Stoichiometric and Crystallographic Investigations of Isonitrile Biosynthesis by ScoE

Ian Gutierrez, Siyue Yang

Mentor: Maanasa Narayanamoorthy

University of California, Berkeley

Abstract

Isonitrile, an electron-rich functional group, has a variety of biological functions — including metal acquisition, detoxification, and virulence — in numerous organisms. For a long time, the understanding of its biosynthetic pathway was thought to be restricted to the isonitrile synthase (IsnA) family. However, a series of recent studies have revealed a distinct pathway for the biosynthesis of a unique class of isonitrile lipopeptides (INLPs), which are critical for the virulence of pathogenic mycobacteria, such as *Mycobacterium tuberculosis*, via a mononuclear Fe(II)/α-ketoglutarate (α-KG) dependent dioxygenase, which utilizes an Fe(IV)-oxo intermediate species to initiate oxidative transformations with various biological functions (*2*). This research uses ScoE, an enzyme from the mononuclear Fe(II)/α-KG dependent dioxygenase superfamily, to investigate the reaction mechanism and enzyme-substrate interactions of this novel pathway. The stoichiometry of the ScoE-catalyzed isonitrile formation was determined, with the general scheme of the substrate, (*R*)-3-((carboxylmethyl)amino)butanoic acid (CABA), and cosubstrate, α-KG, into (R)-3-isocyanobutanoic acid (INBA) and succinate, respectively. Crystal structural analysis of ScoE was used to identify specific binding sites for CABA and α-KG. Together, through comprehensive computational analysis of reaction stoichiometry and crystal structure analysis, this research offers insight into how ScoE performs this unprecedented reaction.

Introduction

Isonitriles, also known as isocyanides, are unique compounds that consist of the functional group -C≡N which has two resonance structures. One of these structures demonstrates carbenic character with a negatively charged carbon that has a lone pair of electrons and a positively charged nitrogen. This allows for unique chemistry by this electron-rich functional

group, and this group is found in an array of natural products and metabolites produced by both terrestrial and marine organisms as well as plants. Considering well-known examples such as the antiviral agent xanthocillin or the antibiotics aerocyanidin and amycomicin (*3*), it is clear the positive impact isonitriles have had on the pharmaceutical industry, and thus, why these compounds are of such big interest. However, despite this wide variety of reactions and products, isonitrile biosynthesis was known to only be done by the isonitrile synthase family, IsnA. This pathway converts an amino acid to an isonitrile and uses ribulose-5-phosphate as a cosubstrate.

Research on a gene cluster widely present in Actinobacteria (gram-positive bacteria with high amounts of guanine and cytosine in their DNA(*4*)), however, revealed a novel pathway for the synthesis of isonitriles as part of the formation of nonribosomal isonitrile lipopeptides (INLPs) (*5*). These nonribosomal peptides are one of the most important secondary metabolites — secondary indicating that these small molecules produced by organisms are not essential for their growth and development (*6*) — and are often biosynthesized by nonribosomal peptide synthases (NRPSs) (*7*). Upon further research of this topic in *Streptomyces coeruleorubidus* led to the discovery of a mononuclear Fe(II)/α-KG dependent dioxygenase known as ScoE, which was speculated to have similarities to the well-known member of this family, TauD. TauD was the first in this family to be understood, and is still considered the ultimate example when it comes to what this family does mechanistically. This large family is known to use the oxidative decarboxylation of α-KG to activate molecular oxygen to catalyze various reactions. It has also been shown that this family maintains a common 2-histidine-1-carboxylate motif coordinated at the iron atom often referred to as a facial triad (*8*). Sure enough, it was observed that ScoE was no exception to this in that while it catalyzes the formation of the isonitrile compound, INBA, from the carboxylic acid substrate, CABA, the similar 2-histidine-1-aspartate facial triad is

maintained throughout the course of the reaction (*5*). The pathway in which ScoE carries out this catalysis differs from the originally known IsnA pathway in that only one substrate is used, a β-glycine adduct attached to a short fatty acyl chain.

In this paper, we begin to understand this novel pathway by investigating stoichiometry and crystal structures to gain insight into what is consumed, what is produced, and the binding interactions taking place between CABA, α-KG, iron, and ScoE.

