

MIT Open Access Articles

Collagen Prolyl 4-Hydroxylase as a Therapeutic Target

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Vasta, James D. et al. "Collagen Prolyl 4-Hydroxylase as a Therapeutic Target." *Journal of Medicinal Chemistry* 61, 23 (December 13, 2018): 10403-10411 © 2018 American Chemical Society

As Published: <http://dx.doi.org/10.1021/acs.jmedchem.8b00822>

Publisher: American Chemical Society (ACS)

Persistent URL: <https://hdl.handle.net/1721.1/123552>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.





Published in final edited form as:

J Med Chem. 2018 December 13; 61(23): 10403–10411. doi:10.1021/acs.jmedchem.8b00822.

Collagen Prolyl 4-Hydroxylase as a Therapeutic Target

James D. Vasta[†] and Ronald T. Raines^{†,‡,#,*}

[†]Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

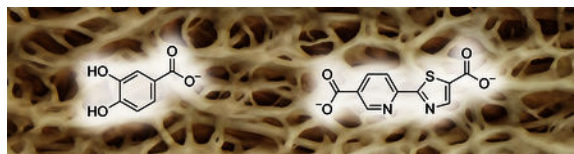
[‡]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[#]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Abstract

Collagen is the dominant protein of the extracellular matrix. Its distinguishing feature is a three-stranded helix of great tensile strength. (2*S*,4*R*)-4-Hydroxyproline residues are essential for the stability of this triple helix. These residues arise from the post-translational modification of (2*S*)-proline residues by collagen prolyl 4-hydroxylases (CP4Hs), which are members of the Fe(II)- and α -ketoglutarate (AKG)-dependent dioxygenase family. Here, we provide a framework for the inhibition of CP4Hs as the basis for treating fibrotic diseases and cancer metastasis. We begin with a summary of the structure and enzymatic reaction mechanism of CP4Hs. Then, we review the metal ions, metal chelators, mimetics of AKG and collagen strands, and natural products that are known to inhibit CP4Hs. Our focus is on inhibitors with potential utility in the clinic. We conclude with a prospectus for more effective inhibitors.

Abstract



Comprising a third of the proteins in the human body and three-quarters of the dry weight of human skin, collagen is the most abundant protein in humans and other animals.¹ Different types of collagen play myriad roles in biology. Collagen is, however, predominantly a structural protein, serving as the principle component of connective tissue, bone, and the extracellular matrix (ECM).

The distinguishing feature of collagen is its triple helix (Figure 1), an elegant structural motif in which three polypeptide chains are interwoven to form a right-handed superhelical structure of great tensile strength (Figures 1A and 1B).¹ The individual polypeptide chains

*Corresponding Author: rtraines@mit.edu.

The authors declare no competing financial interest.

that form the triple helix are composed of numerous triplet repeats of the sequence Xaa-Yaa-Gly, where Xaa is often (2*S*)-proline (Pro) and Yaa is often (2*S*,4*R*)-4-hydroxyproline (Hyp), two amino acids discovered by Emil Fischer.^{2,3} The presence of Hyp in collagen is due to the post-translational hydroxylation of Pro residues in the Yaa position by collagen prolyl 4-hydroxylases (CP4Hs).^{4,5} This hydroxylation enhances greatly the conformational stability of the triple helix⁶ by imposing stereoelectronic effects that preorganize individual strands (Figure 1C).⁷ Thus, the activity of CP4H is critical to the structure of the most prevalent protein in the human body.

Collagen Biosynthesis

As collagens are secretory proteins, their biosynthesis begins in the rough endoplasmic reticulum (ER) (Figure 2).⁸ Mature mRNAs are translated into individual procollagen polypeptides (also called preprocollagens) that undergo processing in the ER lumen. In addition to the excision of the signal peptide, the processing includes extensive post-translational modifications, many of which are essential for various aspects of triple-helix or supramolecular fibril formation. Modified procollagen strands fold into triple helices in a process facilitated by globular N- and C-terminal propeptide domains present in each procollagen strand. Typically, the C-terminal domains from three procollagen strands form a complex that is stabilized by interdomain interactions that enforce a particular homotrimeric or heterotrimeric composition.⁹ This complex also establishes the register for the pendant (Xaa-Yaa-Gly)_{*n*} domains, which “zip up” from C-terminus to N-terminus to form a procollagen triple helix. Triple-helix assembly appears to benefit from general chaperone proteins in the secretory pathway as well as from more specific proteins.^{10,11} To date, 48 proteins have been identified as participants in collagen proteostasis, including Hsp47, cyclophilin B, FKBP10, and enzymes involved in post-translational modifications, including CP4Hs.¹² During assembly, the N- and C-terminal propeptide domains serve to enhance solubility and prevent premature aggregation.⁸ Procollagen triple helices, which are substantially longer (0.3 μm) than typical secretory proteins,¹ are transported from the endoplasmic reticulum to the Golgi in enlarged COPII-coated vesicles.¹³ Subsequent steps in processing and secretion await further clarification and could vary depending on collagen type. Ultimately, the N- and C-terminal propeptide domains are excised by specific proteases, and the resulting collagen triple helices aggregate into fibrils that associate with growing fibers in the ECM.

CP4H as a Drug Target

Abnormalities in the structure and metabolism of collagen are associated with a variety of heritable diseases. For example, the rare genetic disorders osteogenesis imperfecta (OI) and Ehlers–Danlos syndrome (EDS) involve mutations that decrease the production or structural stability of fibrillar collagens in a tissue.¹ Diseases within the OI and EDS spectra are most commonly associated with single-base substitutions that lead to the substitution of an obligate glycine with another residue.

As catalysts of a key step in collagen biosynthesis, CP4H has been a drug target since the discovery of the enzyme itself.¹⁴ Of primary interest are diseases in which normal collagen

is overproduced or produced in a manner that exacerbates a pathological process. The disruption of collagen biosynthesis could be especially beneficial in the treatment of such fibrotic diseases.^{15–19} Recent links between CP4Hs and cancer metastasis suggest new opportunities. For example, the remodeling of collagen in the ECM near tumors has been linked to cancer metastasis.²⁰ Moreover, high levels of two CP4H isoforms correlate well with cancer patient mortality, and an inhibitor of CP4H activity decreases cancer metastasis in mice.²¹ The mechanism-of-action could entail normalization of the vasculature and decompression of blood vessels.²²

CP4Hs are susceptible to inhibition by a variety of chemical strategies.^{23,24} Here, we begin by reviewing the structure and enzymatic reaction mechanism of CP4H. We then discuss the metal ions, metal chelators, mimetics of the cosubstrate AKG, mimetics of collagen strands, and natural products that are known to inhibit CP4H, focusing on inhibition strategies of therapeutic value and those that have been tested in live cells. Lastly, we discuss the state-of-the-art for inhibition of CP4Hs *in vivo*.

