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

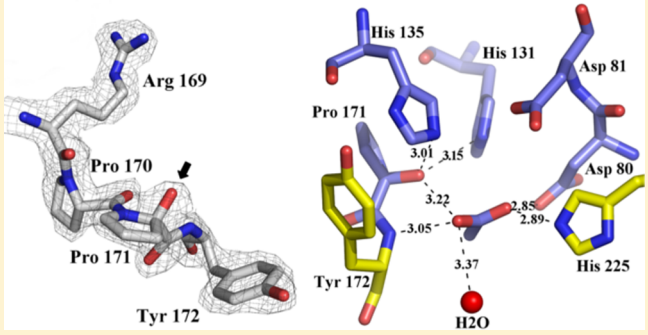
# Unusual $\alpha$-Carbon Hydroxylation of Proline Promotes Active-Site Maturation

Vasiliki E. Fadouloglou,[†,▽] Stavroula Balomenou,[†,‡,▽] Michalis Aivaliotis,[†,▽] Dina Kotsifaki,[†] Sofia Arnaouteli,[‡,#] Anastasia Tomatsidou,[‡] Giorgos Efstathiou,[†] Nikos Kountourakis,[†] Sofia Miliara,[†,⊗] Marianna Griniezaki,[‡] Aleka Tsalafouta,[‡] Spiros A. Pergantis,[§] Ivo G. Boneca,[∥,×] Nicholas M. Glykos,[⊥] Vassilis Bouriotis,*[,†,‡] and Michael Kokkinidis*[,†,‡]

[†]Institute of Molecular Biology and Biotechnology, 70013 Heraklion, Crete, Greece
[‡]Department of Biology, University of Crete, Voutes University Campus, 70013 Heraklion, Crete, Greece
[§]Department of Chemistry, University of Crete, Voutes University Campus, 70013 Heraklion, Crete, Greece
[∥]Biology and Genetics of the Bacterial Cell Wall Unit, Institut Pasteur, 75015 Paris, France
[×]INSERM, Equipe Avenir, Paris, France
[⊥]Department of Molecular Biology and Genetics, Democritus University of Thrace, University Campus, 68100 Alexandroupolis, Greece

Supporting Information

**ABSTRACT:** The full extent of proline (Pro) hydroxylation has yet to be established, as it is largely unexplored in bacteria. We describe here a so far unknown Pro hydroxylation activity which occurs in active sites of polysaccharide deacetylases (PDAs) from bacterial pathogens, modifying the protein backbone at the $C_\alpha$ atom of a Pro residue to produce 2-hydroxyproline (2-Hyp). This process modifies with high specificity a conserved Pro, shares with the deacetylation reaction the same active site and one catalytic residue, and utilizes molecular oxygen as source for the hydroxyl group oxygen of 2-Hyp. By providing additional hydrogen-bonding capacity, the Pro→2-Hyp conversion alters the active site and enhances significantly deacetylase activity, probably by creating a more favorable environment for transition-state stabilization. Our results classify this process as an active-site "maturation", which is highly atypical in being a protein backbone-modifying activity, rather than a side-chain-modifying one.




## INTRODUCTION

Post-translational modifications (PTMs) of proteins result in a substantial increase of the diversity and functional complexity of the proteome.[1] In eukaryotes the post-translational hydroxylation of proline (Pro) side chains by prolyl hydroxylases produces hydroxyproline (Hyp), a major component of collagen[2] which significantly contributes to its stability.[3] Hyp also plays a key role in plant cell wall architecture,[4] in signaling processes linked to hypoxia response,[5] and in physiological pathways associated with diseases such as cancer.[6] In bacteria Pro hydroxylation remains largely understudied[7] and occurs mainly on free Pro.[8] One putative bacterial peptidyl−prolyl hydroxylase has been reported, however, for *Bacillus anthracis*,[9] although its substrate and biological role are presently unknown. The establishment of Pro hydroxylation as a crucial PTM makes the elucidation of its full extent and its various biological roles an important issue, in particular since several recent studies imply[7,10] that hydroxylation may be a significantly more extensive PTM than previously perceived.

We recently identified by X-ray crystallography[11,12] a PTM representing a new type of Pro hydroxylation that targets the main chain of protein molecules, modifying the $C_\alpha$ atom of Pro residues to produce 2-hydroxyproline (2-Hyp). This unusual hydroxylation occurs in the active sites of the putative polysaccharide deacetylases (PDAs) *Bc*0361 and *Ba*0330, from the bacterial pathogens *B. cereus* and *B. anthracis*, respectively. PDAs are members of the carbohydrate esterase family 4 (CE4), and they play crucial roles in various mechanisms of bacterial pathogens, including the evasion of a host's innate immune system.[12] Due to its occurrence in the active sites of PDAs, the new hydroxylation activity could be potentially linked to various aspects of bacterial pathogenicity, thus making its detailed characterization particularly important.