Methods

General Method

A Polymerase Chain Reaction (PCR) was performed to amplify the ScoE gene, and then these were embedded into vectors (pET24 B, pETCDFDuet-1, or pET30) in order to make plasmids as to allow for cloning. The UC Berkeley Sequencing Facility used sequencing to verify that the plasmids constructed were correct. pET30-ScoE specifically was used to make all mutations, which were done by PCR using Phusion HighFidelity DNA Polymerase.

Obtaining ScoE for experiments

ScoE was expressed in BL21 cells as to allow many copies to be generated in order to have a sufficient amount to carry out this research.

Making the holo-enzyme

Ethylenediaminetetraacetic acid (EDTA) was added to ScoE during purification to prevent degradation by metalloproteases. ScoE was then desalted to remove EDTA before being immediately added to the biochemical assay of interest.

Preparing pure ScoE for stoichiometric experiments

Reactions were initiated by the addition of iron.

Quantifying Oxygen consumption

Oxygen consumption was collected as a function of time by an Oxygen Hansatech probe, and the reaction was allowed to incubate until the profile reached a steady state with, initially, only α-KG. Once the concentration of O_2 plateaued, CABA was added and measuring continued.

Quantifying α-KG, CABA, Succinate, and INBA

A click reaction was used to quantify INBA by introducing tetrazine to the biochemical assay of interest, and allowing it to react with INBA. This reaction yields py-aminopyrazole which is easier to detect on the mass spectrometer than INBA itself. It was observed that no free-INBA remained upon completion of the click reaction so using LC-HRMS on py-aminopyrazole was used to indirectly quantify INBA formation. Assays with varying concentrations of α-KG, succinate, CABA, and py-aminopyrazole were prepared and analyzed by LC-HRMS. For quantification of α-KG, succinate, CABA, only part of the assay mixture was analyzed to allow the remaining mixture to be used for the click reaction to quantify INBA (as mentioned previously).

Isotope labeling of CABA

Bromoacetic acid-2-^{13}C was used to form our isotope-labeled substrate, CABA-6-^{13}C, to help determine if this carboxyl group was partially lost as CO or completely lost as CO_2; GC-MS was used to detect this. This question was further investigated by creation of a Sorbet band that was

then compared to the literature. In addition to this, to rule out the possibility that the carboxyl-group is loss by deformylation, two assays were analyzed, and LC-HRMS was used.

Reduction of Oxygen

Two assays were once again prepared and analyzed to see how O_2 in this ScoE catalyzed reaction was reduced; it is speculated to be H_2O_2 or H_2O.

Crystallization

The sitting drop vapor technique was used to grow quality crystals to be used for crystallography experiments. Initially, crystals were oriented to provide insight into how CABA, Fe(II), and ScoE bind before adding α-KG into the complex.

