Dioxygenases

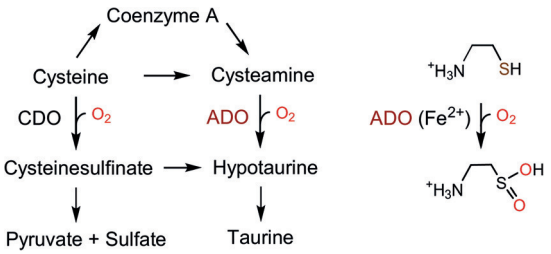
Deutsche Ausgabe: DOI: 10.1002/ange.201803907
Internationale Ausgabe: DOI: 10.1002/anie.201803907

Cofactor Biogenesis in Cysteamine Dioxygenase: C–F Bond Cleavage with Genetically Incorporated Unnatural Tyrosine

*Yifan Wang, Wendell P. Griffith, Jiasong Li, Teruaki Koto, Daniel J. Wherritt, Elizabeth Fritz, and Aimin Liu**

Abstract: *Cysteamine dioxygenase (ADO) is a thiol dioxygenase whose study has been stagnated by the ambiguity as to whether or not it possesses an anticipated protein-derived cofactor. Reported herein is the discovery and elucidation of a Cys-Tyr cofactor in human ADO, crosslinked between Cys220 and Tyr222 through a thioether (C–S) bond. By genetically incorporating an unnatural amino acid, 3,5-difluoro-tyrosine (F_2-Tyr), specifically into Tyr222 of human ADO, an autocatalytic oxidative carbon–fluorine bond activation and fluoride release were identified by mass spectrometry and ^{19}F NMR spectroscopy. These results suggest that the cofactor biogenesis is executed by a powerful oxidant during an autocatalytic process. Unlike that of cysteine dioxygenase, the crosslinking results in a minimal structural change of the protein and it is not detectable by routine low-resolution techniques. Finally, a new sequence motif, C-X-Y-Y(F), is proposed for identifying the Cys-Tyr crosslink.*

The nonheme iron-dependent enzymatic oxidation of 2-amino-ethanethiol (cysteamine) to hypotaurine has been known since 1963 (Scheme 1).[1] The responsible mammalian enzyme, cysteamine dioxygenase (ADO), was purified to homogeneity in 1971, and the characterization of the human enzyme was described in 2007.[2] Human ADO is one of two thiol dioxygenases, alongside cysteine dioxygenase (CDO), which plays critical roles in maintaining proper thiol levels in cells, especially in managing metabolic cysteine and taurine concentrations. The dysfunction of ADO is associated with autoimmune and neurological conditions,[3] fat metabolism,[4] and oxidative stress.[5]



Scheme 1. Thiol dioxygenases in cellular thiol metabolism. The cysteamine dioxygenase (ADO) reaction is shown on the right.

The presence of a protein-derived cysteine-tyrosine (Cys-Tyr) cofactor crosslinked between two distant residues of the protein sequence, Cys93 and Tyr157, was revealed a decade ago for murine, rat, and human CDOs from independent studies by using X-ray crystallography.[6] The Cys-Tyr cofactor, formed through an autocatalytic process, increases the dioxygenase catalytic efficiency of CDO by an order of magnitude.[7]

ADO is a paralogue of CDO.[2b] These two thiol dioxygenases belong to the Cupin superfamily with a conserved β-barrel fold[8] and share similar S-atom-based dioxygenation with parallel substrate structures. Thus, it is sensible to anticipate that a corresponding Cys-Tyr cofactor is also formed through a similar self-processing reaction in human ADO, as previously found in CDO. The difficulty in crystallizing ADO hinders our further understanding towards this enzyme, especially if it possesses a Cys-Tyr cofactor. Notably, after nearly over a half-century since the discovery of ADO, its catalytic machinery remains elusive. The Cys-Tyr cofactor does not give rise to any discernible spectroscopic signature with which its presence can be exclusively identified and quantified. If the cofactor is present, its formation is likely to be an uncoupled, single-turnover reaction out of hundreds of cysteamine oxygenation (i.e., the coupled reaction) cycles. With a ferrous center, ADO is not amenable to characterization by optical spectroscopy since the absorbance of both its substrate and product overlap with the protein absorbance. Thus, the answer to the question as to whether or not a Cys-Tyr cofactor is present in ADO entails strenuous effort.

An EMBOSS pairwise alignment[9] suggested that ADO and CDO share 15.8% sequence identity and 26.5% similarity (see Figure S1 in the Supporting Information). Besides the preserved 3-His iron-binding motif, there is a pair of the conserved tyrosine and cysteine residues (Tyr40 and Cys130 in CDO, Tyr91 and Cys183 in ADO). However, both Tyr40 and Cys130 are biologically irrelevant to CDO catalysis. Also, neither of the crosslinked residues of CDO, Cys93 and Tyr157, are conserved in the ADO sequences.

We cloned and overexpressed human ADO in *Escherichia coli* to determine if a Cys-Tyr cofactor is present. The as-isolated protein presented two significant bands by SDS-PAGE. However, these two bands appear about 40 kDa away from each other, and are indicative of oligomerization changes rather than crosslinking. Further, the two bands coalesced after incubation with excess dithiothreitol (DTT; Figure 1 A), implicating disulfide bond(s). In contrast, the two

[*] Y. Wang, Dr. W. P. Griffith, Dr. J. Li, Dr. T. Koto, Dr. D. J. Wherritt, E. Fritz, Prof. Dr. A. Liu
Department of Chemistry, The University of Texas at San Antonio
San Antonio, TX (USA)
E-mail: Feradical@utsa.edu
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201803907.

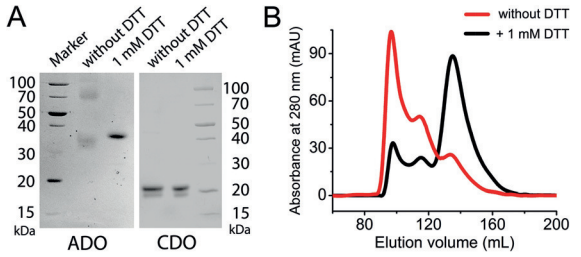


Figure 1. A) SDS PAGE of ADO and CDO incubated with and without excess DTT. B) SEC of ADO eluting from the buffer with and without 1 mM DTT.

bands from CDO are insensitive to DTT because its Cys-Tyr crosslink is irreversible and cannot be reduced by DTT. The higher-molecular-weight band in CDO corresponds to uncrosslinked polypeptide while the faster migrating band is mature CDO with a crosslink.[10] Further processing of ADO with cysteamine and O_2 did not lead to noticeable changes. Similarly, ADO eluted as at least three distinct fractions through size-exclusion chromatography (SEC). However, the inclusion of DTT in the elution buffer substantially reduces other portions to one (Figure 1B). The different ADO conformations observed in solution and denatured forms are most likely a result of the formation of reversible intermolecular disulfide bond(s). Thus, the experiments with low-resolution traditional approaches on crosslinks are inconclusive.