**Received:** November 27, 2016
**Published:** March 23, 2017



DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

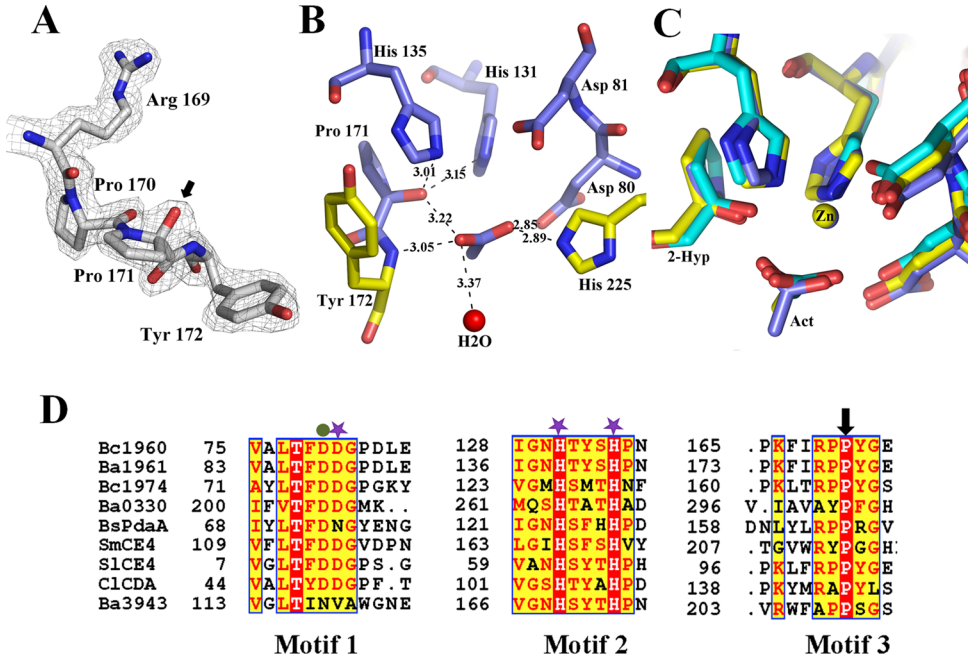



**Figure 1.** 2-Hyp sites in PDAs. (A) $2F_o - F_c$ electron density map of *Bc*1960 at 2.5$\sigma$, showing the 2-Hyp modification (arrow) at Pro171. (B) Active site of *Bc*1960, showing the metal-binding triad (Asp81, His131, His135), the catalytic Asp80, and the hydroxylated Pro171. In the hydrogen-bonding network, also shown are the acetate molecule (Act) and the amido group of residue Tyr172. (C) Superposition of the active sites of *Bc*1960, *Bc*0361, and *Bc*0330. (D) Conserved sequence motifs 1−3 of PDAs showing the conserved Pro residue (black arrow, motif 3) targeted by the hydroxylation reaction, the catalytic Asp (green circle, motif 1), and the metal-binding triad (violet stars, motifs 1 and 2).

Our present study focuses on a comprehensive analysis of this new PTM. We present structural and biochemical information on Pro $C_\alpha$-hydroxylation derived from known or putative PDAs and show that the Pro→2-Hyp conversion is a widespread occurrence in PDAs of the CE4 family affecting a highly conserved Pro residue of the active site. We also provide evidence suggesting that the new Pro hydroxylation mechanism is probably autocatalytic, occurring at the PDA active sites and sharing at least one catalytic residue with the deacetylation reaction. Furthermore, we show that the origin of the hydroxyl group is molecular oxygen and that the new PTM represents an atypical maturation process of PDAs, which establishes an intertwining between the hydroxylation and deacetylation reactions, through which deacetylase activity is enhanced by approximately a factor of 10 for the PDA *Bc*1960.

## EXPERIMENTAL SECTION

**Heterologous Expression and Protein Purification.** *B. cereus* and *B. anthracis* PDAs were heterologously expressed in *Escherichia coli* as previously reported[12−15] and purified by Ni-affinity chromatography. For further details, please refer to SI Materials and Methods.

**Overexpression in *B. anthracis*.** The *ba1961* gene was placed under the transcriptional control of a xylose-inducible promoter, and *Ba*1961 protein was overproduced in *B. anthracis* cells after induction by D-xylose. The *B. anthracis* cells were disrupted by sonication, and protein fractions were collected by centrifugation (see SI Materials and Methods).

***In Vitro* Protein Synthesis.** *Bc*1960 was produced *in vitro* with the S30 T7 High-Yield Protein Expression System (Promega). The following reaction mixtures were set up in a DNase-RNase-free environment: 10 $\mu$L of S30 Premix Plus; 9 $\mu$L of T7 S30 Extract, circular; 1 $\mu$g of DNA *bc1960* or 1 $\mu$g of S30 T7 Positive Control DNA (included in the kit); and $H_2O$ to a total volume of 25 $\mu$L. Reactions were thoroughly mixed, centrifuged for 5 s, and incubated overnight at 25 °C using a shaking rotor. Reactions were terminated by incubating the tubes on ice for 5 min. *Bc*1960 production was confirmed by SDS-PAGE on a 12.5% acrylamide gel.

Cell-free production of the *Bc*1960 protein (a) in $H_2{}^{18}O$ (Aldrich) or (b) in nitrogen atmosphere was used to explore the origin of the oxygen atom of the hydroxyl group in the Pro→2-Hyp conversion:

For procedure (a), *in vitro* synthesis was performed in $H_2{}^{18}O$, and subsequently mass spectrometry was used to characterize the abundance of the isotopes $^{16}O$ and $^{18}O$ at the 2-Hyp residue. To ensure that protein expression occurs in a $H_2{}^{18}O$ environment, the following pretreatment of the system was introduced: specific quantities of the cell-free system reaction reagents that are identical to those described above were separately lyophilized in a SpeedVac concentrator. The reaction mixture was set up by dissolving the ingredients to a final volume of 25 $\mu$L using $H_2{}^{18}O$. Control experiments with the same pretreatment of the cell-free system were also performed for expression of *Bc*1960 in $H_2{}^{16}O$. For procedure (b), cell-free expression was performed in nitrogen atmosphere; mass spectrometry was used to determine the hydroxylation levels of *Bc*1960, which were compared to the levels observed when the process was performed in the presence of oxygen.

**Measurement of Deacetylase Activity in the Presence or Absence of Proline Hydroxylation.** *Bc*1960 was produced *in vitro* with the S30 T7 High-Yield Protein Expression System (Promega) either under aerobic conditions or in nitrogen atmosphere. The enzymatic activity of *Bc*1960 against radiolabeled *O*-hydroxyethylated chitin (glycol chitin) of the protein produced under these conditions was estimated as previously described[15] with only minor changes. Briefly, the *in vitro* reactions (25 $\mu$L) were mixed with 25 mM Mes−NaOH pH 6.0, 1 mM $CoCl_2$, and 5 $\mu$L of glycol chitin (1 mg mL$^{-1}$) radiolabeled in *N*-acetyl groups using $^3$H-acetic anhydride,[16] and the reaction mixtures were incubated for 16 h at 50 °C. This temperature has been previously determined to be optimal for the enzymatic assay of *Bc*1960[15], while the long incubation time is necessitated by the low enzyme concentration obtained from cell-free expression. After reaction termination, enzymatic activity was measured in counts per minute by liquid scintillation counting. An *in vitro* reaction without the expression plasmid of *bc1960* served as negative control. It should be pointed out that the assay used can only provide relative activities (e.g.,


DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

between the hydroxylated and non-hydroxylated forms of *Bc*1960), because radiolabeled glycol chitin is a heterogeneous substrate in terms of molecular weight, degree of chitin substitution by hydroxyethyl groups, and degree of $^{3}H$ labeling.

**X-ray Crystallography.** The crystallization of the *B. cereus* PDA *Bc*1960 and the collection of X-ray diffraction data have been reported earlier.[13] The structure was solved by molecular replacement and refined to 2.3 Å resolution. Details are provided in Table S1.

**Tandem Mass Spectrometry.** Protein/PTM identification and relative quantitation by nanoscale liquid chromatography coupled online with electrospray ionization tandem mass spectrometry (nLC-ESI-MS/MS) was done on a high-mass resolution LTQ-Orbitrap XL (Thermo Fisher Scientific) mass spectrometer coupled to an Easy nLC instrument (Thermo Fisher Scientific). The sample preparation and the LC separation were performed as described in earlier reports[17] with minor modifications. Relative quantitation of hydroxylated versus non-hydroxylated tryptic peptides was performed manually using their MS intensities (see SI Materials and Methods).