CP4H Structure

In mammals, CP4Hs exist as $\alpha_2\beta_2$ tetramers. The α -subunit is the catalytic subunit and the β -subunit is protein disulfide isomerase, which is a multifunctional chaperone protein that maintains the α -subunit in a soluble and active conformation.^{5,17} Three isoforms of the α -subunit, α (I), α (II), and α (III), have been identified in humans.^{5,17} All α -subunit isoforms form tetramers with the β -subunit, which we refer to herein as the CP4H1, CP4H2, and CP4H3 holoenzymes. The structure of the tetrameric complex is unknown, though low-resolution small-angle X-ray scattering data together with models derived from homologous structures (like those in Figure 3) and bioinformatic analyses suggest an elongated assembly with a $\beta\alpha\alpha\beta$ architecture.²⁵

CP4H Mechanism

As in other members of the Fe(II)- and AKG-dependent dioxygenase (FAKGD) superfamily,^{26,27} CP4H couples the oxidative decarboxylation of AKG to the hydroxylation of a hydrocarbon substrate via a radical mechanism involving a highly reactive Fe(IV)=O species (ferryl ion). The ferryl ion abstracts a hydrogen atom from a hydrocarbon substrate and effects hydroxylation via a radical rebound process (“Coupled Reaction” in Figure 4).²⁶ Although human CP4Hs have been the object of insightful analyses,^{17,28,29} much of what we understand about CP4Hs comes from studies of the homolog from chicken¹⁴ and of other FAKGD enzymes,^{24,26,27} as well as theoretical considerations.³⁰ Classical kinetic studies of chicken CP4H1 have suggested that catalysis proceeds by an ordered *Ter Ter* mechanism in which AKG binds to the CP4H-Fe(II) complex, followed by a protocollagen strand and O₂.^{26,30,31} The peptide bond of the Pro residue that is bound in the enzymic active site is likely in the *cis* (that is, *E*) conformation.^{32,33} After formation of the ferryl ion and subsequent hydroxylation of the Pro residue, the enzyme releases the nascent Hyp-containing strand and succinate.

The ferryl ion appears to decay via an unknown mechanism to an inactive Fe(III) state.³⁴ This aberrant pathway (“Uncoupled Reaction” in Figure 4) is linked to scurvy.³⁵ Reduced vitamin C (L-ascorbic acid) can reactivate the enzyme. During catalysis by the typical *Ter Ter* mechanism, there is no requirement for vitamin C, as the radical rebound process returns the iron center to its Fe(II) state. For example, chicken CP4H catalyzes prolyl hydroxylation at a maximal rate for at least a few cycles in the complete absence of vitamin C.³⁴ Yet, even at saturating concentrations of the peptide substrate, CP4Hs (and other FAKGDs²⁷) have been observed to decarboxylate AKG oxidatively in a manner that does not result in the hydroxylation of the peptide substrate.^{34,36} Some ligands for the iron center enhance the rate of vitamin C-dependent reactivation.³⁷

Alternative Metal Ions

Most FAKGD enzymes have only modest affinity for Fe(II). Accordingly, other transition metal ions, including Mn(II), Co(II), Ni(II), Cu(II), and Zn(II), are modest inhibitors of these enzymes.²⁴ CP4Hs have weak affinity for Fe(II) (*e.g.*, apparent $K_M = 4 \mu\text{M}$ for human CP4H1 and $6 \mu\text{M}$ for human CP4H2²⁹), and are particularly susceptible to inhibition in this manner (*e.g.*, $K_i = 0.6 \mu\text{M}$ for inhibition by Zn(II),³⁸ which has an ionic radius similar to that of Fe(II)). Transition metal ions are, however, unlikely to be useful as CP4H inhibitors in a therapeutic context.

Metal Chelators

As with metal ions, virtually all FAKGD enzymes are inhibited by metal chelators (Figure 5A). These inhibitors include simple chelators such as ethylenediaminetetraacetic acid (EDTA), deferoxamine, 2,2'-bipyridine (bipy), 1,10-phenanthroline, and catechol. Simple metal chelators bind to and sequester free iron, and typically display K_i values in the micromolar range depending on the amount of free iron added to assays *in vitro*.^{39,40} Indeed, bipy itself was used early-on to disrupt collagen biosynthesis in a biological context,^{41,42} providing initial evidence that iron was required for the process. Although metal affinity is an important component of some CP4H inhibitors (*vide infra*), the simple metal chelators used clinically to treat conditions such as iron overload⁴³ and heavy metal toxicity⁴⁴ are too nonspecific to be of therapeutic value as inhibitors of CP4Hs.²⁴ Nevertheless, functionalization or modification of chelators can generate more potent inhibitors of the enzyme.²⁴ For example, 2,2'-bipyridine-5,5'-dicarboxylic acid (bipy55'DC) is one of the most potent known inhibitors of a CP4H ($K_i = 0.14 \mu\text{M}$).^{40,45} Yet, the high residual iron affinity of bipy55'DC is a concern and complicates the characterization of the inhibitory mechanism of this and other bipy derivatives.^{40,45,46}

AKG Mimetics

The putative mechanism of catalysis suggests that analogues of AKG that cannot undergo decarboxylation could inhibit catalysis by CP4Hs.³⁰ Moreover, the mechanism explains why AKG mimics that inhibit CP4Hs do not inhibit the essential enzyme AKG dehydrogenase, which does not require a metal for its catalytic activity.^{30,47} Soon after the enzymatic reaction mechanism was put forth,³⁰ a variety of AKG-mimetic inhibitors of chicken CP4H

were reported, including *N*-oxalylglycine (NOG), pyridine-2,4-dicarboxylic acid (24PDC), pyridine 2,5-dicarboxylic acid (25PDC),³⁹ and 3,4-dihydroxybenzoic acid (DHB) (Figure 5B).⁴⁸ Like AKG ($K_M = 20 \mu\text{M}$), these compounds possess a metal-chelating moiety, and thus resemble the structure of AKG in the context of the enzymatic reaction mechanism. All inhibit human CP4Hs *in vitro*.¹⁷ Whereas 24PDC, 25PDC, and NOG are simple competitive inhibitors with respect to AKG,^{39,48} DHB is competitive with both AKG and vitamin C.⁴⁹

Pyridinedicarboxylic Acids

Unlike 24PDC, which inhibits many members of the FAKGD with similar low micromolar potency,²⁴ 25PDC is selective for CP4Hs. Thus, considerable effort has been made to explore 25PDC derivatives. In particular, a variety of substitutions have been made at both the 2-position⁵⁰ and the 5-position⁵¹ of the pyridine ring (Figure 5B). Regrettably, all of the substitutions for the 2-carboxyl group (*i.e.*, *N*-acyl sulfonamides, tetrazole, hydroxamic acids, 2-imidazolyl, 2-pyrrolyl, and others) were made in the absence of the 5-carboxyl group, thereby obfuscating comparisons. In general, though, modification of the 2-position in the absence of the 5-position carboxyl group did not yield potent inhibitors.⁵⁰ For example, many inhibitors were more potent than the parent pyridine-2-carboxylic acid, but most still exhibited IC_{50} values in the mid-micromolar range,⁵⁰ and have not been developed further.

In comparison to 2-position modifications, those at the 5-position have been more fruitful. Substitution of the 5-carboxyl group in *N*-acyl sulfonamide derivatives yielded 25PDC analogues with similar or improved potency compared to the 25PDC parent.⁵¹ Most of these compounds inhibited with IC_{50} values in the low micromolar range, with phenyl and 4-methoxyphenyl sulfonamides (Figure 5B) displaying the highest potency (IC_{50} values of 1.1 μM and 1.0 μM , respectively, compared to 5.5 μM for 25PDC). Yet, considering the binding mode of AKG, it was highly surprising that substitution of the 5-carboxyl group with neutral derivatives was tolerable in inhibitors. This finding raised the question of whether or not 25PDC and these analogues bound to the active site in the same manner as did AKG itself. No clinical inhibitors based upon these sulfonamide derivatives of 25PDC have been reported to date.

