JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Anaerobic 5-Hydroxybenzimidazole Formation from Aminoimidazole Ribotide: An Unanticipated Intersection of Thiamin and Vitamin B_{12} Biosynthesis

Angad P. Mehta,[†] Sameh H. Abdelwahed,[†,∥] Michael K. Fenwick,[‡] Amrita B. Hazra,[§] Michiko E. Taga,[§] Yang Zhang,[‡] Steven E. Ealick,[‡] and Tadhg P. Begley*[,†]

[†]Department of Chemistry, Texas A&M University, College Station, Texas 77843, United States
[‡]Department of Chemistry and Chemical Biology, Cornell University, 120 Baker Lab, Ithaca, New York 14853, United States
[§]Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, United States
[∥]Therapeutic Chemistry Department, National Research Centre, Dokki, Cairo, Egypt

Supporting Information

ABSTRACT: Comparative genomics of the bacterial thiamin pyrimidine synthase (*thiC*) revealed a paralogue of *thiC* (*bzaF*) clustered with anaerobic vitamin B_{12} biosynthetic genes. Here we demonstrate that BzaF is a radical *S*-adenosylmethionine enzyme that catalyzes the remarkable conversion of aminoimidazole ribotide (AIR) to 5-hydroxybenzimidazole (5-HBI). We identify the origin of key product atoms and propose a reaction mechanism. These studies represent the first step in solving a long-standing problem in anaerobic vitamin B_{12} assembly and reveal an unanticipated intersection of thiamin and vitamin B_{12} biosynthesis.

The formation of dimethylbenzimidazole (DMB, **3**), the lower axial ligand of vitamin B_{12} (cobalamin, **1**) (Figure 1A) is the last remaining major unsolved puzzle in B_{12} biosynthesis. In aerobic organisms, DMB is biosynthesized from reduced flavin mononucleotide ($FMNH_2$, **2**) in a remarkable reaction catalyzed by the BluB gene product (Figure 1B). This oxygen-requiring DMB synthase has been biochemically and structurally characterized, but its mechanism is not yet known.[1,2] The anaerobic pathway to DMB involves different chemistry, and labeling studies have demonstrated that FMN is not the precursor (Figure 1C).[3−6] The anaerobic DMB synthase has not been yet identified.

Our approach to elucidating the anaerobic biosynthesis of DMB was indirect and derived from our studies of the bacterial thiamin pyrimidine synthase (ThiC). This radical *S*-adenosylmethionine (SAM) enzyme catalyzes a remarkable rearrangement of aminoimidazole ribotide (AIR, **10**) to the thiamin pyrimidine (**12**) (Figure 2).[7−10]

While looking for a ThiC paralogue to facilitate our mechanistic studies of this complex reaction, we noticed that some anaerobes in the SEED database (http://theseed.uchicago.edu/FIG/) have a *thiC* paralogue (*bzaF*) clustered with vitamin B_{12} biosynthetic genes (Figure 3). This suggested the possibility that BzaF might catalyze the conversion of AIR to a DMB precursor. Previous in vivo labeling studies in *Eubacterium limosum* using 4-aminoimidazole eliminated this heterocycle as a DMB precursor.[12] In this Communication, we report the successful reconstitution of the BzaF-catalyzed conversion of AIR to 5-hydroxybenzimidazole (HBI) and

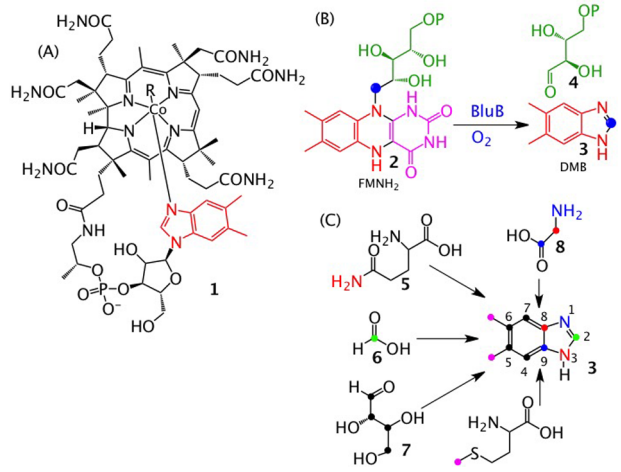

Figure 1. Oxygen-dependent and oxygen-independent biosyntheses of the DMB ligand of vitamin B_{12}. (A) The structure of vitamin B_{12} (**1**) with the DMB ligand shown in red. (B) Oxygen-dependent DMB formation from $FMNH_2$ (**2**). (C) Labeling studies showing the oxygen-independent biosynthetic origin of the atoms of DMB in *Eubacterium limosum*. 5-Hydroxybenzimidazole (HBI) and 5-hydroxy-6-methylbenzimidazole have been identified as precursors.[11]