**Metal Detection.** The metal content of *Bc*1960 was studied using size exclusion chromatography (SEC) online with inductively coupled plasma mass spectrometry (ICP-MS). The SEC column separated the enzyme-bound metal from the unbound metal. An inductively coupled plasma mass spectrometer (NexION 300xx, PerkinElmer) was employed for metal determination as it eluted from the SEC column. A Flexar HPLC system (PerkinElmer) was used for mobile-phase delivery to the SEC column.

**Identification of Structural Homologies.** Pro residues positioned in environments very similar to those of 2-Hyp residues in the active site of PDAs were identified in the Protein Data Bank (PDB, www.pdb.org) using specific orientational angles that are introduced in Figure 5A and programs written by us (see SI Materials and Methods).

The atomic coordinates and structure factors have been deposited with the PDB (PDB ID: 4L1G).

## RESULTS

**2-Hyp Sites in PDAs Exhibit Extensive Structural Homologies.** Following the crystallographic identification of 2-Hyp[11,12] in two putative PDAs, we undertook a broader investigation of *B. anthracis* and *B. cereus* PDAs to determine whether this PTM is a special feature of the homologous *Bc*0361 and *Ba*0330 proteins or a more widespread property. We crystallized[13] *B. cereus* *Bc*1960, a known PDA, and determined its structure in the metal-free (catalytically inactive) form at 2.3 Å resolution (Figure S1 and Table S1). *Bc*1960 adopts a NodB domain fold.[18−20] The most prominent feature of its active site is a clearly identifiable 2-Hyp residue at the position of the conserved Pro171 (Figures 1A and S1). One acetate molecule (Figure 1B), probably originating from the crystallization buffer, is found in an active-site position, where acetate molecules produced by the deacetylation reaction are expected.[19] The −OH group of 2-Hyp is hydrogen-bonded to active-site/metal-binding-site residues and to the acetate molecule (Figure 1B). Acetate is also hydrogen-bonded to the backbone −NH group of the residue following 2-Hyp (Tyr172). The hydrogen bonds formed between the −OH/−NH groups and the acetate occur in the "oxyanion hole" which is crucial for the stabilization of the transition state in the deacetylation reaction.[19] The *Bc*1960 active site exhibits extensive similarities to its counterparts from *Bc*0361 and *Bc*0330 (Figure 1C). The 2-Hyp residues are similarly positioned in the interior of the active site, with the hydrogen-bonding network being broadly conserved. Structural homologies (Figure 1C) and multiple sequence alignments (Figure 1D) reveal a relation between conserved sequence motifs[19] and the site of the new PTM: The hydroxylation activity modifies with utmost specificity the highly conserved Pro of motif 3 (Figure 1D), for which no role had been previously proposed.

**Partial Hydroxylation of the Conserved Pro Is a Frequent Occurrence in PDAs of Various Bacteria.** We subsequently mapped and quantified Hyp sites for several *B. anthracis* and *B. cereus* PDAs (Table 1) using nLC-ESI-MS/MS.

**Table 1. Mass Spectrometric Characterization of the Pro Hydroxylation Level in PDAs from *B. anthracis* and *B. cereus*, Produced by Different Expression Systems and under Different Conditions**

| source organism/expression system | enzyme (conserved Pro) | % Hyp[a] |
|---|---|---|
| *B. cereus*/*E. coli* (freshly produced protein) | *Bc*1960 (Pro171) | 80 ± 9 |
| *B. cereus*/*E. coli* (1 yr old protein stored at 4 °C) | | 95 ± 2 |
| *B. cereus*/cell-free expression (aerobic) | | 72 ± 9 |
| *B. cereus*/cell-free expression (in $H_2{}^{18}O$) | | 13 ± 3 |
| *B. cereus*/cell-free expression (anaerobic) | | 1.4 ± 0.3 |
| *B. anthracis*/*B. anthracis* | *Ba*1961 (Pro169) | 23 ± 7 |
| *B. cereus*/*E. coli* | *Bc*1974 (Pro166) | 31 ± 6 |
| | *Bc*1974 D77N | 26 ± 6 |
| *B. anthracis*/*E. coli* | *Ba*0330 (Pro302) | 63 ± 15 |
| | *Ba*0330 D205A | 6 ± 2 |
| | *Ba*3943 (Pro209) | 9 ± 2 |

[a]The percent occurrence of 2-Hyp refers to the conserved Pro position (given in parentheses in column 2) from motif 3, which is the sole prominent hydroxylation target.

As shown in Figures 2 and S2, MS/MS unambiguously confirmed the crystallographically observed hydroxylation of Pro171 in *Bc*1960 and of Pro302 in *Ba*0330. Pro hydroxylation is partial, occurring at a level of approximately 80% for Pro171 and 63% for Pro302. Partial hydroxylation of Pro characterizes not only the novel $C_\alpha$-targeting activity observed by us but also the more common hydroxylation of Pro side chains catalyzed by prolyl hydroxylases.[21−23]

For all other Pro residues of *Bc*1960 and *Ba*0330 (including the Pro residue which immediately precedes 2-Hyp), hydroxylation does not occur above noise level, confirming the remarkable specificity of the Pro→2-Hyp conversion. For *Bc*1974, the conserved Pro166 is modified at a level of 31% (Table 1). On the other hand, we detected no Pro hydroxylation in *Ba*3943 (Table 1), an inactive deacetylase lacking conserved residues of the active site, including the key Asp residues of motif 1 (Figure 1D).

To rule out the possibility of hydroxylation being an artifact of the heterologous expression in *E. coli*, we overexpressed *Ba*1961, a homologue of *Bc*1960, in the host organism *B. anthracis*, and similarly, we detected significant levels of Pro hydroxylation (Table 1).