N-Oxalyl Glycine and Its Derivatives

NOG does not display the selectivity properties of 25PDC and is known to inhibit AKG dehydrogenase⁵² and other enzymes that utilize AKG.⁵³ Nevertheless, the NOG scaffold has provided a means to evaluate stereochemistry in the context of the AKG binding pocket. In contrast to the 25PDC scaffold, substitutions to the ω -carboxyl group of the glycine moiety abolished inhibition,⁴⁸ suggesting that NOG likely binds in the conventional AKG binding mode. Still, the α carbon of the glycine moiety has a tetrahedral geometry, so substitutions were made to evaluate the nature of the binding pocket on either side of the AKG substrate.⁴⁸ Upon substitution of the glycine moiety with a variety of other amino acids (*e.g.*, alanine, phenylalanine, leucine, methionine, proline, and sarcosine), the only tolerated substitution was that of alanine, which diminished potency ($\text{IC}_{50} = 90.7 \mu\text{M}$ for racemic *N*-oxalyl alanine (NOA), $\text{IC}_{50} = 2.89 \mu\text{M}$ for NOG).⁴⁸ A striking difference was observed between the 2*R* and

2*S* enantiomers, with (*S*)-NOA ($IC_{50} = 38.2 \mu\text{M}$) (Figure 5B) being substantially more potent than (*R*)-NOA ($IC_{50} = 621 \mu\text{M}$).⁴⁸

Attempts have been made to fuse elements of the *N*-oxalyglycine and pyridinedicarboxylic acid scaffolds. For example, amides of glycine with 4-hydroxycinnoline 3-carboxylic acid or 4-hydroxyisoquinoline 3-carboxylic acid were reported to be potent inhibitors of chicken CP4H *in vitro* ($IC_{50} = 0.02 \mu\text{M}$ and $IC_{50} = 0.19 \mu\text{M}$, respectively).⁵⁴ As with NOA, the 2*S* enantiomer of alanine was preferred in place of glycine, which is consistent with the binding of these alkaloid-like compounds in the enzymic active site. A 7-bromo derivative of the isoquinolyglycine was reported to be especially effective at inhibiting the enzyme in chicken cells and live rats, perhaps because of protection against metabolic inactivation afforded by a 7-bromo group (Figure 5B). This isoquinolyglycine has, however, not been developed further, perhaps because of intrinsic cytotoxicity.

Collagen Mimetic Peptides

CP4H inhibitors based upon the structure of collagen strands could offer enhanced selectivity compared to that of AKG mimetics. Whereas many features of the reaction catalyzed by CP4Hs are well characterized, a detailed understanding of the manner in which the protocollagen substrate is recognized by these CP4Hs remains incomplete, partially due to the limited structural information available for mammalian CP4Hs.

Although there is still not a clear understanding of the peptide substrate, a few individual examples of peptide-based inhibitors have been reported. For example, CP4Hs are known to be inhibited by polyproline strands in the PP-II conformation, which have been observed to bind in the PSB domain of human CP4H-1.⁵⁵ Moreover, some simple analogues of the collagen-repeat sequence in which the composition is modified or rearranged inhibit the enzyme from chicken. For example, peptides of the sequence (Pro-Gly)_{*n*}, (Gly-Pro-Gly)_{*n*}, and (Pro-Pro-Gly-Gly)_{*n*} have all been observed to inhibit chicken CP4H, though (Pro-Pro-Gly-Gly)_{*n*} has also been observed to be a poor substrate.⁵⁶ Moreover, short collagen-like polymers with triplet repeats containing Pro derivatives with methyl or fluoro groups at the 3- or 4-position serve as inhibitors of chicken CP4H.⁵⁷ A small tetrapeptide containing the sequence Gly-Pro-flp-Gly (where flp refers to a (2*S*,4*S*)-4-fluoroproline residue) was found to be a substrate for human CP4H-1, furnishing a product containing (2*S*)-4-ketoproline.³² This product results from the decomposition of the fluorohydrin that is generated by enzymatic catalysis. Thus, a Gly-Pro-flp-Gly substrate can serve as the basis for a continuous assay of enzymatic activity based on the detection of fluoride ion with a selective electrode.⁵⁸ Moreover, the orthogonal reactivity of the ensuing keto group could serve as the basis for a high-throughput assay.⁵⁹

Irreversible Inhibitors

Most known inhibitors of CP4H are reversible. Nonetheless, a few inhibitors are time-dependent and irreversible. These molecules (Figure 5C) include certain anthracyclines, peptides containing (2*S*)-5-oxaproline (OPr), and electrophilic analogues of AKG.

Amongst anthracyclines, doxorubicin and daunorubicin are irreversible inhibitors of chicken CP4H.⁶⁰ They appeared to function via a mechanism that requires Fe(II) in solution but that is competitive with vitamin C, with possible alkylation of the enzymic active site via a postulated radical intermediate. Doxorubicin is used clinically as a chemotherapeutic agent for the treatment of a variety of types of cancer. One of the common side effects of doxorubicin treatment is poor wound healing,⁶¹ which is consistent with the notion that a pleiotropic effect of doxorubicin is inactivation of CP4Hs (as has been observed in human skin fibroblasts⁶²).

Tetrapeptides containing OPr (Figure 5C) are irreversible inhibitors of chicken CP4H.^{63,64} The chemical mechanism of this inactivation is not known but requires Fe(II) and AKG, and is diminished by high concentrations of (Pro-Pro-Gly)₁₀ or vitamin C.⁶³ Peptides containing OPr were even found to inhibit CP4H activity in cultured fibroblasts,⁶⁵ though much less effectively than the inhibition observed *in vitro*.⁶³

Finally, two electrophilic analogues of AKG inactivate CP4H. One is coumalic acid (Figure 5C), which resembles AKG in its chain length and placement of a pendant carboxyl group, but appears to alkylate the enzyme irreversibly by a postulated nucleophilic attack on the coumarin ring.⁶⁶ AKG protects against the inactivation, suggesting that the alkylation likely occurred in or near the enzymic active site. The second is racemic 4-oxo-5,6-epoxyhexanoic acid, in which a carboxy group of AKG is replaced with an epoxy group that could be activated for alkylation upon chelation to the active-site Fe(II).⁶⁷ A nematode, *Caenorhabditis elegans*, exposed to the esterified compound displays the phenotype of animals lacking CP4H.

Inhibition of CP4Hs *In Vivo*

In addition to some compounds described above, a few other CP4H inhibitors have been tested *in vivo*. Few, however, have proven to be sufficiently safe or efficacious to move into clinical studies.²⁴ HOE-077 (Figure 5D), which is a prodrug of 24PDC, has been explored in preclinical studies in rats and dogs and was found to inhibit collagen synthesis in the liver.²⁴ HOE-077 was, however, found to yield many metabolites, and was found to disrupt collagen synthesis by inactivating cells involved in producing collagen, rather than by direct inhibition of CP4H.²⁴ FG-0041 (Figure 5D), which is a derivative of 1,10-phenanthroline, was shown to aid in the recovery of left ventricular function in rats that have suffered post myocardial infarction.²⁴ This compound likely functions, however, by inhibiting hypoxia inducible factor prolyl hydroxylases, rather than CP4Hs.

Arguably, the state-of-the-art compound for inhibiting CP4Hs *in vivo* is ethyl 3,4-dihydroxybenzoate (EDHB) (Figure 5D). EDHB is a prodrug of DHB, and is commonly used in cell culture experiments. Moreover, EDHB has even been shown to be safe enough to use in mouse models of cancer metastasis.²¹ A dose of 40 mg/kg per day in mice was found to decrease primary tumor growth, reduce the collagen content of the primary tumor, and significantly reduce metastasis to the lung, all without affecting overall body weight.²¹ Like other catechols, however, EDHB perturbs iron metabolism⁶⁸ and affects redox

homeostasis.⁶⁹ Thus, EDHB is unlikely to be sufficiently selective for the therapeutic applications of interest here.