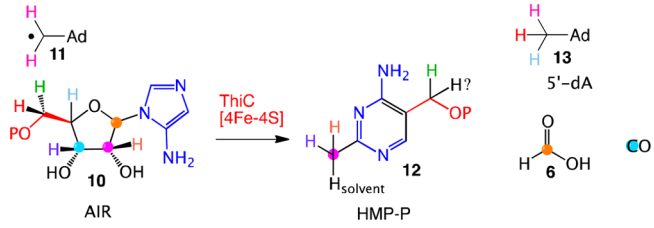

Figure 2. The remarkable ThiC-catalyzed conversion of AIR (**10**) to thiamin pyrimidine (**12**).

Received: April 10, 2015
Published: August 3, 2015

DOI: 10.1021/jacs.5b03576
J. Am. Chem. Soc. 2015, 137, 10444−10447

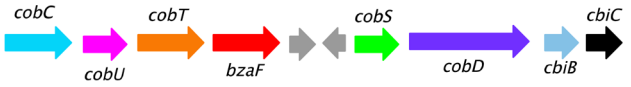

Figure 3. The *bzaF* gene cluster in *Desulfuromonas acetoxidans*. *cobCUT*, *cobSD*, and *cbiBC* have established functions in vitamin B_{12} biosynthesis. BzaF has 38% sequence identity and 53.6% sequence similarity to *D. acetoxidans* ThiC. (Bza is the IUPAC–IUB abbreviation for "benzimidazolyl".)

identify the hydrogen atom initially abstracted by the 5′-deoxyadenosyl radical and the origin, in AIR, of key HBI atoms. These results are integrated into a mechanistic hypothesis for the BzaF (HBI synthase)-catalyzed reaction.

The *bzaF* gene from *Desulfuromonas acetoxidans* was synthesized with codon optimization (for gene details, see Figure S33), cloned in the THT vector, and coexpressed in the presence of a plasmid encoding the *suf* operon for in vivo [4Fe-4S] reconstitution in *Escherichia coli* BL21-(DE3) (see p 14 in the Supporting Information (SI)). The enzyme was then purified under anaerobic conditions using Ni-NTA chromatography (see the SI for experimental details). The enzyme yield was 3 mg/L of culture. BzaF has a molecular mass of 48 kDa and contains 3 irons and 2.6 sulfides per monomer. The UV–vis spectrum of the purified protein showed the long-wavelength absorption characteristic of the $[4Fe\text{-}4S]^{2+}$ cluster (Figure 4B).

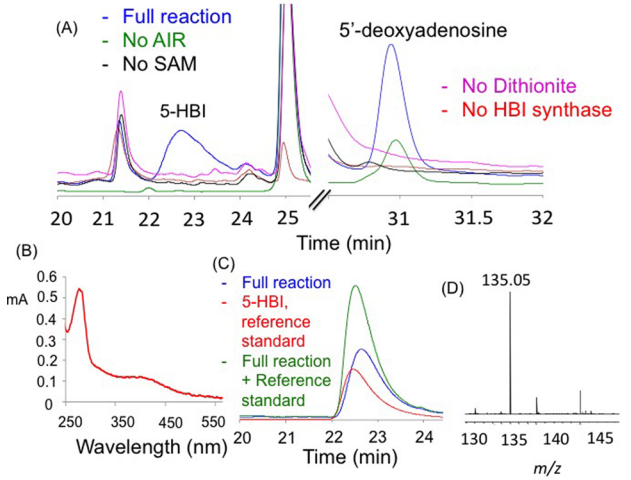

Figure 4. Characterization of the product of the BzaF-catalyzed reaction. (A) HPLC (254 nm) analysis of the HBI synthase reaction mixture showing a new product eluting at 22.7 min and 5′-dA eluting at 31 min (blue trace). This product was not formed in reaction mixtures lacking dithionite (red trace). (See Figure S31 for all controls). (B) UV–vis spectrum of as-isolated HBI synthase. (C) Product identification by HPLC analysis (red trace, HBI; blue trace, full enzymatic reaction; green trace, coinjection of HBI and the reaction product). (D) MS analysis of the 22.7 min enzymatic product showing $[M + H]^+$ = 135.05 Da, consistent with HBI formation.