**The Catalytic Asp of the Deacetylation Reaction Is Crucial Also for the Pro Hydroxylation Reaction.** To explore possible links between the highly conserved Asp residues in motif 1 (Figure 1D) and hydroxylation, we studied two mutations: (i) Mutant D205A in *Ba*0330 is the Ala substitution of the catalytic Asp205, which is critical for the deacetylation reaction, since mutations at this position completely abolish PDA activity.[19,24,25] (ii) Mutant D77N in *Bc*1974 substitutes Asp77 of motif 1, a residue from the metal-binding triad, by Asn. This substitution is consistent with the amino acid preferences of the PDA metal-binding site (Figure


DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

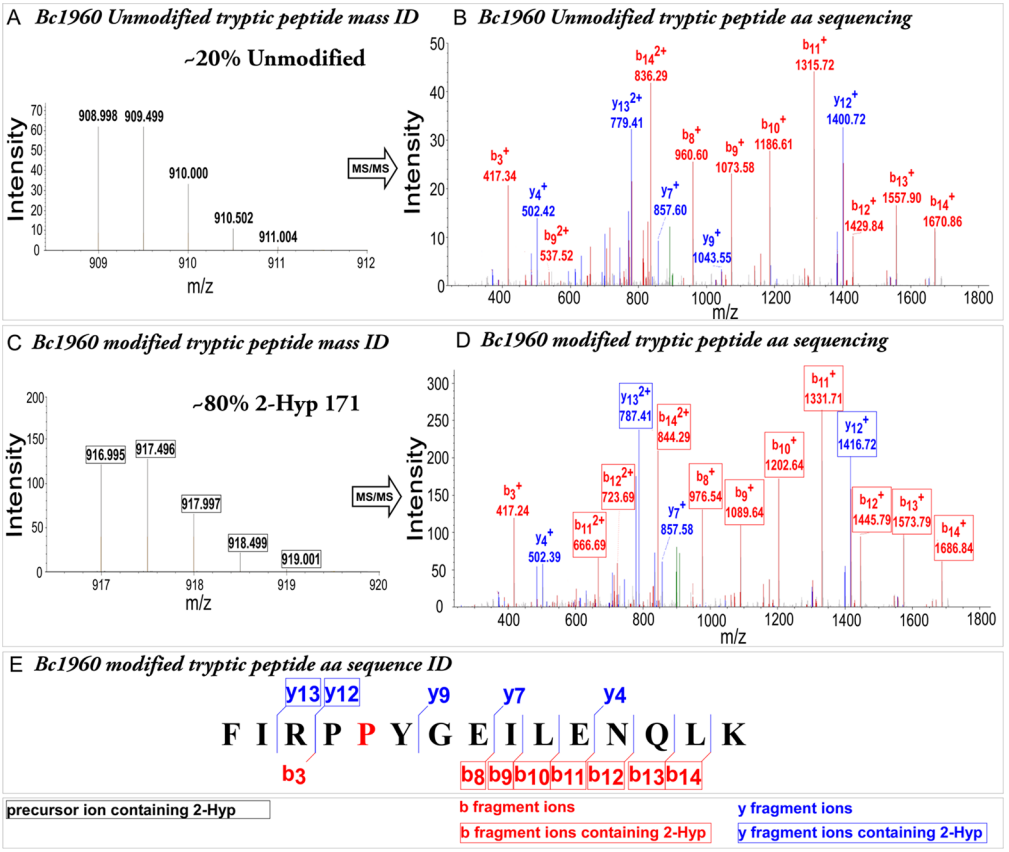



**Figure 2.** Mass spectrometric identification and relative quantitation of residue 2-Hyp 171 in *Bc*1960 by nLC-ESI-MS/MS. The protein was found to be partially hydroxylated on residue Pro171 by the identification of modified and unmodified tryptic peptides containing Pro171 after proteolysis and subsequent nLC-ESI-MS/MS analysis. The exact mass of the modified tryptic peptide (C) was found to be 16 Da higher than the exact mass of the unmodified peptide (A) by the high mass resolution and accuracy Orbitrap analyzer, corresponding to the substitution of a hydrogen by a hydroxyl group. The peptides were subsequently sequenced by tandem mass spectrometric analysis (MS/MS) for protein identification and the detection of the exact site of hydroxylation on Pro171 (B,D,E). The relative abundance of the 2-Hyp form in position 171 is approximately 80%.

1D).[18] The two Asp mutations affect the hydroxylation activity at the conserved Pro of motif 3 very differently: Pro hydroxylation nearly disappears in D205A, while it is only slightly reduced in D77N.

The catalytic Asp of motif 1 has thus a key role not only in deacetylation but also in Pro hydroxylation. On the other hand, hydroxylation is insensitive to conservative Asp→Asn substitutions in the metal binding triad, although deacetylation against glycol chitin is significantly reduced. The absence of Pro hydroxylation in *Ba*3943 is also consistent with the substitution of the catalytic Asp by Asn.

**Cell-Free Expression Supports an Autocatalytic Pro Hydroxylation Process.** *Bc*1960 was also produced *in vitro*, and a high level of hydroxylation (72%) at the conserved Pro was detected (Table 1). This is comparable to the hydroxylation level (80%) observed in *E. coli*. It is thus possible that the $C_\alpha$-hydroxylation of Pro is catalyzed not by another enzyme but rather through an intrinsic, i.e., autocatalytic, hydroxylation process. Self-hydroxylation is also supported by the crucial importance on the hydroxylation level of one active-site residue, the catalytic Asp (Table 1, *Ba*0330, *Ba*3943). Furthermore, since cell-free expression was performed in the absence of PDA substrate, it must be concluded that, for the $C_\alpha$-hydroxylation of Pro, the deacetylation reaction is not required.

**The Pro→2-Hyp Reaction Depends on Molecular Oxygen.** The origin of the oxygen atom of the 2-Hyp −OH group was elucidated using cell-free expression of *Bc*1960. Two possible sources for oxygen were considered, i.e., water and molecular (atmospheric) oxygen. For cell-free expression performed in $H_2{}^{18}O$, mass spectrometry revealed a reduced hydroxylation level for Pro171; reduced hydroxylation was also observed in the control experiment performed in $H_2{}^{16}O$ (see Experimental Section). This reduction can be probably attributed to the specific pretreatment of the cell-free system in both cases, which includes a lyophilization step. However, for the 2-Hyp residues formed in the environment of $H_2{}^{18}O$, the incorporation of only the isotope $^{16}O$ and not $^{18}O$ could be detected (Figure 3). On the other hand, cell-free expression under nitrogen atmosphere, i.e., in the absence of atmospheric oxygen, reduced the hydroxylation of Pro to practically noise levels (Table 1, Figure 3). These results preclude water as the origin of the 2-Hyp −OH group and strongly suggest that molecular oxygen is the source of the −OH group oxygen.

**Pro Hydroxylation Is an Active-Site Maturation Process.** 2-Hyp introduces via its −OH group an additional hydrogen-bond donor in the oxyanion hole (Figure 1). The hydrogen-bonding network in the active site (Figure 1B) suggests that this −OH is probably involved in interactions stabilizing the oxyanion intermediate,[19] as required by the catalytic mechanism of PDAs.