A recent screen of biheteroaryl compounds revealed that 2-(5-carboxythiazol-2-yl)pyridine-5-carboxylic acid (pythiDC), in which one pyridyl group of bipy55' DC is replaced with a thiazole moiety, retained potency as a CP4H inhibitor ($K_i = 0.39 \mu\text{M}$) but had a negligible affinity for iron.⁷⁰ Unlike EDHB, the diethyl ester of pythiDC inhibited collagen biosynthesis in human breast cancer cells and human embryonic kidney cells at concentrations that neither cause general toxicity nor disrupt iron homeostasis. pythiDC awaits testing *in vivo*.

Prospectus

The need to treat fibrotic diseases and cancer metastasis along with the importance of CP4Hs in both of these pathologies provide a clear imperative for the development of CP4H inhibitors with therapeutic potential. Although known inhibitors can be effective under some circumstances, most fall short of clinical relevance for a variety of reasons, including low potency in cells, mixed modes of activity, prohibitive toxicity, or other intolerable off-target effects.

Perhaps the most fruitful route to CP4H inhibitors with therapeutic potential entails endowing extant scaffolds with pharmacological attributes. Compounds that not only mimic proline or AKG but also interact with enzymic residues that are unique to CP4H hold promise. We also look forward to the revelation of a high-resolution crystal structure of CP4H, which would provide highly desirable guidance towards filling an important gap in the pharmacopeia.

ACKNOWLEDGEMENT

We are grateful to M. D. Shoulders and I. C. Tanrikulu for comments on the manuscript.

Funding

J.D.V. was supported by Molecular Biosciences Training Grant T32 GM007215 [National Institutes of Health (NIH)] and a fellowship from the Department of Biochemistry at the University of Wisconsin–Madison. Work on collagen prolyl 4-hydroxylase in the Raines laboratory is supported by grant R01 AR044276 (NIH).

BIOGRAPHIES

James D. Vasta received his B.S. degree in Biochemistry from Lafayette College in May 2010. He received his Ph.D. degree in Biochemistry from the University of Wisconsin–Madison in December 2105. His doctoral thesis work was focused on the development of selective inhibitors and modulators of collagen prolyl 4-hydroxylase. He is now a scientist at Promega Corp. in Madison, Wisconsin.

Ronald T. Raines is the Firmenich Professor of Chemistry at the Massachusetts Institute of Technology. Prior to July 2017, he was a professor in the biochemistry and chemistry departments at the University of Wisconsin–Madison. His research group is focused on the chemistry and biology of proteins, including collagen and collagen prolyl 4-hydroxylase.

ABBREVIATIONS USED

AKG	α -ketoglutarate
CP4H	collagen prolyl 4-hydroxylase
FAKGD	Fe(II)- and AKG-dependent dioxygenase

References

- (1). Shoulders MD; Raines RT Collagen structure and stability. *Annu. Rev. Biochem* 2009, 78, 929–958. [PubMed: 19344236]
- (2). Fischer E Ueber die Hydrolyse des Caseins durch Salzsäure. *Z. Physiol. Chem* 1901, 33, 151–176.
- (3). Fischer E Über eine neue Aminosäure aus Leim. *Chem. Ber* 1902, 35, 2660–2665.
- (4). Myllyharju J; Kivirikko KI Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* 2004, 20, 33–43. [PubMed: 14698617]
- (5). Gorres KL; Raines RT Prolyl-4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol* 2010, 45, 106–124. [PubMed: 20199358]
- (6). Berg RA; Prockop DJ The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple helix of collagen. *Biochem. Biophys. Res. Comm* 1973, 52, 115–120. [PubMed: 4712181]
- (7). Bretscher LE; Jenkins CL; Taylor KM; DeRider ML; Raines RT Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc* 2001, 123, 777–778. [PubMed: 11456609]
- (8). Gelse K; Pöschl E; Aigner T Collagens—structure, function, and biosynthesis. *Adv. Drug Deliv. Rev* 2003, 55, 1531–1546. [PubMed: 14623400]
- (9). Sharma U; Carrique L; Vadon-Le Goff S; Mariano N; Georges R-N; Delolme F; Koivunen P; Myllyharju J; Moali C; Aghajari N; Hulmes DJS Structural basis of homo- and heterotrimerization of collagen I. *Nat. Commun* 2017, 8, 14671. [PubMed: 28281531]
- (10). Makareeva E; Nydea AA; Leikin S Chaperoning osteogenesis: New protein-folding disease paradigms. *Trends Cell Biol.* 2011, 21, 168–176. [PubMed: 21183349]
- (11). Wong MY; Doan ND; DiChiara AS; Papa LJ III; Cheah JH; Soule CK; Watson N; Hulleman JD; Shoulders MD A high-throughput assay for collagen secretion suggests an unanticipated role for Hsp90 in collagen production. *Biochemistry* 2018, 57, 2814–2827. [PubMed: 29676157]
- (12). DiChiara AS; Taylor RJ; Wong MY; Doan N-D; Del Rosario AM; Shoulders MD Mapping and exploring the collagen-I proteostasis network. *ACS Chem. Biol* 2016, 11, 1408–1421. [PubMed: 26848503]
- (13). Gorur A; Yuan L; Kenny SJ; Baba S; Xu K; Schekman R COPII-Coated membranes function as transport carriers of intracellular procollagen I. *J. Cell Biol.* 2017, 216, 1745–1759. [PubMed: 28428367]
- (14). Kivirikko KI; Prockop DJ Hydroxylation of proline in synthetic polypeptides with purified procollagen hydroxylase. *J. Biol. Chem* 1967, 242, 4007–4012. [PubMed: 6061694]
- (15). Cohen IK Can collagen metabolism be controlled: Theoretical considerations. *J. Trauma* 1985, 25, 410–412. [PubMed: 3923206]
- (16). Hanauske-Abel HM Prolyl 4-hydroxylase, a target enzyme for drug development. Design of suppressive agents and the *in vitro* effects of inhibitors and proinhibitors. *J. Hepatol* 1991, 13 (Suppl. 3), S8–S16.
- (17). Myllyharju J Prolyl 4-hydroxylases, key enzymes in the synthesis of collagens and regulation of the response to hypoxia, and their roles as treatment targets. *Ann. Med* 2008, 40, 402–417. [PubMed: 19160570]
- (18). Rosenbloom J; Castro SV; Jimenez SA Narrative review: Fibrotic diseases: Cellular and molecular mechanisms and novel therapies. *Ann. Intern. Med* 2010, 152, 159–166. [PubMed: 20124232]