We then attempted to solve the puzzle with high-resolution mass spectrometry. The expected 2 Da difference between the Cys-Tyr crosslinked and uncrosslinked forms of the WT protein is difficult to determine from intact protein MS.[11] Hence, we conducted a rigorous MS/MS analysis on human ADO. The native full-length protein was digested with trypsin and carbamidomethylated with all free cysteines. After examination of all the resulting peptides by high-resolution LC-MS/MS, a short hydrophilic peptide DCHYYR stood out with its early retention time (see Figure S2).

At a +2 charge state, both the crosslinked form and the uncrosslinked form with a carbamidomethyl cysteine were detected (Figure 2A). The crosslinked cysteine covalently bonds with another residue, and thus could not be carbamidomethylated. The two-hydrogen loss (2 Da) in the crosslinked form and one additional carbamidomethylation (57 Da) in the uncrosslinked form resulted in a total of 59 Da difference, indicating a crosslink present in the peptide and that Cys220, which is the only cysteine in this peptide, is involved.

Since two tyrosine residues, Tyr222 and Tyr223, are present in the peptide we conducted further fragmentation

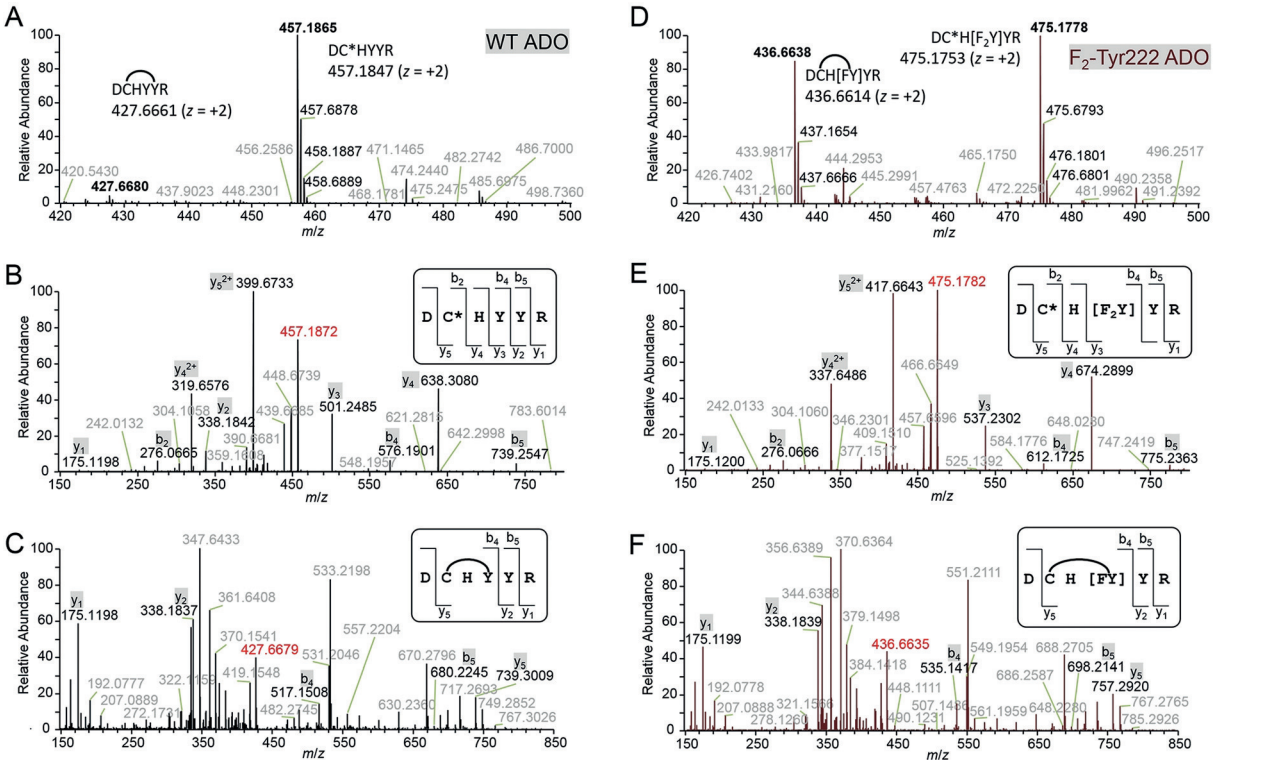


Figure 2. High-resolution mass spectra of WT (A–C) and F_2-Tyr222 ADO (D–F). In WT ADO, both the crosslinked form at *m/z* 427.6680 and the uncrosslinked form at *m/z* 457.1865 were observed by LC-MS/MS (A). The predicted values are 427.6661 and 457.1847. CID spectra of the uncrosslinked (B) and crosslinked (C) forms identified the crosslink between Cys220 and Tyr222. In F_2-Tyr222 ADO, both the crosslinked form at *m/z* 436.6638 and the uncrosslinked form at *m/z* 475.1778 were observed by LC-MS/MS (D). The predicted values are 436.6614 and 475.1753. CID spectra of uncrosslinked (E) and crosslinked (F) forms identified the crosslink was formed by Cys220 and F_2-Tyr222 with only one fluorine. C* represents the carbamidomethylated cysteine. Arc represents crosslinking. The intact peptides ($z = +2$) are highlighted by red; specified fragments are shown in black while other observed values are in gray.

with collision-induced dissociation (CID) to identify the exact crosslink position. The CID spectrum of the uncrosslinked form indicated the presence of each amino acid in this peptide (Figure 2B). But the CID spectrum of the crosslinked form was quite different (Figure 2C), which suggested that the fragmentation ions not containing Cys220, His221, Tyr222 residues (CHY) were conserved in the uncrosslinked peptide (e.g., y1, y2), while the ones containing CHY differed by 59 Da (e.g., y5, b4, b5). Fragments breaking CHY were not observable in the spectrum because of the sidechain crosslink. These data explicitly suggested that Cys220 and Tyr222 are crosslinked. Moreover, the electron-transfer dissociation (ETD) spectrum also confirmed the crosslink is located between these two residues (see Figure S3A).

To facilitate a definitive mass spectrometry analyses and explore the oxidizing power of the oxidant for C–H bond activation during cofactor biogenesis, we incorporated an unnatural tyrosine into ADO through a genetic method. Using a tRNA(MjtRNATyrCUA)/F_2-TyrRS system to target the TAG codon,[12] we generated an expression system that efficiently synthesizes human ADO with a specific substitution of Tyr222 to 3,5-difluoro-tyrosine (F_2-Tyr). The presence of two fluorine atoms will potentially facilitate the mass spectrometry analysis if the same crosslink is formed because the mass difference would be 20 instead of 2 Da after a C–F bond and an S–H bond cleavage, and formation of a new C–S bond. After trypsin digestion on F_2-Tyr222 ADO, we observed a similar elution profile by LC-MS/MS as for WT ADO (see Figure S2B). Interestingly, we still found two forms of the targeting peptide. At $z=+2$, a mono-fluoro-substituted crosslinked form and a di-fluoro-substituted uncrosslinked form with a carbamidomethylated cysteine were observed (Figure 2D). The 77 Da difference was caused by S–H and C–F bond cleavage (20 Da for loss of an H and F) in the crosslinked form and one additional carbamidomethylation (57 Da difference) in the uncrosslinked form. The mass increase of a total of 36 Da in the F_2-Tyr222 ADO uncrosslinked form, compared to WT ADO, was expected for the successful incorporation of F_2-Tyr222, the *ortho* positions of which are substituted by two fluorine atoms instead of two hydrogen atoms.