A possible intertwining of the deacetylation and hydroxylation reactions is shown in Figure 4: The 2-Hyp residue (Figure 4A) gives rise to a hydrogen bond (Figure 4B), which


DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

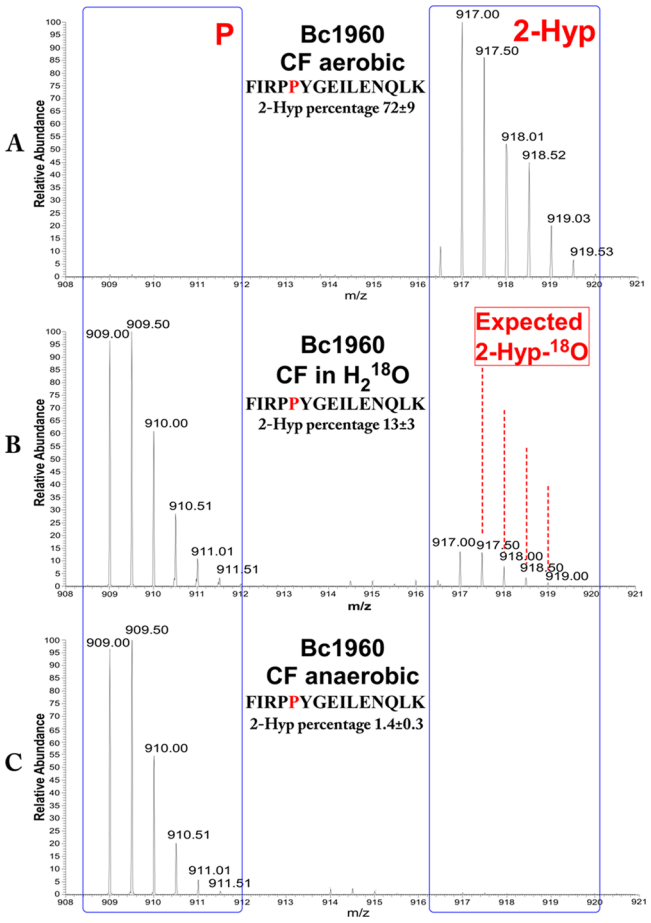



**Figure 3.** Occurrence of hydroxyproline at position 171 after cell-free (CF) expression of *Bc*1960. Mass spectra of the tryptic peptide containing residue 171 are shown for protein expression under normal (aerobic) conditions (A), in the presence of $H_2{}^{18}O$ (B), and anaerobically in a nitrogen atmosphere (C). The hydroxylated (2-Hyp) and non-hydroxylated (P) forms of Pro171 are shown. Relative abundances (Table 1) are suggested by the intensity of their MS signals. In (B), red lines indicate the expected signal of a 2-Hyp group with an $^{18}OH$ group (2-Hyp-$^{18}O$). Since no $^{18}O$ was detected, water was ruled out as origin of the O atom of the 2-Hyp hydroxyl group. In (C), the absence of Pro171 hydroxylation for *in vitro* protein expression under nitrogen atmosphere suggests that the O atom of the 2-Hyp hydroxyl group originates from molecular (atmospheric) oxygen.

provides additional stabilization to the transition state of the deacetylation reaction, complementing its interactions with the divalent metal and with the backbone −NH group (from the residue following 2-Hyp). The deacetylation mechanism in Figure 4 is based upon earlier work[26] on zinc-dependent deacetylases: $Zn^{2+}$ plays the role of Lewis acid, and Asp and His residues play the roles of a catalytic base, activating the nucleophilic water (magenta), and of an acid, respectively, which aid the generation (Figure 4C) of a free amine and of an acetate product which is hydrogen-bonded to 2-Hyp and the main-chain −NH group, as observed in the crystal structure (Figure 1B).

Thus, the 2-Hyp interactions complement those involving the −NH group of the residue following the conserved Pro, e.g., Tyr172 in *Bc*1960 (Figure 1B), which have been suggested earlier as transition-state stabilizers.[19] The additional capacity for transition-state stabilization provided via the formation of the 2-Hyp residue represents a "maturation" process of PDA active sites (Figure 4), which is expected to affect the deacetylase activity of the enzyme. The implications of the enzyme maturation process were tested by comparing the enzymatic activities of the hydroxylated and non-hydroxylated forms of *Bc*1960. The protein was produced *in vitro* by a cell-free system, either under aerobic conditions, favoring the formation of 2-Hyp at position Pro171, or in nitrogen atmosphere, which suppresses Pro $C_\alpha$-hydroxylation (Table 1). The relative deacetylase activities of the two forms were estimated using the enzymatic assay described in the Experimental Section. Some parameters of this assay are optimized for the low enzyme concentration produced by the cell-free expression system (long incubation time) and the specific properties of the *Bc*1960 enzyme (assay temperature).[15] The deacetylase activity of the highly hydroxylated (aerobically produced) form of *Bc*1960 shows an enhancement by approximately a factor of 10 relative to the activity of the non-hydroxylated (anaerobically produced) form. It may be thus concluded that the Pro hydroxylation reaction is functionally strongly intertwined with the deacetylation reaction. Both activities share the PDA active site and the catalytic Asp residue. Apart from relative activities, however, it was not possible to measure the kinetic properties of the two forms of *Bc*1960, since radiolabeled glycol chitin is a heterogeneous substrate, as pointed out in the Experimental Section. Likewise, an alternative assay[19] based on fluorogenic labeling of the free amines generated by deacetylase activity

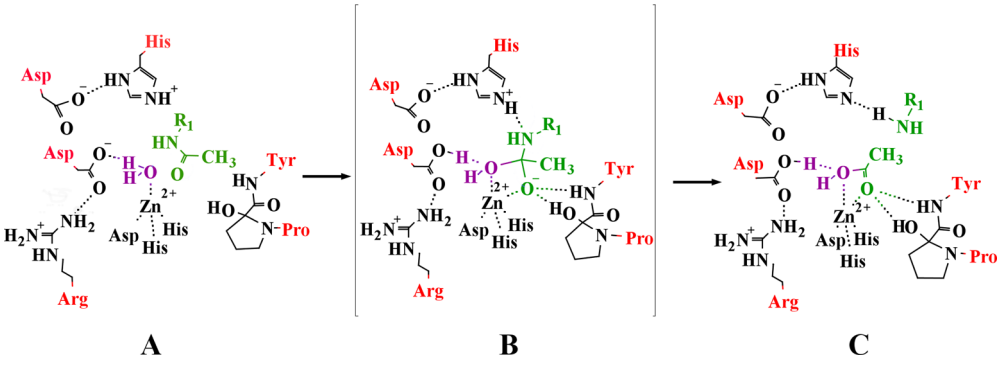



**Figure 4.** Proposed intertwining of deacetylation with the Pro→2-Hyp conversion in PDAs. All residues shown are conserved in PDAs with the exception of the residue following 2-Hyp (Tyr in the case of *Bc*1960); however, it contributes to the interaction network shown exclusively through the protein backbone −NH group, which is invariant.



DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

with *N*-acetyl chitooligosaccharides is not suitable, as the two *Bc*1960 forms produced by the cell-free expression system are not purified to homogeneity.

**Potential Extent of the New PTM.** Our results raise the question of whether the new form of Pro hydroxylation can also occur in other proteins. The absence of 2-Hyp in earlier published crystal structures of CE4 family members could simply reflect relatively low hydroxylation levels and thus low crystallographic occupancy for 2-Hyp residues, thus complicating their crystallographic identification on electron density maps. For this reason we searched the Protein Data Bank (www.pdb.org) for protein structures comprising geometries similar to those shown in Figure 1C. The analysis (SI Materials and Methods, Figure 5A) revealed a cluster of structures (Figure 5B), each including a Pro residue positioned in a very similar orientation relative to the metal-binding site (mostly a $Zn^{2+}$ site), as in Figure 1C. The cluster consists overwhelmingly of deacetylases of the CE4 family for which no 2-Hyp residues have been reported. Accordingly, these proteins (Table S2) are prime candidates for MS analysis aiming at the discovery of additional enzymes converting Pro to 2-Hyp.

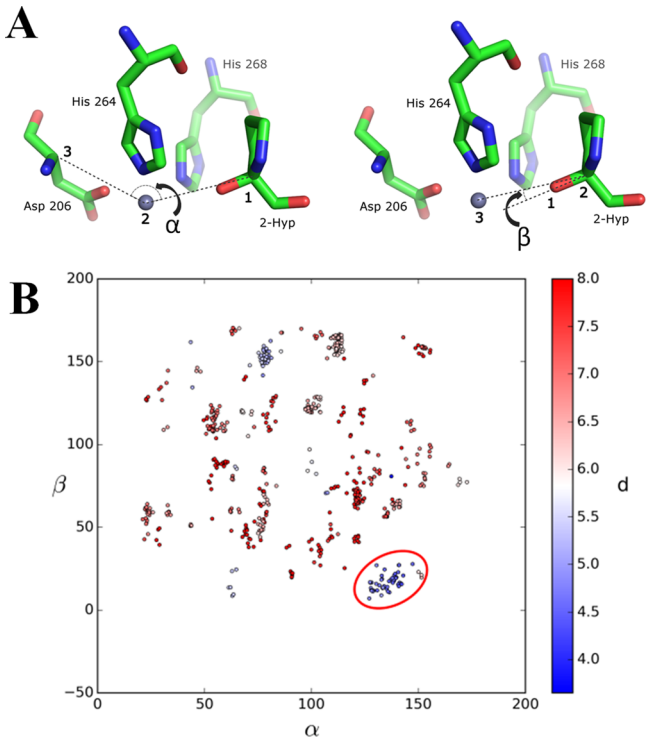



**Figure 5.** (A) Definition of the orientational angles $\alpha$ and $\beta$ for the *Bc*0361 structure (PDB ID: 4HD5). $Zn^{2+}$ (sphere), the metal-coordinating triad, and the 2-Hyp residue are shown. The orientational angles $\alpha$, and $\beta$ are defined as follows: angle $\alpha$ is formed by the Pro $C_{\alpha}$, the metal, and Asp $C_{\alpha}$; angle $\beta$ is formed by the Pro $C_{\alpha}$-hydrogen (or Pro $C_{\alpha}$-oxygen in the case of a 2-Hyp residue), Pro $C_{\alpha}$, and metal atoms. The position of the $C_{\alpha}$-bonded hydrogen atom is calculated via PyMol. (B) Structural homologies of 2-Hyp sites: identification of Pro residues positioned in an environment similar to that of 2-Hyp in active sites of PDAs (see Figure 1B). PDB structures with Pro residues in the vicinity of divalent metal binding sites (Zn, Fe, Co, Cd, Ni, etc.) and with the Pro−metal distance not exceeding 8 Å are shown as dots in the diagram, which gives the Pro−metal distance and the two orientation parameters ($\alpha$, $\beta$) relative to the metal binding site. The color code for the metal−$C_{\alpha}$(Pro) distance is shown on the vertical bar. The cluster of Pro-comprising sites that are structurally very similar to active sites of 2-Hyp-containing PDAs is included in the red ellipse and in Table S2.

## DISCUSSION

Our results provide insights into the $C_{\alpha}$-hydroxylation of Pro residues, a new and highly atypical PTM which has been identified in the active sites of PDAs from pathogenic bacteria. This PTM enhances the deacetylase activity of the enzymes.

Many details of the hydroxylation mechanism are still unclear. Its dependence on molecular oxygen has some analogies to that of Fe(II)-utilizing mechanisms of prolyl hydroxylases from the superfamily of Fe(II)- and 2-oxoglutarate-dependent dioxygenases.[27] The fact that the active-site $Zn^{2+}$ of deacetylases is known to occasionally switch to Fe(II)[28] could be suggestive of an autocatalytic mechanism for the $C_{\alpha}$-hydroxylation of Pro utilizing the presence of Fe(II). Therefore, we used, as described in the Experimental Section, SEC online with ICP-MS for metal detection in *Bc*1960 (Figure S3). The method has been used successful in various studies for the detection of Fe and other metals in similar proteins.[29,30] The enzyme-bound metals were detected at around 445 s elution time, whereas unbound metals were detected at around 540 s elution time. The analysis revealed the presence of enzyme-bound $Zn^{2+}$ in *Bc*1960, which can be removed by EDTA. However, no uptake of Fe(II) by metal-free (EDTA-treated) *Bc*1960 could be detected. The observed peak profile when Fe(II) was added to the protein was similar to the profile observed for the ionic Fe standard which was also analyzed, suggesting that only unbound Fe was detected. This behavior may reflect either very weak or no Fe(II) binding to *Bc*1960, as is the case with many other enzymes,[31] but it could also suggest that the $C_{\alpha}$-hydroxylation mechanism in PDAs is not comparable to Fe(II)-utilizing mechanisms.

Thus, for the $O_2$-dependent $C_{\alpha}$-hydroxylation of Pro, a metal-independent mechanism similar to mechanisms described in a number of reports for a small class of oxygenation enzymes[32−34] could provide a reasonable explanation for our observations.