- (19). Friedman SL; Sheppard D; Duffield JS; Violette S Therapy for fibrotic diseases: Nearing the starting line. *Sci. Transl. Med* 2013, 5, 1–17.
- (20). Esbona K; Yi Y; Saha S; Yu M; Van Doom RR; Conklin MW; Graham DS; Wisinski KB; Ponik SM; Eliceiri KW; Wilke LG; Keely PJ The presence of cyclooxygenase 2, tumor-associated macrophages, and collagen alignment as prognostic markers for invasive breast carcinoma patients. *Am. J. Pathol* 2018, 188, 559–573. [PubMed: 29429545]
- (21). Gilkes DM; Chaturvedi P; Bajpai S; Wong CC; Wei H; Pitcairn S; Hubbi ME; Wirtz D; Semenza GL Collagen prolyl hydroxylases are essential for breast cancer metastasis. *Cancer Res.* 2013, 73, 3285–3296. [PubMed: 23539444]
- (22). Stylianopoulos T; Munn LL; Jain RK Engineering the physical microenvironment of tumors to improve drug delivery and efficacy: From mathematical modeling to bench to bedside. *Trends Cancer* 2018, 4, 292–319. [PubMed: 29606314]
- (23). Günzler V; Weidmann K, Prolyl 4-hydroxylase inhibitors In *Prolyl Hydroxylase, Protein Disulfide Isomerase, and Other Structurally Related Proteins*, Guzman NA, Ed. Marcel Dekker: New York, 1998; pp 65–95.
- (24). Rose NR; McDonough MA; King ONF; Kawamura A; Schofield CJ Inhibition of 2-oxoglutarate dependent oxygenases. *Chem. Soc. Rev* 2011, 40, 4364–4397. [PubMed: 21390379]
- (25). Koski MK; Anantharajan J; Kursula P; Dhavala P; Murthy AV; Bergmann U; Myllyharju J; Wierenga RK Assembly of the elongated collagen prolyl 4-hydroxylase $\alpha_2\beta_2$ heterotetramer around a central α_2 dimer. *Biochem. J* 2017, 474, 751–769. [PubMed: 28093469]
- (26). Hausinger RP Fe(II)/ α -ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol* 2004, 39, 21–68.
- (27). Hausinger RP; Schofield CJ, Eds. *2-Oxoglutarate-Dependent Oxygenases*. The Royal Society of Chemistry: Cambridge, UK, 2015.
- (28). Kivirikko KI; Myllyharju J Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol.* 1998, 16, 357–368. [PubMed: 9524356]
- (29). Myllyharju J Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol.* 2003, 22, 15–24. [PubMed: 12714038]
- (30). Hanauske-Abel HM; Günzler V A stereochemical concept for the catalytic mechanism of prolylhydroxylase: Applicability to classification and design of inhibitors. *J. Theor. Biol* 1982, 94, 421–455. [PubMed: 6281585]
- (31). Myllylä R; Tuderman L; Kivirikko KI Mechanism of the prolyl hydroxylase reaction. 2. Kinetic analysis of the reaction sequence. *Eur. J. Biochem* 1977, 80, 349–357. [PubMed: 200425]
- (32). Gorres KL; Edupuganti R; Krow GR; Raines RT Conformational preferences of substrates for human prolyl 4-hydroxylase. *Biochemistry* 2008, 47, 9447–9455. [PubMed: 18702512]
- (33). Vasta JD; Choudhary A; Jensen KH; McGrath NA; Raines RT Prolyl 4-hydroxylase: Substrate isosteres in which an (*E*)- or (*Z*)-alkene replaces the prolyl peptide bond. *Biochemistry* 2017, 56, 219–227. [PubMed: 28001367]
- (34). De Jong L; Albracht SPJ; Kemp A Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron: The role of ascorbate in peptidyl proline hydroxylation. *Biochim. Biophys. Acta* 1982, 704, 326–332. [PubMed: 6285984]
- (35). Kivirikko KI; Risteli L Biosynthesis of collagen and its alterations in pathological states. *Med. Biol* 1976, 54, 159–186. [PubMed: 181643]
- (36). Myllylä R; Majamaa K; Gunzler V; Hanauske-Abel HM; Kivirikko KI Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *J. Biol. Chem* 1984, 259, 5403–5405. [PubMed: 6325436]
- (37). Vasta JD; Raines RT Human collagen prolyl 4-hydroxylase is activated by ligands for its iron center. *Biochemistry* 2016, 55, 3224–3233. [PubMed: 27183028]
- (38). Kivirikko KI; Myllylä R Recent developments in posttranslational modification: Intracellular processing. *Methods Enzymol.* 1987, 144, 96–114. [PubMed: 3041180]
- (39). Majamaa K; Hanauske-Abel HM; Günzler V; Kivirikko KI The 2-oxoglutarate binding site of prolyl 4-hydroxylase. Identification of distinct subsites and evidence for 2-oxoglutarate decarboxylation in a ligand reaction at the enzyme-bound ferrous ion. *Eur. J. Biochem* 1984, 138, 239–245. [PubMed: 6321169]

- (40). Hales NJ; Beattie JF Novel inhibitors of prolyl 4-hydroxylase. 5. The intriguing structure–activity relationships seen with 2,2′-bipyridine and its 5,5′-dicarboxylic acid derivatives. *J. Med. Chem* 1993, 36, 3853–3858. [PubMed: 8254616]
- (41). Whitson TC; Peacock EE Jr. Effect of alpha alpha dipyridyl on collagen synthesis in healing wounds. *Surg. Gynecol. Obstet* 1969, 128, 1061–1064. [PubMed: 5779738]
- (42). Switzer BR; Summer GK Inhibition of collagen synthesis by α, α' -dipyridyl in human skin fibroblasts in culture. *In Vitro* 1973, 9, 160–166. [PubMed: 4779003]
- (43). Sheth S Iron chelation: An update. *Curr. Opin. Hematol* 2014, 21, 179–185. [PubMed: 24504090]
- (44). Flora SJS; Pachauri V Chelation in metal intoxication. *Int. J. Environ. Res. Public Health* 2010, 7, 2745–2788. [PubMed: 20717537]
- (45). Vasta JD; Raines RT Selective inhibition of prolyl 4-hydroxylases by bipyridinedicarboxylates. *Bioorg. Med. Chem* 2015, 23, 3081–3090. [PubMed: 26022078]
- (46). Franklin TJ; Morris WP; Edwards PN; Large MS; Stephenson R Inhibition of prolyl 4-hydroxylase *in vitro* and *in vivo* by members of a novel series of phenanthrolinones. *Biochem. J* 2001, 353, 333–338. [PubMed: 11139398]
- (47). Majamaa K; Turpeenniemi-Hujanen TM; Latipää P; Günzler V; Hanauske-Abel HM; Hassinen IE; Kivirikko KI Differences between collagen hydroxylases and 2-oxoglutarate dehydrogenase in their inhibition by structural analogues of 2-oxoglutarate. *Biochem. J* 1985, 229, 127–133. [PubMed: 2994628]
- (48). Cunliffe CJ; Franklin TJ; Hales NJ; Hill GB Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analogue *N*-oxaloglycine and its derivatives. *J. Med. Chem* 1992, 35, 2652–2658. [PubMed: 1321909]
- (49). Majamaa K; Günzler V; Hanauske-Abel HM; Myllylä R; Kivirikko KI Partial identity of the 2-oxoglutarate and ascorbate binding sites of prolyl 4-hydroxylase. *J. Biol. Chem* 1986, 261, 7819–7823. [PubMed: 3011801]
- (50). Dowell RI; Hales NH; Tucker H Novel inhibitors of prolyl 4-hydroxylase. Part 4. Pyridine-2-carboxylic acid analogues with alternative 2-substituents. *Eur. J. Med. Chem* 1993, 28, 513–516.
- (51). Dowell RI; Hadley EM Novel inhibitors of prolyl 4-hydroxylase. *J. Med. Chem* 1992, 35, 800–804. [PubMed: 1312598]
- (52). Rife JE; Cleland WW Determination of the chemical mechanism of glutamate dehydrogenase for pH studies. *Biochemistry* 1980, 19, 2328–2333. [PubMed: 7190025]
- (53). Northrop DB; Cleland WW The kinetics of pig heart triphosphopyridine nucleotide-isocitrate dehydrogenase. *J. Biol. Chem* 1974, 249, 2928–2931. [PubMed: 4151309]
- (54). Franklin TJ; Hales NJ; Johnstone D; Morris WB; Cunliffe CJ; Millest AJ; Hill GB Approaches to the design of anti-fibrotic drugs. *Biochem. Soc. Trans* 1991, 19, 812–815. [PubMed: 1665444]
- (55). Anantharajan J; Koski MK; Kursula P; Hieta R; Bergmann U; Myllyharju J; Wierenga RK The structural motifs for substrate binding and dimerization of the α subunit of collagen prolyl 4-hydroxylase. *Structure* 2013, 21, 2107–2118. [PubMed: 24207127]
- (56). Rapaka RS; Urry DW; Ryu RKN; Bhatnagar RS Collagen and elastin analogues: Syntheses, hydroxylation, calcium binding, and coacervation properties of (Pro-Gly)_n, (Gly-Pro-Gly)_n, and (Pro-Pro-Gly-Gly)_n. *Ala. J. Med. Sci* 1978, 15, 196–212. [PubMed: 665943]
- (57). Hutton JJ; Marglin A; Witkop B; Kurtz J; Berger A; Udenfriend S Synthetic polypeptides as substrates and inhibitors of collagen proline hydroxylase. *Arch. Biochem. Biophys* 1968, 125, 779–785. [PubMed: 5671042]
- (58). Gorres KL; Raines RT Direct and continuous assay for prolyl 4-hydroxylase. *Anal. Biochem* 2009, 386, 181–185. [PubMed: 19111518]
- (59). Zhou J; Xu G; Han R; Dong J; Zhang W; Zhang R; Ni Y Carbonyl group-dependent high-throughput screening and enzymatic characterization of diaromatic ketone reductase. *Catal. Sci. Technol* 2016, 6, 6320–6327.
- (60). Günzler V; Hanauske-Abel HM; Myllylä R; Kaska DD; Hanauske A; Kivirikko KI Syncatalytic inactivation of prolyl 4-hydroxylase by anthracyclines. *Biochem. J* 1988, 251, 365–372. [PubMed: 2840891]