Next, we performed CID fragmentation on both forms to further confirm the displacement of a fluorine atom upon crosslinking. Compared to the uncrosslinked form of WT ADO (Figure 2B), the uncrosslinked form of F_2-Tyr222 ADO (Figure 2E) has an overall 36 Da increase only in fragments containing F_2-Tyr222 (e.g., y3, y4, y5, b4, b5) and no increase in fragments without it (e.g., y1, b2). As for the crosslinked form, however, F_2-Tyr222 ADO (Figure 2F) shows only an 18 Da increase on the fragments containing the crosslink (e.g., y5, b4, b5) when compared to the WT (Figure 2C). The same results were obtained in the ETD spectrum (see Figure S3B). Thus, there is only a single fluorine in the crosslinked peptide. Together with the WT results, the presence of crosslink between Cys220 and Tyr222 is unambiguously established.

To reconcile the mass spectrometry results, we employed ^{19}F NMR spectroscopy to attempt to detect the fate of the fluorine during cofactor biogenesis in F_2-Tyr222 ADO. Under anaerobic conditions, concentrated F_2-Tyr222 ADO was mixed with the substrate cysteamine in a sealed NMR tube. The resulting ^{19}F NMR spectrum was completely silent (Figure 3). The absence of a F_2-Tyr222 signal may result from the coupling with the adjacent paramagnetic iron center.

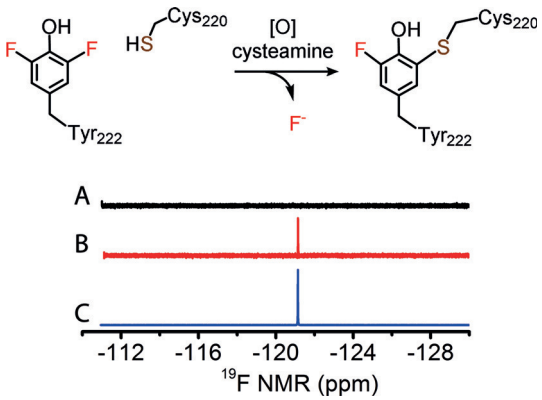


Figure 3. ^{19}F NMR spectra of the F_2-Tyr222 enzyme–substrate complex under anaerobic conditions (A), with exposure to air for an overnight reaction (B), and with a spike of aqueous KF added to the reaction mixture (C).

Then the NMR tube was opened and exposed to air overnight for an in-tube reaction. The oxidized F_2-Tyr222 ADO gave rise to a new signal at $\delta=-121.14$ ppm. Spiking in aqueous potassium fluoride overlaid with the observed new signal, which confirms the fluorine departed in the fluoride form during autocatalytic Cys-Tyr cofactor formation, which was promoted in the presence of the enzyme substrates (i.e., cysteamine and O_2).

The C–F bond is one of the most durable single bonds in organic chemistry. The oxidative cleavage of one of the C–F bonds in F_2-Tyr222 ADO is an intriguing observation. To estimate the bond strength of the C–F bond with an aromatic carbon center, we performed density functional theory calculations on L-Tyr and F_2-Tyr. The results show that the dissociation energy of the C–F bond (132.66 kcal mol^{-1}) is 10.12 kcal mol^{-1} greater than that of the corresponding C–H bond (122.54 kcal mol^{-1}; see Table S1), which is more significant than a phosphoanhydride bond in ATP (7.3 kcal mol^{-1}), yet the nonheme iron-bound oxidant possesses sufficient power to cleave a C–F bond.

We obtained a predicted model structure by using the Phyre2 server[13] to estimate the locations and relative distances of the cysteine and tyrosine residues. The model suggests that Tyr87, Tyr222, Tyr223, and Cys220 are within the second coordination sphere of the iron center while others such as Tyr91 and Cys183, conserved in CDO, are likely located on the protein surface and remote from the iron center (Figure 4A). Notably, the side chains of Cys220 and Tyr222 are close to each other (ca. 3 Å) but located on a different side of the iron as compared to CDO (Figure 4B). Moreover, our multiple sequence alignment revealed that Cys220 and Tyr222 are both conserved in the eukaryotic ADOs and a new Cys-Tyr motif C-X-Y-Y(F) is identified (see Figure S1B).

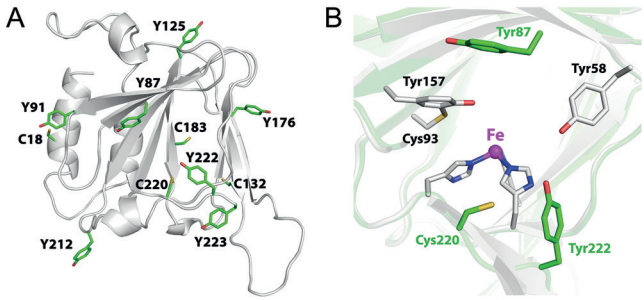


Figure 4. ADO structural model. A) The cysteine and tyrosine residues are highlighted. Tyr87, Tyr222, Tyr223, and Cys220 are within the second coordination sphere of the iron center while others are likely located in protein surface and remote from the iron center. B) Superposition of the active site of ADO model (green) and CDO crystal structure (PDB ID: 2IC1, gray).

We produced site-directed mutates of Tyr87, Tyr222, and Tyr223, altering them to alanine. The resulting variants were subjected to an activity assay using an oxygen electrode. We quantitated the oxygen consumption and plotted Michaelis–Menten curves (see Figure S4). The catalytic activity was most significantly reduced in Y222A ADO (ca. 23% of WT), with a modest reduction in Y87A (ca. 51% of WT), and no significant reduction in the activity of Y223A was observed (see Table S2). To exclude the possibility of decreased activity resulting from backbone distortions, we constructed a Y222F variant and again obtained similar assay results that showed that mutation of Tyr222 led to a significant reduction of the ADO activity. Alternatively, F_2-Tyr222 ADO only showed a modest decrease in activity (ca. 66% of WT). Thus, the structural modeling and site-directed mutagenesis analysis further corroborate the presence of Cys220-Tyr222 cofactor in human ADO and its enhancement to catalysis of the cysteamine deoxygenation is proven.

In the past decade, thiol dioxygenases, including ADO and CDO, have attracted a great deal of attention because of the potential of novel sulfur-based chemistry. The structure and function study of ADO has been hindered by ambiguity on its catalytic cofactor in the mature protein, and the enzymatic mechanism remains unknown even after nearly half a century of the discovery of its enzymatic activity. The ADO cofactor is less accessible relative to that of CDO because it is not formed from two remote residues and little impact made to the overall structure, so that there is seemingly only one form of the protein at the reduced state. Of note, the protein-derived Cys-Tyr redox cofactor is present in a growing number of iron- and copper-containing proteins. Its presence either enables or significantly enhances the capacity of the host metalloprotein to mediate a specific redox process, which may be achieved through correctly positioning substrate(s) or stabilizing the catalytic intermediates. It was reported that there are nearly three hundred candidate proteins identified for probable Cys-Tyr crosslinking based on an examination of published protein structures,[11] whereas few have been validated to possess such a crosslinking. The major hurdle for such a limited number of the reported protein-derived Cys-Tyr cofactors is the dependence on crystallography as for identification, and it is not always accessible to a specific system. A recent study has provided an alternative method that utilizes native fluorescence, however, there is limited choice for the studied proteins since it requires small size and low tryptophan composition.[14] Herein, a high-resolution mass-spectrometry-centered method coupled with the genetic incorporation of an unnatural amino acid is described,[12,15] and may be potentially developed into an efficient method to detect the presence of a Cys-Tyr cofactor.