The $C_{\alpha}$-hydroxylation of the conserved Pro of motif 3 is partial and the factors which determine the hydroxylation level are not clear yet. Hydroxylation may vary considerably (Table 1), depending on the specific enzyme, the particular protein expression system, and the experimental conditions. For example, lyophilization of the cell-free system for *Bc*1960 expression appears to affect the $C_{\alpha}$-hydroxylation of the conserved Pro by lowering the hydroxylation level.

Although the variation of hydroxylation levels for different enzymes is not understood yet, it can be speculated that it can be at least in part attributed to differences in the structures and sequences of the active sites where the $C_{\alpha}$-hydroxylation reaction occurs. Specific properties of the protein expression systems probably also play a role, as well as the localization pattern of each enzyme in the cell, as its microenvironment could affect the hydroxylation reaction. The time frame of the Pro→2-Hyp conversion might also affect the observed hydroxylation level; this aspect will be addressed by future experiments.

Our results also provide evidence that the protein-backbone-modifying Pro→2-Hyp conversion in PDAs might be autocatalytic. Autocatalytic hydroxylation mechanisms exist in many other enzymes, in which, however, only side chains are modified.[35]


DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

The only other known case of $C_\alpha$-hydroxylation is the enzymatic conversion of Gly-extended precursor peptides (prohormones) to a peptidyl-$\alpha$-hydroxyglycine intermediate state that is subsequently decomposed.[36] In contrast to the latter process, however, the Pro $C_\alpha$-hydroxylation reported here produces stably modified proteins, which may retain very high hydroxylation levels even after storage for a period of 1 year at 4 °C (Table 1).

The most important aspect of the new PTM is probably the enhancement of deacetylase activity via maturation of the active site through modification of the protein backbone. All known post-translational maturation processes in other enzymes involve side-chain modifications.[37,38] Active-site maturation processes in the absence of PTMs, e.g., the maturation of [NiFe]-hydrogenase via metal chaperones,[39] are rare.

Cases of active sites shared by two activities, as we found in PDAs for hydroxylation and deacetylation, are also very rare and include the multifunctional Pk-Rec enzyme, in which the ATPase and deoxyribonuclease activities share the same site.[40] Also, in the human O-linked *N*-acetylglucosamine (O-GlcNAc) transferase, the activities of protein glycosylation and Host Cell Factor-1 (HCF-1) cleavage share the same active site.[41]

In summary, 2-Hyp expands the role of Pro in proteins, which is usually a structural one. Through its additional hydrogen-bonding capacity, 2-Hyp can acquire the role of a functional/catalytic residue in PDAs, a property so far associated only with non-proteinogenic Pro[42] in chemical catalysis. Furthermore, the −OH group of 2-Hyp is expected to affect its accessible conformational space through the introduction of additional steric constraints[43] compared to normal amino acids. The implications in protein folding, stability, and catalysis remain to be explored.

It is now evident that Nature employs various strategies for modifying catalytic sites in proteins, with the post-translational Pro→2-Hyp conversion being one example. The properties of 2-Hyp reveal previously unanticipated aspects of the catalytic mechanisms of deacetylases and raise important questions regarding the evolution and full extent of hydroxylation processes in proteins. In particular, our observations provide more refined concepts about the important role of Pro hydroxylation in the enzymatic activity of PDAs. The physiological role of the intertwining of the deacetylation and the Pro $C_\alpha$-hydroxylation reactions can be only speculated upon for some PDAs, as the exact roles of several *B. cereus* and *B. anthracis* PDAs are still unclear. For example, *Ba*1977 from *B. anthracis* (a homologue of *Bc*1974 from *B. cereus*) is a *bona fide* peptidoglycan deacetylase, which contributes to the resistance of the pathogen to host lysozyme via deacetylation of *N*-acetylglucosamine (GlcNAc) residues from the protective peptidoglycan (PG) layer of the bacterial cell wall. The enzyme localizes extracellularly at the cell membrane,[14] being accessible to molecular oxygen. $C_\alpha$-hydroxylation of the conserved Pro could enhance the deacetylase activity, and thus also the protective properties of *Ba*1977, by making it a more efficient modifier of the PG barrier compared to the non-hydroxylated form. This would make the intertwining of deacetylation and $C_\alpha$-hydroxylation an important element of bacterial pathogenicity. Proline hydroxylation may thus provide an attractive new drug target, with applications in the control of a wide range of pathogens such as *B. anthracis* and *B. cereus*.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b12209.

Detailed materials and methods for protein expression and purification in *E. coli* and host organism, measurement of enzymatic activity, tandem mass spectrometry, and identification of structural homologies (PDF)

X-ray crystallographic data for *Bc*1960 (CIF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

*bouriotis@imbb.forth.gr

*kokkinid@imbb.forth.gr

### ORCID

Nicholas M. Glykos: 0000-0003-3782-206X

Michael Kokkinidis: 0000-0002-3112-6904

### Present Addresses

#S.A.: Division of Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

⊗S.M.: Department of Biosciences and Nutrition, Karolinska Institute, 141 57 Huddinge, Sweden

### Author Contributions

▽V.E.F., S.B., and M.A. contributed equally to this work.


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the EU/FP7 project *InnovCrete* and the GSRT projects SYN11_1_31, SYN11_3_1321, THALIS, and 12CHN396. We thank Dr. K. Petratos and Prof. N. Panopoulos for helpful discussions and the beamline staff at ALBA/Spain for assistance with data collection.

## ■ REFERENCES


(1) Walsh, T. C. *Posttranslational modification of proteins: expanding nature's inventory*; Roberts and Co.: Englewood, CO, 2006.

(2) Myllyharju, J.; Kivirikko, K. I. *Ann. Med.* **2001**, *33*, 7.

(3) Engel, J., Bachinger, H. P., Muller, P. K. *Collagen*; Topics in Current Chemistry *247*; Springer: Berlin, 2005; p 7

(4) Lamport, D. T.; Kieliszewski, M. J.; Chen, Y.; Cannon, M. C. *Plant Physiol.* **2011**, *156*, 11.

(5) Palmer, B. F.; Clegg, D. J. *Mol. Cell. Endocrinol.* **2014**, *397*, 51.

(6) Masson, N.; Ratcliffe, P. J. *Cancer Metab* **2014**, *2*, 3.

(7) van Staalduinen, L. M.; Jia, Z. *Front. Microbiol.* **2015**, *5*, 798.

(8) Gorres, K. L.; Raines, R. T. *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 106.

(9) Culpepper, M. A.; Scott, E. E.; Limburg, J. *Biochemistry* **2010**, *49*, 124.

(10) Loenarz, C.; Schofield, C. J. *Trends Biochem. Sci.* **2011**, *36*, 7.