- (61). Bland KI; Palin WE; von Fraunhofer JA; Morris RR; Adcock RA; Tobin GR II. Experimental and clinical observations of the effects of cytotoxic chemotherapeutic drugs on wound healing. *Ann. Surg* 1984, 199, 782–790. [PubMed: 6732317]
- (62). Sasaki T; Holeyfield KC; Uitto J Doxorubicin-induced inhibition of prolyl hydroxylation during collagen biosynthesis in human skin fibroblast cultures. *J. Clin. Invest* 1987, 80, 1735–1741. [PubMed: 2824566]
- (63). Günzler V; Brocks D; Henke S; Myllylä R; Geiger R; Kivirikko KI Syncatalytic inactivation of prolyl 4-hydroxylase by peptides containing the unphysiologic amino acid 5-oxaproline. *J. Biol. Chem* 1988, 263, 19498–19504. [PubMed: 2848831]
- (64). Wu M; Moon H-S; Begley TP Mechanism-based inactivation of the human prolyl-4-hydroxylase by 5-oxaproline-containing peptides: Evidence for a prolyl radical intermediate. *J. Am. Chem. Soc* 1999, 121, 587–588.
- (65). Karvonen K; Ala-Kokko L; Pihlajaniemi T; Helaakoski T; Henke S; Günzler V; Kivirikko KI; Savolainen ER Specific inactivation of prolyl 4-hydroxylase and inhibition of collagen synthesis by oxaproline-containing peptide in cultured human skin fibroblasts. *J. Biol. Chem* 1990, 265, 8415–8419. [PubMed: 2160457]
- (66). Günzler V; Hanauske-Abel HM; Myllylä R; Mohr J; Kivirikko KI Time-dependent inactivation of chick-embryo prolyl 4-hydroxylase by coumalic acid. *Biochem. J* 1987, 242, 163–169. [PubMed: 3036081]
- (67). Vasta JD; Higgin JJ; Kersteen EA; Raines RT Bioavailable affinity label for collagen prolyl 4-hydroxylase. *Bioorg. Med. Chem* 2013, 21, 3597–3601. [PubMed: 23702396]
- (68). Wang J; Buss JL; Chen G; Ponka P; Pantopoulos K The prolyl 4-hydroxylase inhibitor ethyl-3,4-dihydroxybenzoate generates effective iron deficiency in cultured cells. *FEBS Lett.* 2002, 529, 309–312. [PubMed: 12372619]
- (69). Schweigert N; Alexander J; Zehnder B; Eggen RIL Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. *Environ. Microbiol* 2001, 3, 81–91. [PubMed: 11321547]
- (70). Vasta JD; Andersen KA; Deck KM; Nizzi CP; Eisenstein RS; Raines RT Selective inhibition of collagen prolyl 4-hydroxylase in human cells. *ACS Chem. Biol* 2016, 11, 193–199. [PubMed: 26535807]

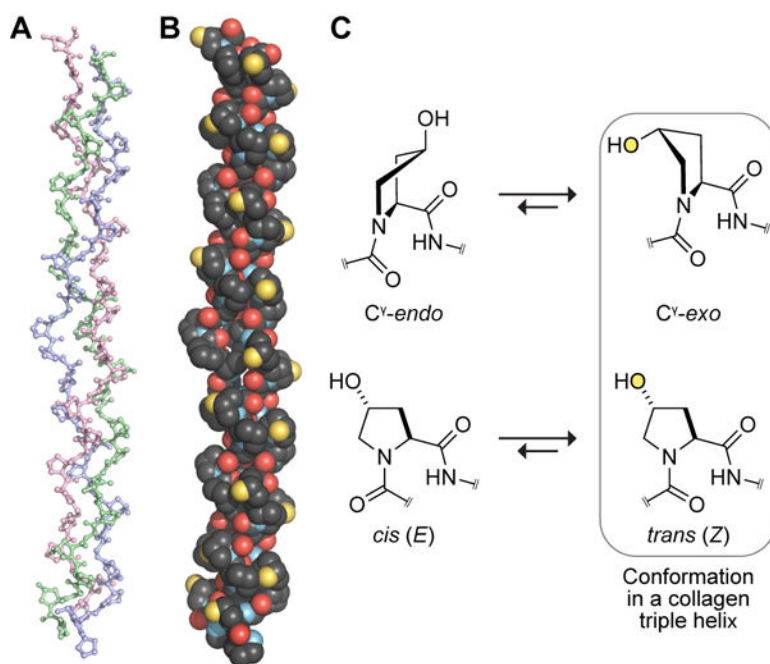


Figure 1.

The collagen triple helix. (A) Ball-and-stick model of a (Pro-Hyp-Gly)_n triple helix with each strand in a different color. Atomic coordinates are from PDB entry 1cgd. (B) Space-filling model of the triple helix from panel A. Oxygen atoms of the Hyp residues installed by CP4H are in yellow; other atoms are in CPK colors. (C) The 4*R* hydroxy group of Hyp residues increases the conformational stability of collagen triple helices by stabilizing the requisite *C^γ-exo* pyrrolidine ring pucker with a *gauche* effect and *trans (Z)* peptide bond with an *n*→*π** interaction.⁷