In summary, the Cys-Tyr cofactor is firmly established as part of the catalytic machinery in human ADO by high-resolution mass spectrometry, mutagenesis analysis, activity assays, and ^{19}F NMR data. ADO is the second human enzyme shown to contain a Cys-Tyr cofactor, and unlike the first one found in CDO, it has its own crosslink motif. The identification of the C-X-Y-Y(F) motif provides one more template for Cys-Tyr crosslink discovery in thiol dioxygenases. Additionally, we determined the metal-bound oxidant in ADO is potent enough to activate either an aromatic C–H or a more durable C–F bond when a genetically substituted fluorotyrosine probe is incorporated into the protein. These findings will facilitate the structure and mechanism studies of ADO.

Acknowledgements

This work was supported in whole or part by the National Institutes of Health (NIH) grants GM107529 and GM108988, the National Science Foundation grant CHE-1623856, and the Lutcher Brown Distinguished Chair Endowment fund (to A.L.). The mass spectrometry facility was sponsored by National Institutes of Health Grant G12MD007591. The NMR spectrometer used is a shared instrument sponsored by the National Science Foundation (NSF) under award no. 1625963. We thank Dr. Jiangyun Wang for the generous gift of pEVOL-F_2-TyrRS plasmid and Ian Davis for constructive discussions.

Conflict of interest

The authors declare no conflict of interest.

Keywords: amino acids · C–H activation · dioxygenases · fluorine · proteins

How to cite: *Angew. Chem. Int. Ed.* **2018**, *57*, 8149–8153
Angew. Chem. **2018**, *130*, 8281–8285

[1] D. Cavallini, R. Scandurra, C. Demarco, *J. Biol. Chem.* **1963**, *238*, 2999–3005.
[2] a) D. Cavallini, R. Scandurra, S. Duprè, *Methods Enzymol.* **1971**, *17 Part B*, 479–483; b) J. E. Dominy, Jr., C. R. Simmons, L. L. Hirschberger, J. Hwang, R. M. Coloso, M. H. Stipanuk, *J. Biol. Chem.* **2007**, *282*, 25189–25198.
[3] a) M. Bousquet, C. Gibrat, M. Ouellet, C. Rouillard, F. Calon, F. Cicchetti, *J. Neurochem.* **2010**, *114*, 1651–1658; b) G. Cisbani, J. Drouin-Ouellet, C. Gibrat, M. Saint-Pierre, M. Lagace, S. Badrinarayanan, M. H. Lavallee-Bourget, J. Charest, A. Chab-

rat, L. Boivin, M. Lebel, M. Bousquet, M. Levesque, F. Cicchetti, *Neurobiol. Dis.* **2015**, *82*, 430 – 444; c) K. Hara, M. Nakamura, Y. Haranishi, T. Terada, K. Kataoka, T. Sata, *Amino Acids* **2012**, *43*, 397 – 404; d) R. M. Coloso, L. L. Hirschberger, J. E. Dominy, J. I. Lee, M. H. Stipanuk, *Adv. Exp. Med. Biol.* **2006**, *583*, 25 – 36.

[4] I. Ueki, M. H. Stipanuk, *J. Nutr.* **2009**, *139*, 207 – 214.

[5] a) T. Nishimura, M. Duereh, Y. Sugita, Y. Yoshida, K. Higuchi, M. Tomi, E. Nakashima, *Placenta* **2015**, *36*, 693 – 698; b) E. T. Donnelly, N. McClure, S. E. Lewis, *Mutagenesis* **2000**, *15*, 61 – 68.

[6] a) J. G. McCoy, L. J. Bailey, E. Bitto, C. A. Bingman, D. J. Aceti, B. G. Fox, G. N. Phillips, Jr., *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3084 – 3089; b) C. R. Simmons, Q. Liu, Q. Q. Huang, Q. Hao, T. P. Begley, P. A. Karplus, M. H. Stipanuk, *J. Biol. Chem.* **2006**, *281*, 18723 – 18733; c) S. Ye, X. a. Wu, L. Wei, D. Tang, P. Sun, M. Bartlam, Z. Rao, *J. Biol. Chem.* **2007**, *282*, 3391 – 3402.

[7] J. E. Dominy, Jr., J. Hwang, S. Guo, L. L. Hirschberger, S. Zhang, M. H. Stipanuk, *J. Biol. Chem.* **2008**, *283*, 12188 – 12201.

[8] M. H. Stipanuk, C. R. Simmons, P. A. Karplus, J. E. Dominy, Jr., *Amino Acids* **2011**, *41*, 91 – 102.

[9] P. Rice, I. Longden, A. Bleasby, *Trends Genet.* **2000**, *16*, 276 – 277.

[10] M. H. Stipanuk, M. Londono, L. L. Hirschberger, C. Hickey, D. J. Thiel, L. Wang, *Amino Acids* **2004**, *26*, 99 – 106.

[11] R. J. Martinie, P. I. Godakumbura, E. G. Porter, A. Divakaran, B. J. Burkhart, J. T. Wertz, D. E. Benson, *Metallomics* **2012**, *4*, 1037 – 1042.

[12] F. Li, P. Shi, J. Li, F. Yang, T. Wang, W. Zhang, F. Gao, W. Ding, D. Li, J. Li, Y. Xiong, J. Sun, W. Gong, C. Tian, J. Wang, *Angew. Chem. Int. Ed.* **2013**, *52*, 3958 – 3962; *Angew. Chem.* **2013**, *125*, 4050 – 4054.

[13] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **2015**, *10*, 845 – 858.

[14] S. E. Hromada, A. M. Hilbrands, E. M. Wolf, J. L. Ross, T. R. Hegg, A. G. Roth, M. T. Hollowell, C. E. Anderson, D. E. Benson, *J. Inorg. Biochem.* **2017**, *176*, 168 – 174.

[15] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498 – 500.

Manuscript received: April 2, 2018
Accepted manuscript online: May 11, 2018
Version of record online: June 5, 2018

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Cofactor Biogenesis in Cysteamine Dioxygenase: C–F Bond Cleavage with Genetically Incorporated Unnatural Tyrosine

*Yifan Wang, Wendell P. Griffith, Jiasong Li, Teruaki Koto, Daniel J. Wherritt, Elizabeth Fritz, and Aimin Liu**

anie_201803907_sm_miscellaneous_information.pdf

Author Contributions

Y.W. Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Visualization: Lead; Writing—original draft: Lead; Writing—review & editing: Lead
W.G. Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Validation: Lead; Writing—original draft: Equal
J.L. Data curation: Equal; Methodology: Lead; Resources: Lead; Validation: Supporting
T.K. Data curation: Lead; Formal analysis: Lead; Validation: Lead
D.W. Data curation: Lead; Formal analysis: Lead
E.F. Data curation: Supporting
A.L. Conceptualization: Lead; Funding acquisition: Lead; Project administration: Lead; Supervision: Lead; Validation: Lead; Writing—original draft: Lead; Writing—review & editing: Lead.