(11) Fadouloglou, V. E.; Kapanidou, M.; Agiomirgianaki, A.; Arnaouteli, S.; Bouriotis, V.; Glykos, N. M.; Kokkinidis, M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 276.

(12) Arnaouteli, S.; Giastas, P.; Andreou, A.; Tzanodaskalaki, M.; Aldridge, C.; Tzartos, S. J.; Vollmer, W.; Eliopoulos, E.; Bouriotis, V. *J. Biol. Chem.* **2015**, *290*, 13465.

(13) Tsalafouta, A.; Psylinakis, E.; Kapetaniou, E. G.; Kotsifaki, D.; Deli, A.; Roidis, A.; Bouriotis, V.; Kokkinidis, M. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2008**, *64*, 203.

(14) Balomenou, S.; Fouet, A.; Tzanodaskalaki, M.; Couture-Tosi, E.; Bouriotis, V.; Boneca, I. G. *Mol. Microbiol.* **2013**, *87*, 867.




DOI: 10.1021/jacs.6b12209
J. Am. Chem. Soc. 2017, 139, 5330−5337

(15) Psylinakis, E.; Boneca, I. G.; Mavromatis, K.; Deli, A.; Hayhurst, E.; Foster, S. J.; Varum, K. M.; Bouriotis, V. *J. Biol. Chem.* **2005**, *280*, 30856.
(16) Araki, Y.; Ito, E. *Eur. J. Biochem.* **1975**, *55*, 71.
(17) Aivaliotis, M.; Gevaert, K.; Falb, M.; Tebbe, A.; Konstantinidis, K.; Bisle, B.; Klein, C.; Martens, L.; Staes, A.; Timmerman, E.; Van Damme, J.; Siedler, F.; Pfeiffer, F.; Vandekerckhove, J.; Oesterhelt, D. *J. Proteome Res.* **2007**, *6*, 2195.
(18) Blair, D. E.; van Aalten, D. M. *FEBS Lett.* **2004**, *570*, 13.
(19) Blair, D. E.; Schuttelkopf, A. W.; MacRae, J. I.; van Aalten, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15429.
(20) Balomenou, S.; Arnaouteli, S.; Koutsioulis, D.; Fadouloglou, V. E.; Bouriotis, V. Polysaccharide Deacetylases: New Antibacterial Drug Targets. In *Frontiers in Anti-Infective Drug Discovery*, Vol. 4; Atta-ur-Rahman, Choudhary, M. I., Eds.; Bentham eBooks, 2015; pp 68−130, DOI: 10.2174/9781681080826115040l.
(21) Xu, X.; Gan, Q.; Clough, R. C.; Pappu, K. M.; Howard, J. A.; Baez, J. A.; Wang, K. *BMC Biotechnol.* **2011**, *11*, 69.
(22) Chan, S. W.; Greaves, J.; Da Silva, N. A.; Wang, S. W. *BMC Biotechnol.* **2012**, *12*, 51.
(23) Li, Z.; Wang, Y.; Yao, Q.; Justice, N. B.; Ahn, T. H.; Xu, D.; Hettich, R. L.; Banfield, J. F.; Pan, C. *Nat. Commun.* **2014**, *5*, 4405.
(24) Fadouloglou, V. E.; Deli, A.; Glykos, N. M.; Psylinakis, E.; Bouriotis, V.; Kokkinidis, M. *FEBS J.* **2007**, *274*, 3044.
(25) Deli, A.; Koutsioulis, D.; Fadouloglou, V. E.; Spiliotopoulou, P.; Balomenou, S.; Arnaouteli, S.; Tzanodaskalaki, M.; Mavromatis, K.; Kokkinidis, M.; Bouriotis, V. *FEBS J.* **2010**, *277*, 2740.
(26) Hernick, M.; Fierke, C. A. *Arch. Biochem. Biophys.* **2005**, *433*, 71.
(27) Longbotham, J. E.; Levy, C.; Johannissen, L. O.; Tarhonskaya, H.; Jiang, S.; Loenarz, C.; Flashman, E.; Hay, S.; Schofield, C. J.; Scrutton, N. S. *Biochemistry* **2015**, *54*, 6093.
(28) Gattis, S. G.; Hernick, M.; Fierke, C. A. *J. Biol. Chem.* **2010**, *285*, 33788.
(29) Yang, Z.; Chi, X.; Funabashi, M.; Baba, S.; Nonaka, K.; Pahari, P.; Unrine, J.; Jacobsen, J. M.; Elliott, G. I.; Rohr, J.; Van Lanen, S. G. *J. Biol. Chem.* **2011**, *286*, 7885.
(30) Yang, Z.; Unrine, J.; Nonaka, K.; Van Lanen, S. G. *Methods Enzymol.* **2012**, *516*, 153.
(31) Cotruvo, J. A., Jr.; Stubbe, J. *Metallomics* **2012**, *4*, 1020.
(32) Du, Y. L.; Singh, R.; Alkhalaf, L. M.; Kuatsjah, E.; He, H. Y.; Eltis, L. D.; Ryan, K. S. *Nat. Chem. Biol.* **2016**, *12*, 194.
(33) Widboom, P. F.; Fielding, E. N.; Liu, Y.; Bruner, S. D. *Nature* **2007**, *447*, 342.
(34) Di Russo, N. V.; Condurso, H. L.; Li, K.; Bruner, S. D.; Roitberg, A. E. *Chem. Sci.* **2015**, *6*, 6341.
(35) Liu, A.; Ho, R. Y.; Que, L., Jr.; Ryle, M. J.; Phinney, B. S.; Hausinger, R. P. *J. Am. Chem. Soc.* **2001**, *123*, 5126.
(36) Takahashi, K.; Satani, M.; Gao, Y.; Noguchi, M. *Int. Congr. Ser.* **2002**, *1233*, 235.
(37) Davidson, V. L. *Biochemistry* **2007**, *46*, 5283.
(38) Davidson, V. L. *Mol. BioSyst.* **2011**, *7*, 29.
(39) Chan, K. H.; Lee, K. M.; Wong, K. B. *PLoS One* **2012**, *7*, e32592.
(40) Rashid, N.; Morikawa, M.; Kanaya, S.; Atomi, H.; Imanaka, T. *FEBS Lett.* **1999**, *445*, 111.
(41) Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S. *Science* **2013**, *342*, 1235.
(42) Panday, S. K. *Tetrahedron: Asymmetry* **2011**, *22*, 1817.
(43) Porter, L. L.; Rose, G. D. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 109.


DOI: 10.1021/jacs.6b12209
*J. Am. Chem. Soc.* 2017, 139, 5330−5337