Anaerobic incubation of AIR and SAM with dithionite-reduced BzaF (room temperature, 60 min) resulted in the formation of 5′-deoxyadenosine (5′-dA) and a new product as determined by HPLC analysis (Figure 4A). This product had an $[M + H]^+$ of 135.05 Da and comigrated with an authentic standard of HBI (Figure 4C,D). This identification was confirmed by scale-up of the reaction followed by NMR analysis of the purified product. The HBI:BzaF and 5′-dA:BzaF ratios were 0.7 and 1.1, respectively. The corresponding ratios for ThiC are 0.4 (HMP-P:*Caulobacter crescentus* ThiC) and 0.6 (5′-dA:*C. crescentus* ThiC).

As a first step in elucidating the mechanism of HBI synthase, we have determined the origin of key atoms of HBI. Our strategy involved the conversion of a set of ^2H- and ^{13}C-AIR isotopologues to HBI followed by the identification of the site of labeling by NMR and MS analysis. The results of these experiments are shown in Figure 5 and Table S1 (also see Figures S20–S29). The key findings, summarized in Figure 5A, are as follows: (1) The pro-*S* hydrogen at C5′ of AIR is abstracted by the adenosyl radical (Table S1, lines 2–4). (2) The C7 hydrogen of HBI is derived from the C5′ pro-*R* hydrogen of AIR (Table S1, line 3, and Figure 5F). (3) The C6 hydrogen of HBI is derived from the C4′ hydrogen of AIR

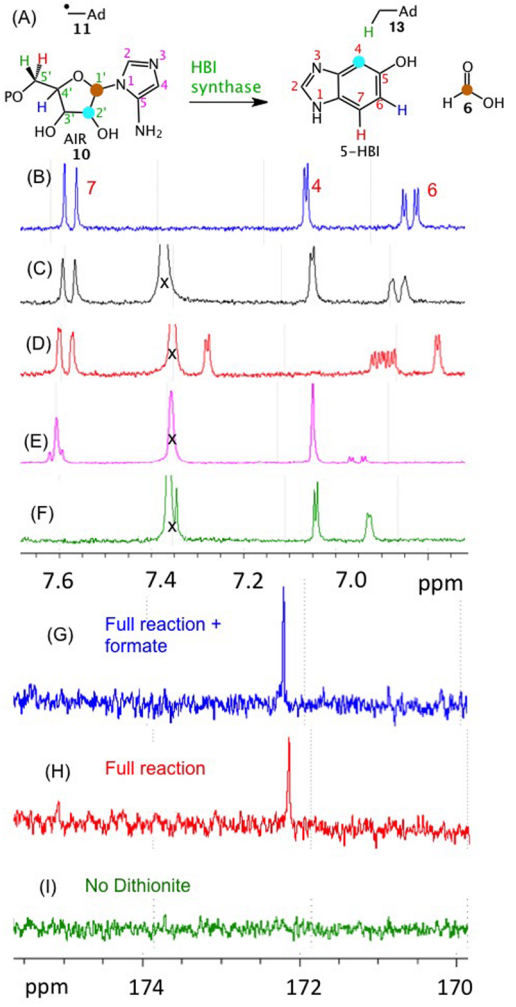

Figure 5. Labeling studies to determine the origin of key HBI atoms. (A) Summary of the labeling studies. (B) Partial ^1H NMR spectrum for authentic HBI. (C–F): ^1H NMR spectra of HBI purified from enzymatic reactions using (C) unlabeled AIR, (D) 2′-^{13}C-AIR, (E) 4-^2H-AIR, and (F) 5′,5′-$^2H_2$-AIR. (G–I): ^{13}C DEPT 90 analyses of the reaction mixtures for (G) the reaction of HBI synthase with 1′,2′,3′,4′,5′-$^{13}C_5$-AIR spiked with formate (10 mM) after 2 h, (H) the reaction of HBI synthase with 1′,2′,3′,4′,5′-$^{13}C_5$-AIR, and (I) the control without dithionite. Peaks labeled with X are due to an unknown impurity. The doublet at 7.6 ppm and the doublet of doublets at 7.05 ppm in (E) are due to a trace of unlabeled AIR in the sample (see Figure S30 for ^{13}C NMR spectra).