Table of Contents

Experimental Procedures

Synthesis of 3,5-Difluorotyrosine

To synthesize F_2-Tyr, we transformed 2,6-difluorophenol to F_2-Tyr by using Citrobacter Freundii (ATCC8090) tyrosine phenol lyase (TPL) following established methods.[1] Detailed synthetic method and F_2-Tyr purification were identical to previous publication.[2] The F_2-Tyr prepared was used in the following experiment for genetic incorporation into human ADO.

Cloning, Expression, Purification and Analysis of Recombinant Proteins

Expression vector pET28a-human ADO was purchased from DNASU Plasmid Repository and transformed into Rosetta2™ (DE3) competent cells (Merck). Expression and purification of His_6-tagged human ADO was performed as follows: Kanamycin (50 μg/mL) and chloramphenicol (30 μg/mL) were used as antibiotics for selection of recombinant strains. the culture was grown at 37 °C in Luria Bertani (LB) media in a baffled flask at 220 rpm till $OD_{600\ nm}$ reached 0.8 AU. Isopropyl-β-thiogalactoside (IPTG) and ferrous ammonium sulfate were added to a final concentration of 600 and 50 μM before the temperature was lowered to 28 °C. After 12 h of further incubation, the cells were harvested and resuspended in lysis buffer (50 mM Tris-HCl, 200 mM NaCl, pH 8.0) and then disrupted by a Microfluidizer LM20 cell disruptor, and the supernatant was recovered after centrifugation (4 °C, 25,000 g for 40 min). ADO was separated using nickel-nitrilotriacetic acid resin. After being washed with washing buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, pH 8.0), protein was eluted with elution buffer (50 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 8.0). Purified protein-containing fractions were further purified by gel filtration column (Superdex 75, 26/65) in storage buffer (50 mM Tris-HCl, 200 mM NaCl, 5% glycerol, pH 8.0) for further use. Protein concentration was determined by UV-vis absorbance at 280 nm ($\varepsilon_{280\ nm}$ = 24,785 $cm^{-1}M^{-1}$).

The sets of mutagenic primers were synthetized by Integrated DNA Technologies. The sequence of the forward primers are listed and its reverse pairs are reverse complement of the forward primers:

5'–GCCTCCGGTTACCGCTATGCACATTTATGAAAC–3' for Y87A
5'–GTCGTGATTGTCATTATGCTCGTGTTCTGGAACCGG–3' for Y223A
5'–GTCGTGATTGTCATTTTTATCGTGTTCTGGAAC–3' for Y222F
5'–GTCGTGATTGTCATGCTTATCGTGTTCTGGAAC–3' for Y222A
5'–GTCGTGATTGTCATTAGTATCGTGTTCTGGAAC–3' for Y222F_2-Tyr

Reagents were purchased from Sigma-Aldrich, New England Biolabs and Thermo, Inc. DNA manipulations in Escherichia coli were carried out according to standard procedures. The expression and purification for all mutants are same as the WT ADO.
For the expression of F_2-Tyr222 ADO, pEVOL-F_2-TyrRS was co-transformed with pET28a-human ADO222TAG into BL21 (DE3). The cells were induced with 0.6 mM IPTG and 0.02% L-arabinose at OD600 nm of 0.8 in the presence of 0.5 mM F_2-Tyr. After growing for 12 h at 28 °C, the F_2-Tyr222 ADO were purified using the protocol described above.

Activity Assay

ADO activity assay was done using oxygen electrode (Oxygraph, Hansatech Instruments). Enzymes were freshly iron reconstituted before assay use. A standard condition derived from previous method was conducted with 0 - 3 μg enzyme, 50 mM Tris borate buffer and varying concentrations of cysteamine (0 – 40 mM) in a total volume of 1 mL (pH8.0, 25 °C).[3] The reaction was initiated by addition of cysteamine into the electrode chamber with constant stirring. The net oxygen consumption was obtained by subtracting initial oxygen consumption rate in the presence of enzyme by blank, the one in the absence of enzyme. Curves of net oxygen consumption over enzyme concentration versus starting cysteamine concentration were fitted with the Michaelis-Menten equation in OriginPro (OriginLab, Northampton, MA) to determine k_{cat} and K_m values.

Mass Spectrometry

WT and F_2-Tyr222 ADO proteins were denatured by incubation at 55 °C for 1 h in a solution of 6 M guanidine HCl containing 10 mM DTT. After cooling to room temperature, iodoacetamide was added to a concentration of approximately 40 mM and incubated in the dark at room temperature for 4 h. Protein samples were exhaustively desalted by filtration with 50 mM ammonium acetate through centrifugal filters with 10 kDa membrane cut-off. Proteomics-grade trypsin was added to the desalted carbamidomethylated proteins (1:100 enzyme:substrate ratio) and the mixture incubated at 37 °C overnight. The trypsin-digested proteins were analyzed on a nano

reversed-phase liquid chromatography (LC) electrospray ionization tandem mass spectrometry (MS) system that consisted of an UltiMate 3000 Nano LC System and an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher, San Jose, CA). Chromatography was performed using a home-made 20 cm × 75 μm ID column packed with XBridgeTM BEH C18 beads (2.5 μm, 130 Å). Solvent A (0.1% aqueous formic acid) and B (0.1% formic acid in acetonitrile) were used to establish the 85-min gradient, comprised of 60 min of 5 - 45% B, then 15 min of 45-90% B, and finally maintained at 90% B for 10 min, with the flow rate at 200 nL/min. The LTQ-Orbitrap Elite mass spectrometer was operated in positive ionization mode with a 2.6 kV applied spray voltage. The temperature of the ion transfer capillary was 300 °C. One microscan was set for each MS and MS/MS scan. A full scan MS acquired in the range 380 ≤ m/z ≤ 2000 was followed by 10 data dependent MS/MS events on the 10 most intense ions. The mass resolution was set at 60000 for full MS. The dynamic exclusion function was set as follows: repeat count, 1; repeat duration, 30 s; exclusion duration, 30 s. Collision-induced dissociation (CID) was performed at a normalized collision energy of 35% and the activation time was set at 0.1 ms.

For offline electron transfer dissociation (ETD) fragmentation analysis of LC-fractionated peptides, sample solutions were introduced into the nanospray source *via* syringe pump at a flow rate of 20 nL/min. Instrument parameters were maintained at the following optimal values for each experiment unless otherwise indicated: source voltage, 2.2 kV; vaporizer temperature, 180 °C; capillary temperature, 200 °C. The Orbitrap was set to 60K resolution for all data collection. The calibration was performed externally using Pierce® ESI Positive Ion Calibration Solution (Thermo Scientific); and mass spectra recorded in the range 50 ≤ *m*/*z* ≤ 1500. Mass spectra were an average of 1 minute of scans with an automatic gain control target value of 1×10^6 ions or maximum injection time of 100 ms. Specific ions of interest were isolated and fragmented by ETD using fluoranthene as reagent and with 100 ms activation time. Data were processed using Qual Browser Xcalibur™ 2.2 (Thermo Fisher Scientific Inc.).