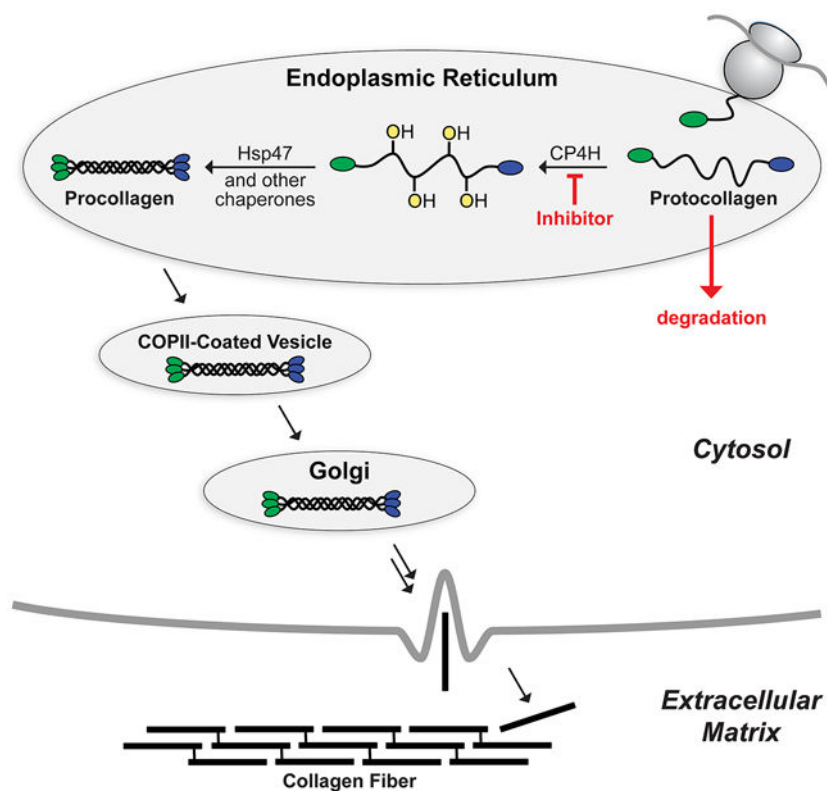


Figure 2. Putative biosynthetic pathway of collagen. Collagen synthesis begins in the rough endoplasmic reticulum. After protocollagen strands are translated by the ribosome, their globular N- and C-terminal propeptide domains establish the composition and register of a nascent triple helix. After extensive post-translational modification by CP4H and other enzymes, procollagens form triple helices, a process facilitated by chaperones such as Hsp47. Procollagen triple helices are transported to the Golgi in enlarged COPII-coated vesicles. Ultimately, processed fibrils are incorporated into growing collagen fibers, which are stabilized by covalent crosslinks between triple helices. The inhibition of CP4H prevents triple-helix formation and likely leads to the degradation of protocollagen strands.

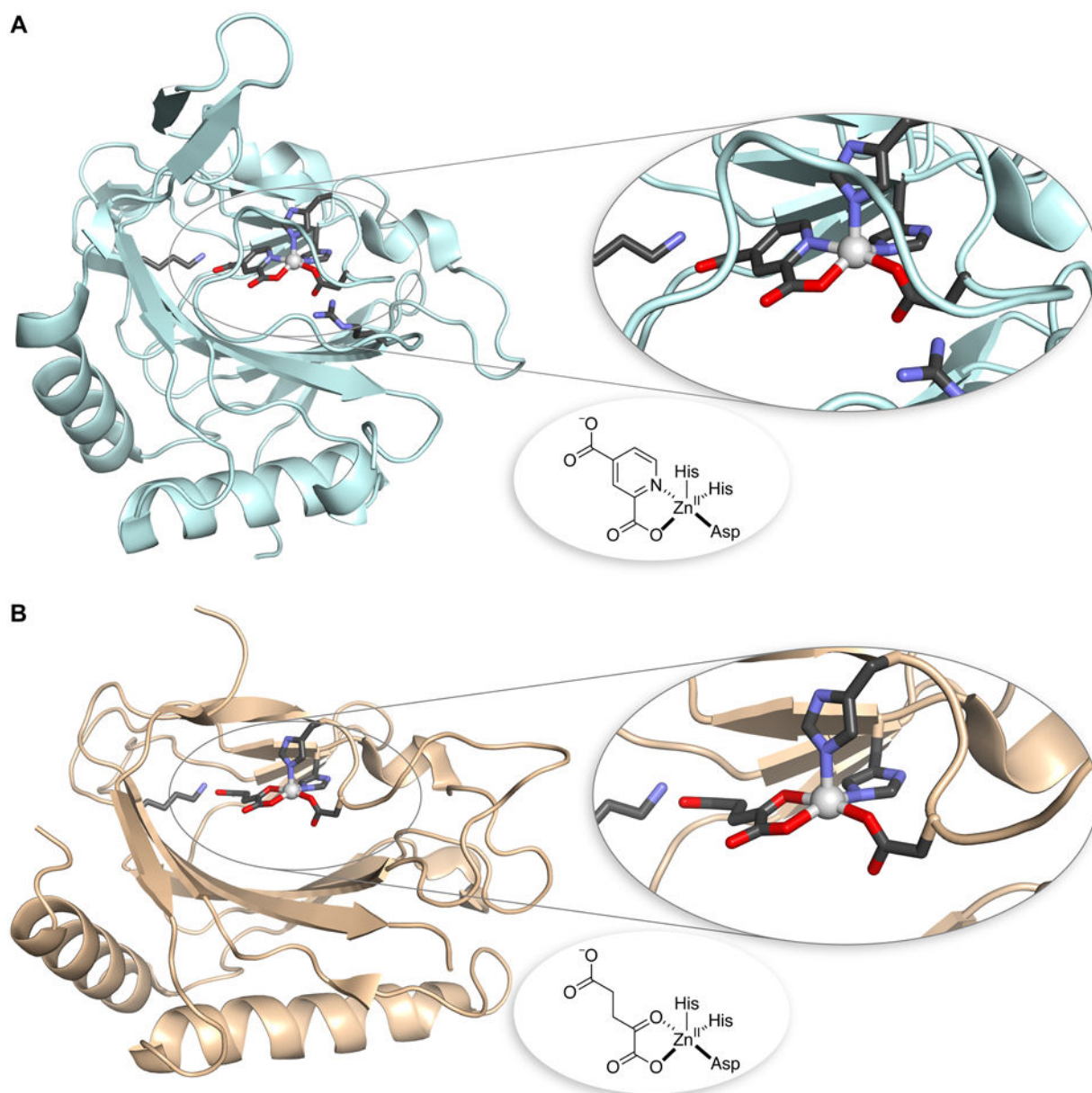


Figure 3. Structural homologs of the catalytic domain of CP4Hs. (A) Crystal structure of *Chlamydomonas reinhardtii* P4H-1 in a complex with 2,4-pyridinedicarboxylate (24PDC) (PDB entry 2jig). In the active site, Zn(II) (sphere) is bound to two His side chains, an Asp side chain, and the C-2 carboxylate group and nitrogen of 24PDC. The C-4 carboxylate group is engaged in a Coulombic interaction with a Lys side chain. (B) Crystal structure of *Paramecium bursaria* chlorella virus-1 P4H in a complex with AKG (PDB entry 5c5t). In the active site, Zn(II) (sphere) is bound to two His side chains, an Asp side chain, and the C-1 carboxylate and C-2 keto groups of AKG. The C-5 carboxylate group is engaged in a Coulombic interaction with a Lys side chain.