DOI: 10.1021/jacs.5b03576
J. Am. Chem. Soc. 2015, 137, 10444–10447

(Table S1, line 5, and Figure 5E). (4) The C1′ carbon of AIR is lost as formate (Table S1, lines 6–9, and Figure 5G–I). (5) The C5 amino group nitrogen of AIR is partially retained in the product (Table S1, line 10; 62% retention, see the legend of Table S1). (6) The C4 carbon of HBI is derived from the C2′ carbon of AIR (Figure 5D). From these experiments, we can deduce the origin of all of the atoms of HBI, as shown in Figure 6.

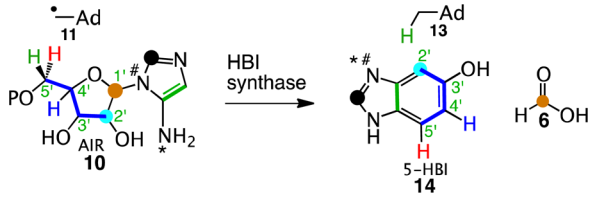

Figure 6. Origin of all of the atoms of HBI based on the results of labeling studies shown in Figure 5 and Table S1.

These labeling studies allow us to propose a mechanistic hypothesis for the HBI-synthase-catalyzed reaction (Figure 7A). As expected from the sequence similarity with ThiC, HBI formation is initiated by the abstraction of the C5′ pro-*S* hydrogen of AIR by the 5′-deoxyadenosyl radical to give **15**. C–O bond cleavage followed by *N*-glycosyl bond cleavage gives **17**. At this point, the chemistry diverges from the proposed ThiC pathway (see Figure S32). Electrophilic addition of the ribose C4′ to the aminoimidazole leads to the thiamin pyrimidine, while electrophilic addition of the ribose C5′ gives **19**, which leads to HBI. Loss of phosphate followed by deprotonation gives **21**, which is quenched by hydrogen atom donation from an unknown source. A ring-opening/closing sequence gives **24**. Loss of ammonia from **24** gives imine **25**, in which 38% of the imine nitrogen is derived from N1 of AIR and the remainder is derived from the C5 amino group. This is consistent with the partial scrambling of the two amidine nitrogens in **23** by C–C bond rotation. Loss of formate, ring closure, and a series of dehydrations and tautomerizations complete the formation of HBI (**14**).

A homology model for HBI synthase was constructed using the structure of the *Arabidopsis thaliana* ThiC/AIR/*S*-adenosylhomocysteine (SAH)/Fe complex as a template.[13] HBI synthase lacks the chloroplast-targeting sequence specific to *Arabidopsis* ThiC as well as the N-terminal ThiC domain.[7,13,14] This N-terminal domain is also missing from ThiCs from anaerobic organisms. HBI synthase is predicted to have the noncanonical active-site architecture recently reported for ThiC with an additional iron chelated by two conserved histidine residues and the amino and carboxylate groups of SAM.[13] The homology model is consistent with the observed abstraction of the 5′ pro-*S* hydrogen atom of the substrate and shows remarkable conservation of active-site residues, with only four differences out of 21 residues in the first shell compared with ThiC (Figure 7B,C). Three of the differences involve contacts with AIR. In ThiC, Asn228 hydrogen-bonds to AIR O3′ and is replaced by Ser66 in HBI synthase. Thr320 hydrogen-bonds to Glu422, which in turn hydrogen-bonds to AIR O2′ and is replaced by Ala161 in HBI synthase; Glu266 is equivalent to Glu422 and still hydrogen-bonds to AIR O2′. Asp383 (if protonated) hydrogen-bonds to AIR N3 (2.8 Å) and is replaced by Asn227 in HBI synthase. The fourth difference is near SAM, where Arg343 hydrogen-bonds to N7 of SAM and is replaced by Lys187, which is also predicted to hydrogen-bond to N7 in HBI synthase. BLAST searches[15] starting with either *D. acetoxidans* HBI synthase or *A. thaliana* ThiC suggest that these four differences alone distinguish the two paralogues.