Nuclear Magnetic Resonance Spectroscopy

^{19}F NMR spectra were recorded on an Bruker (Billerica, MA) Avance III HD 500 MHz spectrometer equipped with a 5-mm Cryoprobe Prodigy at 300 K. Spectra were recorded in 90/10 buffer/D_2O and referenced to internal trifluoroacetic acid (-76.5 ppm). One-dimension ^{19}F spectra (zgflqn) were recorded with 5 s relaxation delay, 64 k data points, and multiplied with an exponential function for a line-broadening of 5 Hz before Fourier transformation. All NMR data were processed using MestReNova NMR v11.0.3 software. 400 μM F_2-Tyr222 ADO was first incubated with 30 mM cysteamine in 100 mM Tris buffer, pH 8.0 anaerobically and took 3000 scans. Then it was exposed to air for overnight to obtain more crosslinked ADO before another 3000 scans.

ADO Model Building and Sequence Alignment

Human ADO primary sequence was submitted to Phyre2 protein fold recognition sever[4] with normal modeling mode. The final model was built mainly on type I cysteine dioxygenase template. 155 residues out of 270 amino acids (57% sequence coverage) have been modeled with 99.7% confidence by the cysteine dioxygenase template. In the predicted human ADO model, four out of six cysteine residues and seven out of eight tyrosine residues were presented. Only Cys239, Cys258, Tyr263 at disordered C-terminal end were not predicted.

The pairwise sequence alignment of human CDO (BAA12873) and human ADO (NP_116193) were done using EMBOSS Needle.[5] The sequences of eukaryotic and putative ADOs are obtained from NCBI non-redundant protein sequence database, and aligned by Clustal Omega multiple sequence alignment.[6] The final alignment figure was generated by ESPript.[7]

Calculation of Dissociation Energy of C-H and C-F Bonds by Density Function Theory

All calculations were implemented with the ORCA (4.0.1) quantum chemistry program package.[8] Full geometry optimizations were performed for all molecules using B3LYP hybrid functional in combination with RIJCOSX approximation with the 6-311+G(d,p) basis set for all atoms, in which tight SCF convergence criteria were used, for obtaining minimum energy structures with zero imaginary frequencies. The solvent effect of aqueous media was estimated using the conductor-like polarizable continuum model (CPCM) in combination with the conductor-like screening model (COSMO)[9] epsilon function (dielectric constant ε = 80.4 for water). Thermal corrections to enthalpies and Gibbs free energies determined at 298.15 K using the B3LYP/6-311+G(d,p) level of theory were also applied to the total energies. Table S1 shows theoretical bond dissociation energy, which was calculated by difference in Gibbs free enthalpy on the assumption that carbon-hydrogen bond and carbon-fluorine bonds are homolytically cleaved in aqueous solution. We calculated also for fluorobenzene and unfluorinated benzene for comparison with reported data.[10]

Supporting Table and Figures

Table S1. Calculation of Theoretical Bond Dissociation Energy. Homolytic cleavage of carbon-hydrogen/carbon-fluorine bond calculated by difference in Gibbs free enthalpy at 298.15 K at the level of B3LYP/6-311+G(d,p), and comparison with reported experimental data.

	Calcd. DH_{298} (kcal/mol)	Expl. DH_{298} (kcal/mol)
C-H bond in Tyr	122.54	-
C-F bond in F_2-Tyr	132.66	-
C-H bond in benzene	120.43	112.9 ± 0.6[a]
C-F bond in fluorobenzene	134.78	127.2 ± 0.7[a]

[a] Reported experimental data.[10]

Table S2. Site-directed Mutagenesis Analysis of ADO. The experiment was done using oxygen electrode at room temperature with 0 - 3 µg enzyme, 50 mM Tris borate buffer and varying concentrations of cysteamine (0 – 40 mM) in a total volume of 1 mL**.**

	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
WT	1.22 ± 0.08	3.4 ± 0.3	0.36 ± 0.06
Y87A	0.62 ± 0.03	3.6 ± 0.4	0.18 ± 0.03
Y223A	1.21 ± 0.14	5.6 ± 1.8	0.25 ± 0.11
Y222A	0.28 ± 0.02	5.3 ± 0.7	0.05 ± 0.01
Y222F	0.51 ± 0.05	14.4 ± 3.0	0.04 ± 0.01
F_2-Try222	0.80 ± 0.03	5.4 ± 0.3	0.15 ± 0.01

Figure S1. Sequence Alignment. (A) Pairwise sequence alignment of human ADO and CDO. The 3-His metal binding motif is conserved; only one pair of tyrosine (Tyr40 in CDO, and Tyr91 in ADO) and one pair of cysteine (Cys130 in CDO, and Cys183 in ADO) are conserved. α-helices and β-strands are predicted using unknown CDO structure (PDB ID 2IC1). (B) Multiple sequence alignment of eukaryotic and putative ADOs. The included species and NCBI protein accession numbers are (1) *Homo sapiens*, NP_116193; (2) *Pan troglodytes*, JAA35409; (3) *Mus musculus*, NP_001005419; (4) *Lepidothrix coronate*, XP_017681447; (5) *Xenopus laevis*, NP_001088077; (6) *Nothobranchius rachovii*, SBR92053; (7) *Spodoptera litura*, XP_022827635; (8) *Leishmania major*, CAJ02170; (9) *Strongylocentrotus purpuratus*; XP_001183066; (10) *Gossypium arboretula*, KHG13152; (11) *Arabidopsis thaliana*, NP_191426; (12) *Medicago truncatula*, ABD28522; (13) *Phytophthora nicotianae*, KUF87667; (14) *Chrysochromulina*, KOO33837. Highly and strictly conserved residues are boxed, and highlighted in black background, respectively.

