kretz, K.; Short, J. M.; Bartneck, F.; Knowles, R.; Kan, L.; Swanson, P. E.; Robertson, D. E. Rapid Evolution of Reversible Denaturation and Elevated Melting Temperature in a Microbial Haloalkane Dehalogenase. *Adv. Synth. Catal.* **2001**, *343*, 607–617.
- (27) Matthews, B. W.; Nicholson, H.; Becktel, W. J. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 6663–6667.
- (28) Das, G.; Hickey, D. R.; McLendon, D.; McLendon, G.; Sherman, F. Dramatic thermostabilization of yeast iso-1-cytochrome c by an asparagine to isoleucine replacement at position 57. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 496–499.
- (29) Williams, J. C.; Zeelen, J. P.; Neubauer, G.; Vriend, G.; Backmann, J.; Michels, P. A. M.; Lambier, A.; Wierenga, R. K. Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power. *Protein Eng., Des. Sel.* **1999**, *12* (3), 243–250.
- (30) Dumon, C.; Varvak, A.; Wall, M. A.; Flint, J. E.; Lewis, R. J.; Lakey, J. H.; Morland, C.; Luginbuhl, P.; Healey, S.; Todaro, T.; DeSantis, G.; Sun, M.; Parra-Gessert, L.; Tan, X.; Weiner, D. P.; Gilbert, H. J. Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. *J. Biol. Chem.* **2008**, *283* (33), 22557–64.