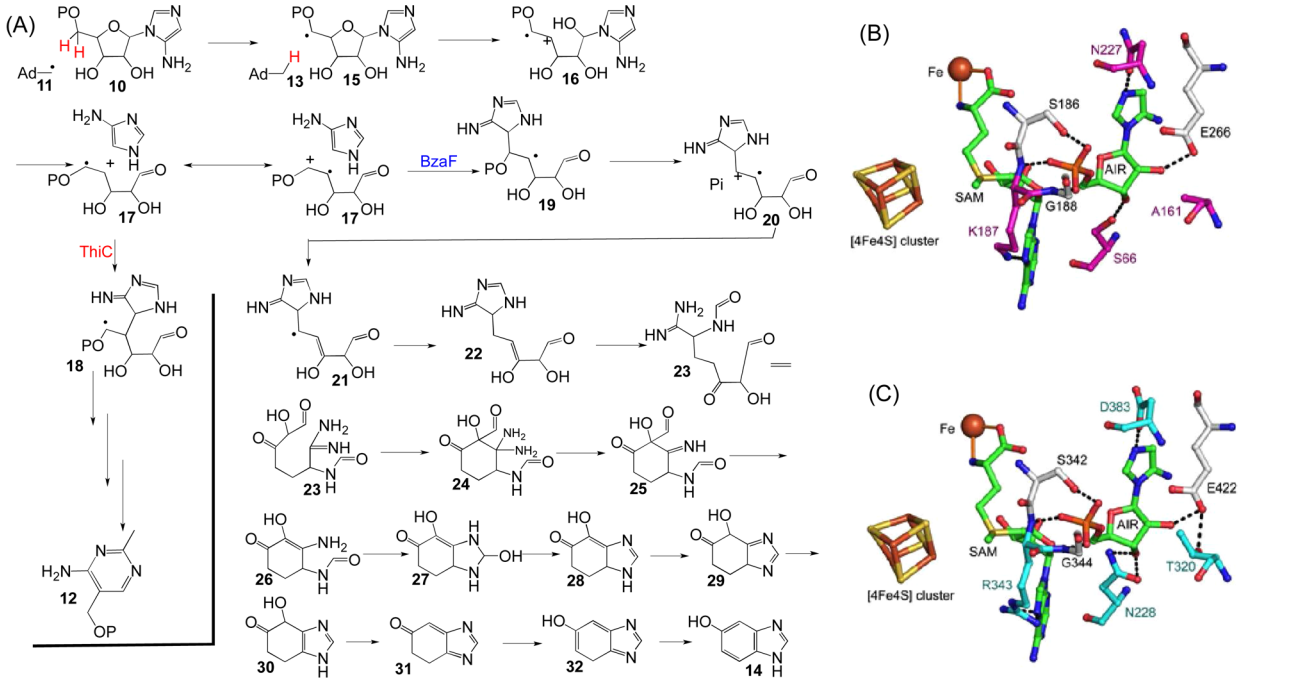

Figure 7. (A) Mechanistic proposal for the radical-triggered conversion of AIR to HBI catalyzed by HBI synthase. (B) Homology model for HBI synthase (BzaF). (C) X-ray structure of the ThiC active site. The [4Fe-4S] cluster, additional iron, SAM, AIR, and key active-site residues are shown. SAM was generated by adding a methyl group in the *S* configuration to SAH. The additional iron is also chelated by two conserved side chains (His270 and His334 in HBI synthase and His426 and His490 in ThiC).