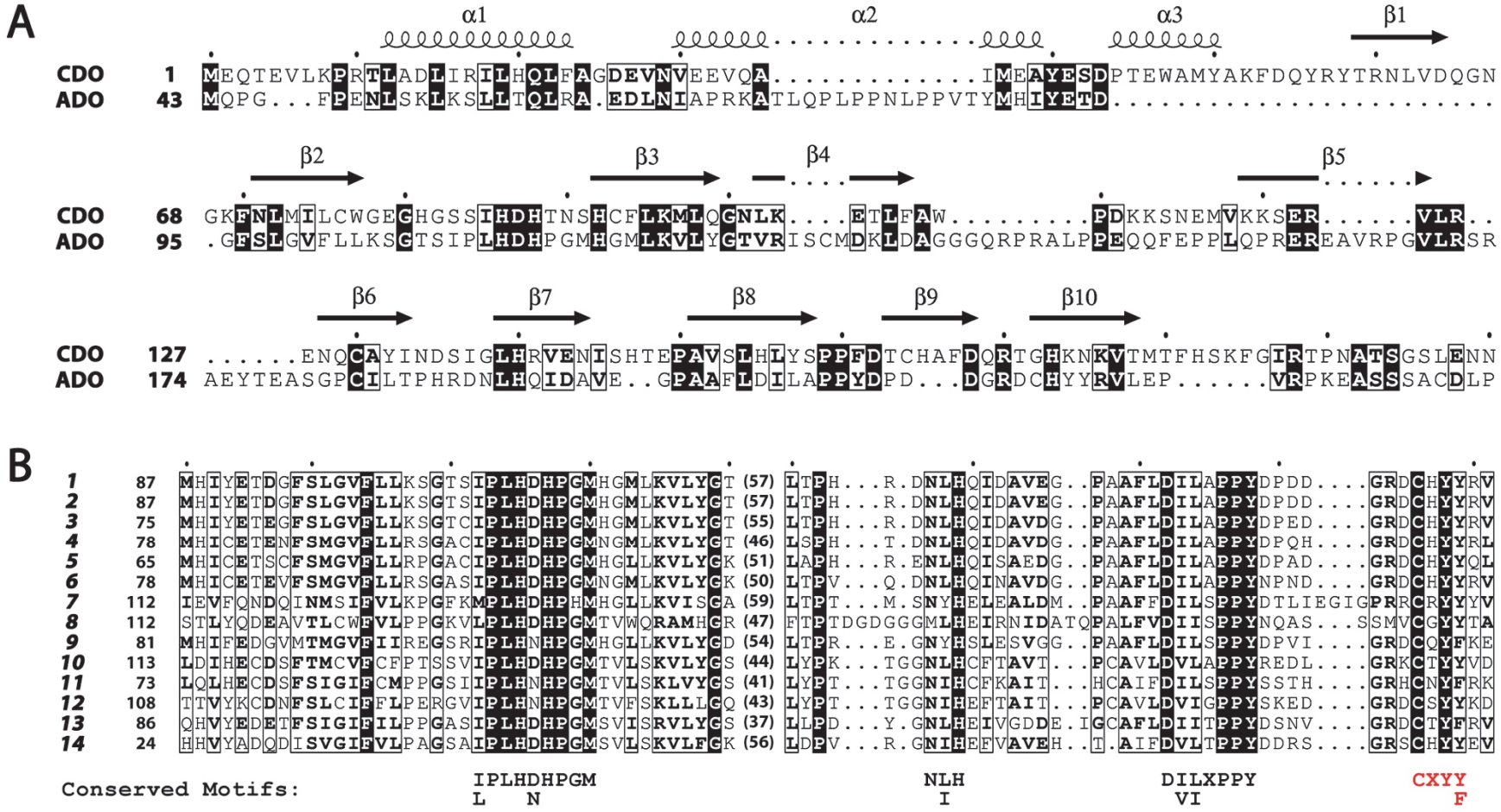


Figure S2. LC-MS of WT and F_2-Try222 ADO. WT (A) and F_2-Try222 ADO (B) peptides eluted form LC-MS after trypsin digestion. Gray colored text represents elution time (min); Black colored text represents *m*/*z* values. The Cys220 and Tyr222 containing peptide of WT ADO DCHYYR eluted at 40-44 min while the F_2-Tyr222 ADO peptide DCH[F_2Y]YR eluted at 49-53 min.

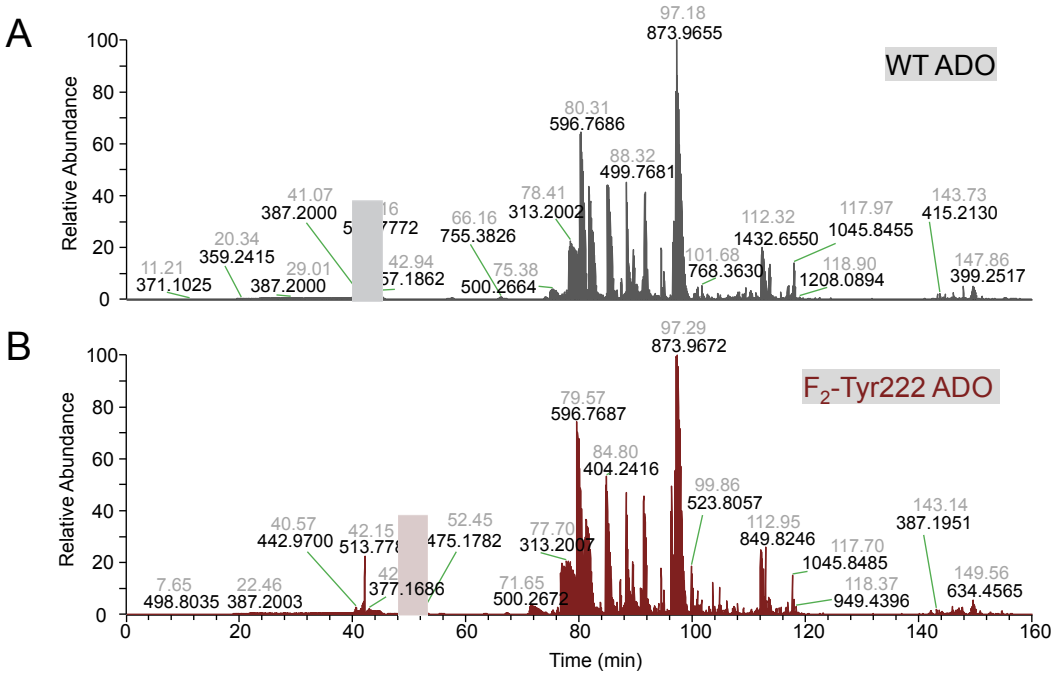


Figure S3. ETD Fragmentation Spectra of LC-fractionated Peptides. (A) Crosslinked peptide of WT ADO and (B) crosslinked peptide of F2-Try222 ADO. The arc represents crosslinking. FY represents a mono-fluoro substituted Tyr222. The intact peptides at +2 charge state are highlighted by red and signature fragmentation ions are colored by black. Other observed *m*/*z* values are colored by gray.