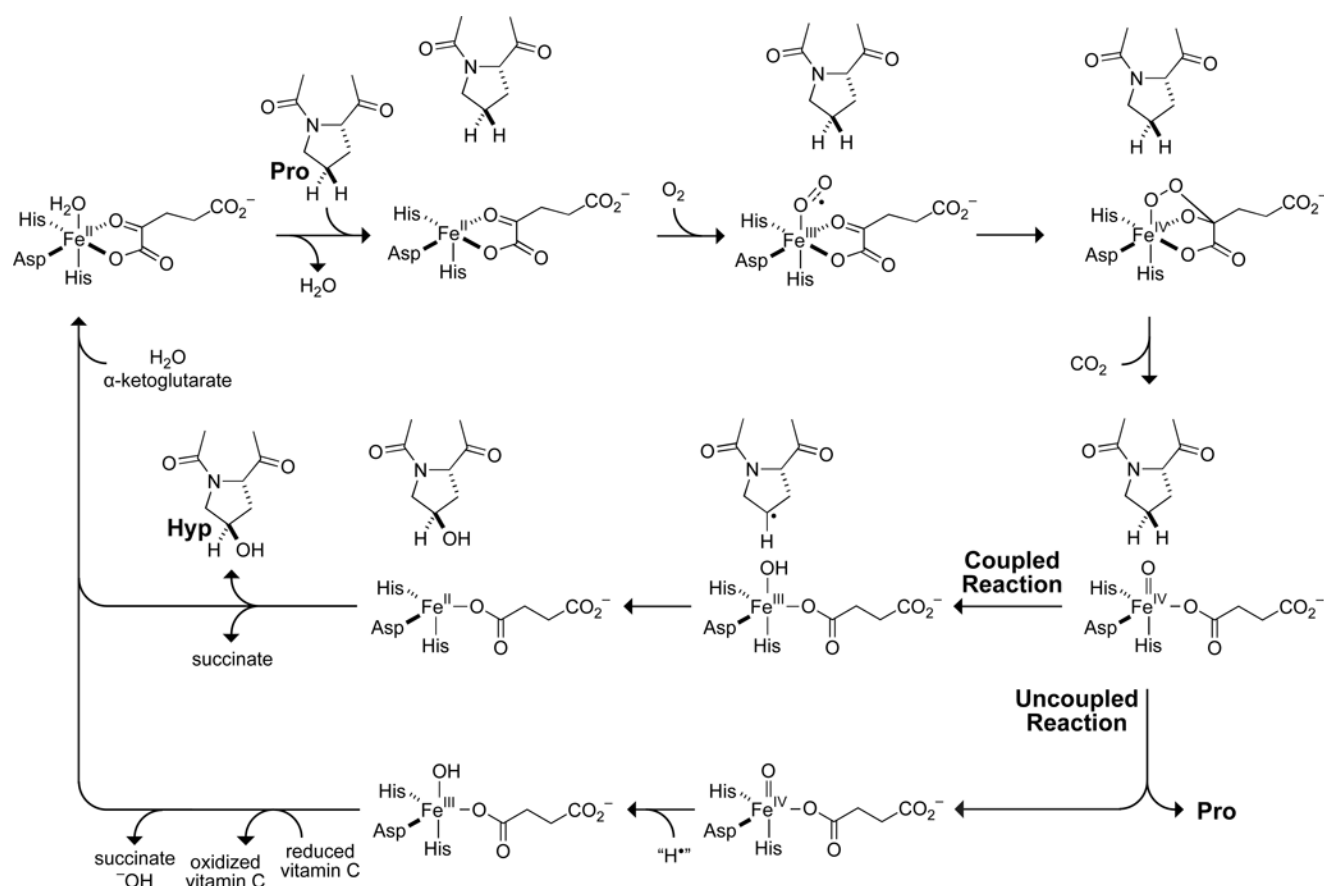


Figure 4. Putative mechanism of catalysis by CP4Hs. The configuration of the iron ligands is based on Figure 3. Neither the precise order of substrate binding and product release^{26,30,31} nor the source of the hydrogen radical shown in the uncoupled reaction is known.

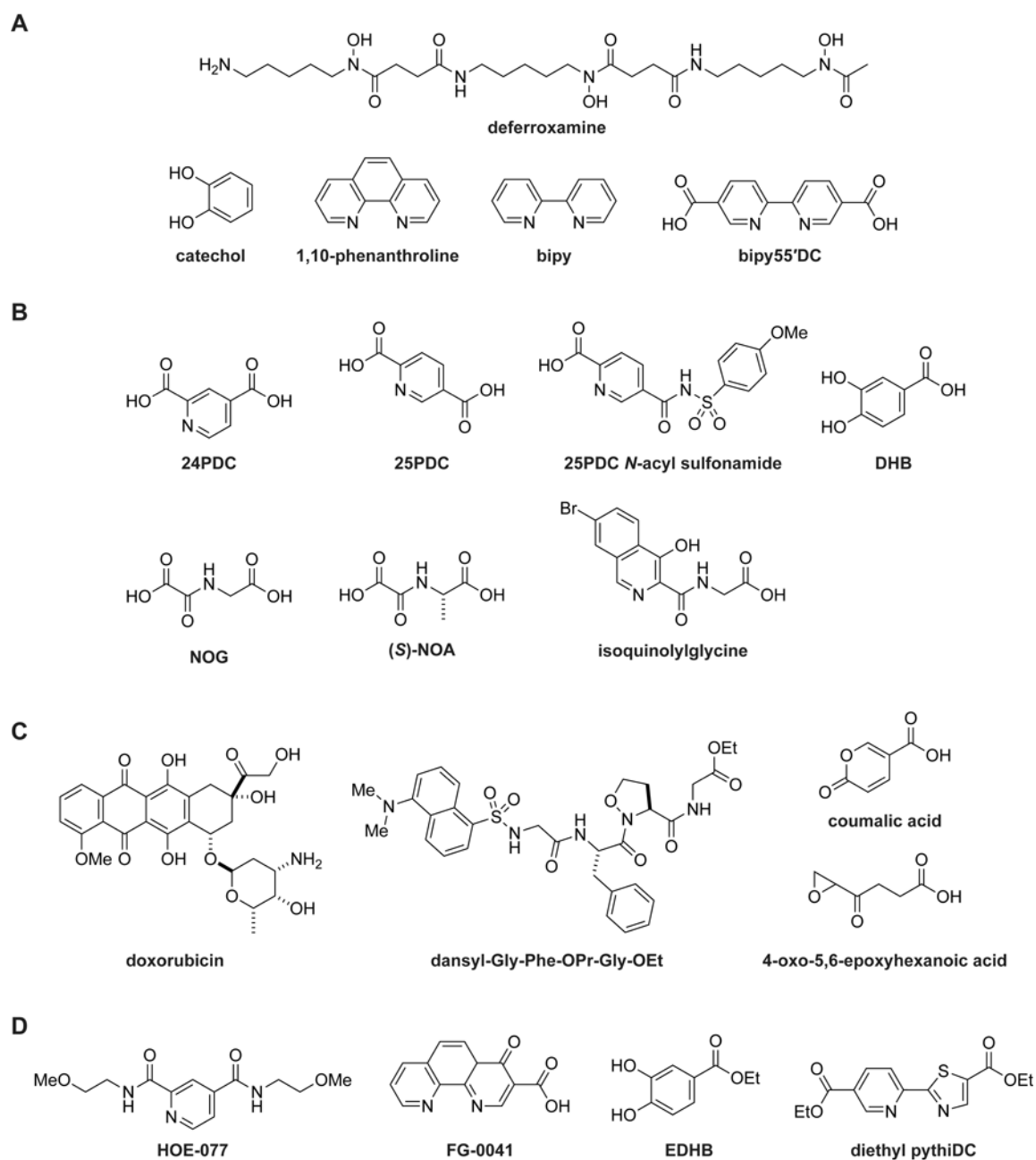


Figure 5. Examples of known CP4H inhibitors. In addition to metal ions and some collagen-like peptides, CP4Hs are inhibited by (A) simple metal chelators and their derivatives, (B) mimetics of the AKG cosubstrate, and (C) a variety of irreversible inhibitors. A few other compounds (D) have been tested *in cellulo* and *in vivo*.