DOI: 10.1021/jacs.5b03576
J. Am. Chem. Soc. 2015, 137, 10444–10447

Conserved active sites suggest similar catalytic mechanisms for most enzymatic reactions. This is clearly not the case for ThiC and BzaF. These proteins reveal unique features of enzymatic catalysis of reactions that proceed via high-energy intermediates. For such reactions, it is likely that the high-energy substrate radical **15** is converted to HMP-P (**12**) or to HBI (**14**) in a complex multistep sequence in which several steps along the reaction coordinate do not require catalytic assistance from the enzyme. This allows for similar active sites to produce different products. This catalytic motif has previously been observed in terpene cyclases, where similar active sites can catalyze the conversion of high-energy carbocations to different products. For example, the D100E mutant of trichodiene synthase generates six terpenes[16] and native baruol synthase produces a library of 23 terpenes.[17]

Our studies on the identification and reconstitution of the vitamin B_{12} HBI synthase represent the first steps in solving a long-standing problem in anaerobic vitamin B_{12} biosynthesis. While this work was in progress, the *bza* genes encoding the anaerobic DMB biosynthesis pathway in *Eubacterium limosum* were reported. The *bza* operon contains homologs of *bzaF*, thus confirming the role of HBI synthase as the first step in the biosynthesis of DMB.[18] The biosynthesis involves an unanticipated intersection of thiamin and vitamin B_{12} biosynthesis and uses chemistry that exceeds in complexity any of the enzymatic reactions used in the assembly of the corrin ring of vitamin B_{12}. Mechanistic and structural studies are in progress to elucidate the details of this remarkable reaction.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b03576.

Detailed experimental procedures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*begley@chem.tamu.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the Robert A. Welch Foundation (A-0034 to T.P.B.) and the National Institutes of Health (DK44083 to T.P.B.; DK067081 to S.E.E.).

■ REFERENCES

(1) Taga, M. E.; Larsen, N. A.; Howard-Jones, A. R.; Walsh, C. T.; Walker, G. C. *Nature (London, U. K.)* **2007**, *446*, 449.
(2) Cooper, A. J. L.; Pinto, J. T. *Chemtracts* **2006**, *19*, 474.
(3) Scott, A. I. *J. Org. Chem.* **2003**, *68*, 2529.
(4) Munder, M.; Vogt, J. R. A.; Vogler, B.; Renz, P. *Eur. J. Biochem.* **1992**, *204*, 679.
(5) Vogt, J. R. A.; Lamm-Kolonko, L.; Renz, P. *Eur. J. Biochem.* **1988**, *174*, 637.
(6) Vogt, J. R. A.; Renz, P. *Eur. J. Biochem.* **1988**, *171*, 655.
(7) Chatterjee, A.; Li, Y.; Zhang, Y.; Grove, T. L.; Lee, M.; Krebs, C.; Booker, S. J.; Begley, T. P.; Ealick, S. E. *Nat. Chem. Biol.* **2008**, *4*, 758.
(8) Chatterjee, A.; Hazra, A. B.; Abdelwahed, S.; Hilmey, D. G.; Begley, T. P. *Angew. Chem., Int. Ed.* **2010**, *49*, 8653.
(9) Lawhorn, B. G.; Mehl, R. A.; Begley, T. P. *Org. Biomol. Chem.* **2004**, *2*, 2538.
(10) Yamada, K.; Kumaoka, H. *J. Nutr. Sci. Vitaminol.* **1983**, *29*, 389.
(11) Renz, P.; Endres, B.; Kurz, B.; Marquart, J. *Eur. J. Biochem.* **1993**, *217*, 1117.
(12) Endres, B.; Würfel, A.; Vogler, B.; Renz, P. *Biol. Chem. Hoppe-Seyler* **1995**, *376*, 595.
(13) Fenwick, M. K.; Mehta, A. P.; Zhang, Y.; Abdelwahed, S. H.; Begley, T. P.; Ealick, S. E. *Nat. Commun.* **2015**, *6*, 6480.
(14) Coquille, S.; Roux, C.; Mehta, A.; Begley, T. P.; Fitzpatrick, T. B.; Thore, S. *J. Struct Biol.* **2013**, *184*, 438.
(15) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.
(16) Cane, D. E.; Xue, Q.; Fitzsimons, B. C. *Biochemistry* **1996**, *35*, 12369.
(17) Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. T. *J. Am. Chem. Soc.* **2007**, *129*, 11213.
(18) Hazra, A. B.; Han, A. W.; Mehta, A. P.; Mok, K. C.; Osadchiy, V.; Begley, T. P.; Taga, M. E. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 2015 Aug 5, pii:201509132 [Epub ahead of print].

DOI: 10.1021/jacs.5b03576
J. Am. Chem. Soc. 2015, 137, 10444−10447