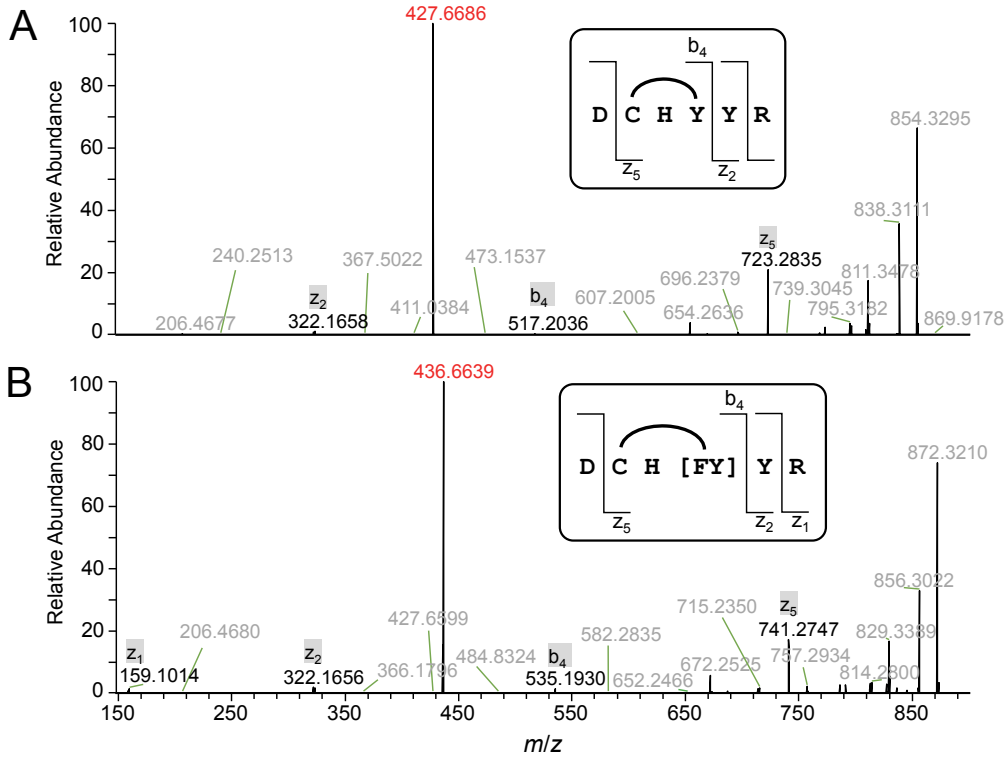
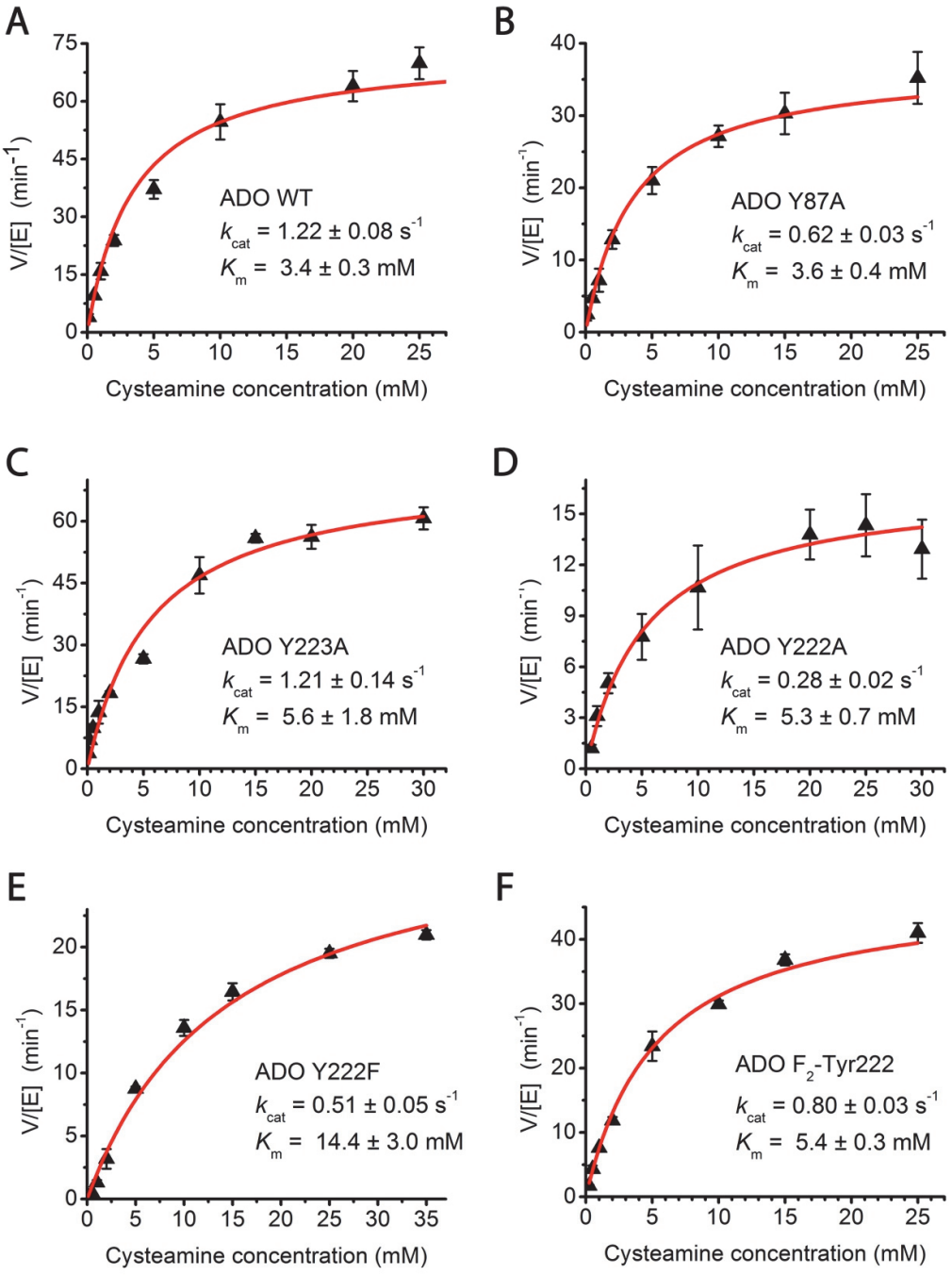


Figure S4. Michaelis-Menten Plots for WT ADO and ADO Mutants. WT ADO (A) showed activity of k_{cat} = 1.22 ± 0.08 s^{-1}, K_m = 3.4 ± 0.3 mM. Y87A (B) and Y223A (C) did not significantly affect the activity while Y222A (D) and Y222F (E) had much slower activity compared with WT. F_2-Try222 ADO (F) was catalytic active but had a slower catalytic rate.



References

[1] M. R. Seyedsayamdost, C. S. Yee, J. Stubbe, *Nat. Protoc.* **2007**, *2*, 1225-1235.
[2] F. Li, P. Shi, J. Li, F. Yang, T. Wang, W. Zhang, F. Gao, W. Ding, D. Li, J. Li, Y. Xiong, J. Sun, W. Gong, C. Tian, J. Wang, *Angew. Chem. Int. Ed.* **2013**, *52*, 3958-3962.
[3] J. E. Dominy, Jr., C. R. Simmons, L. L. Hirschberger, J. Hwang, R. M. Coloso, M. H. Stipanuk, *J. Biol. Chem.* **2007**, *282*, 25189-25198.
[4] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat Protoc* **2015**, *10*, 845-858.
[5] P. Rice, I. Longden, A. Bleasby, *Trends. Genet.* **2000**, *16*, 276-277.
[6] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Z. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson, D. G. Higgins, *Mol. Syst. Biol.* **2011**, *7*.
[7] X. Robert, P. Gouet, *Nucleic Acids Res.* **2014**, *42*, W320-324.
[8] F. Neese, *Wiley Interdiscip. Rev.* **2012**, *2*, 73-78.
[9] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, F. Neese, *J. Phys. Chem. A* **2006**, *110*, 2235-2245.
[10] S. J. Blanksby, G. B. Ellison, *Acc. Chem. Res.* **2003**, *36*, 255-263.

Author Contributions

Y. W. did all the sample preparation (cloning, protein expression, and purification). W. P. G. and Y. W. conducted the mass spectrometry analyses. Y. W. and J. L synthesized F_2-Tyr and incorporated it into human ADO by a genetic method. D.J.W. performed the ^{19}F NMR analysis. Y. W. and E. F. conducted side-directed mutagenesis analyses and kinetic assays. T. K. performed computational estimations on bond dissociation energies. A.L. conceived the research and wrote the manuscript together with Y.W. and W. P. G. All authors approved the final submitted manuscript.