Results

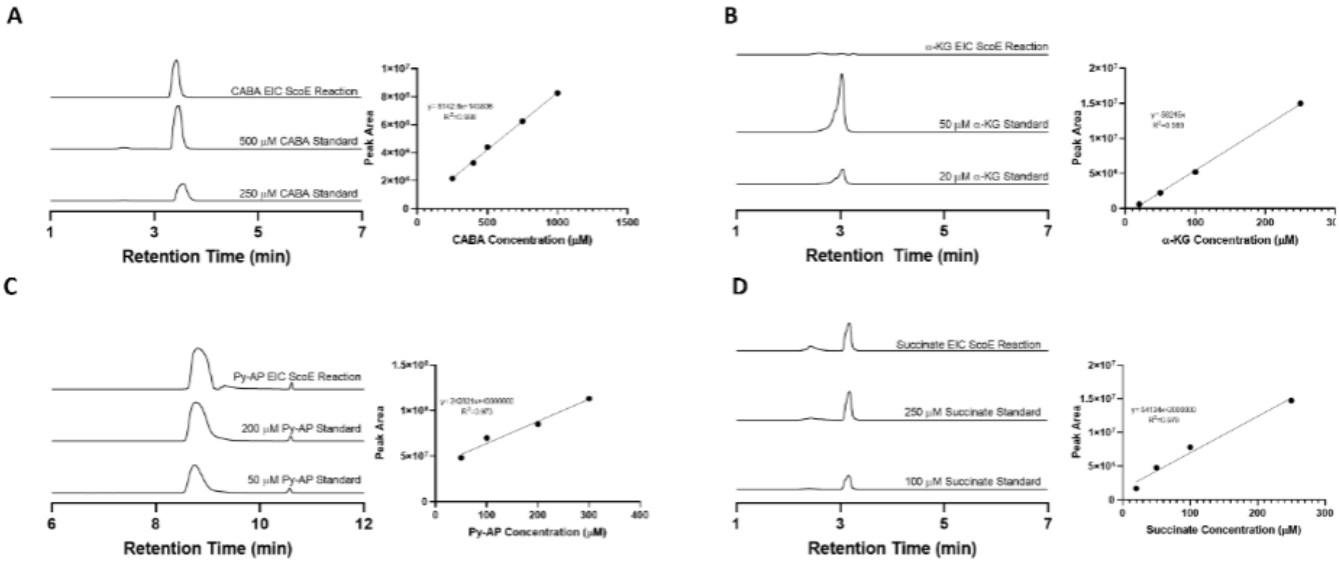
Overall Stoichiometry

From the LC-HRMS and Oxygen probe experiments, a 1:1:1 relationship was confirmed between α-KG consumption, oxygen consumption, and succinate formation as well as the predicted 1:1 correlation between CABA consumption and INBA formation. From this, it can also be seen that 2 equivalents of α-KG are consumed by this reaction for every 1 equivalent of INBA formed. These relationships can be seen in Table 1 below.

Table 1: Observed amounts consumed and produced by ScoE

CABA consumption (μM)	α-KG consumption (μM)	O_2 consumption (μM)	Succinate production (μM)	INBA production (μM)
108 ± 3.3	221.4 ± 28.4	221.5 ± 13.5	230.1 ± 15.1	116.4 ± 15.2

Figure 2: Quantification of the reagents in Table 1



Data obtained by LC-HRMS verifying what is consumed and what is produced by the ScoE catalysis reaction.

Oxygen Consumption

By use of an Oxygen Hansatech probe, it was shown that while oxygen is not consumed with CABA alone, it is consumed solely with α-KG and further consumed upon addition of CABA with the α-KG.

Figure 1: Observed O_2 consumption

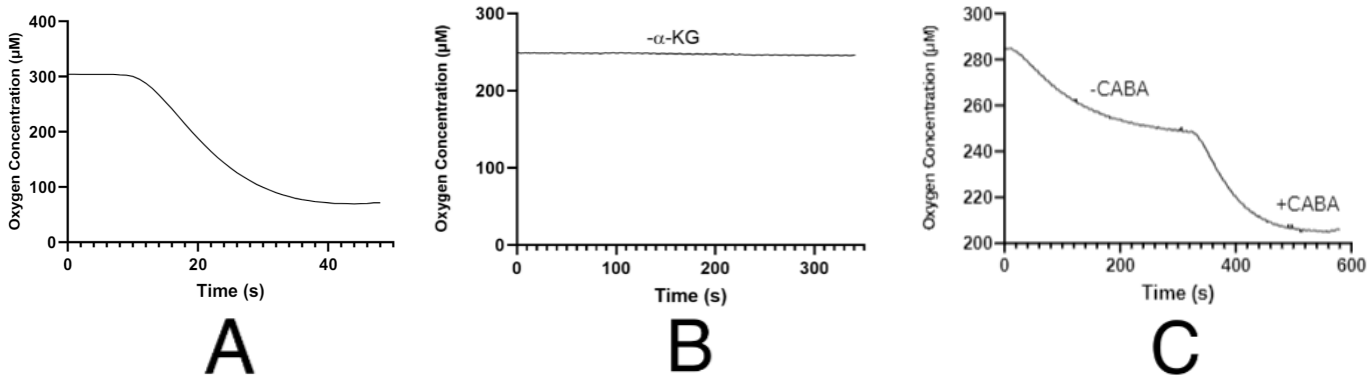


Figure A shows that oxygen is consumed by this reaction. **Figure B** is the control showing that in the absence of α-KG (and thus, only CABA is present), no oxygen is consumed. **Figure C** illustrates that with α-KG in the absence of CABA, oxygen is consumed and this consumption increases after flat-lining upon addition of CABA.

Loss of the carboxyl group on CABA

With isotope labeling and the use of GC-MS as well as LC-HRMS, the carboxylate group on CABA that was lost in this reaction was confirmed to be loss as CO_2 rather than CO or a formate adduct product.

Reduction of Oxygen

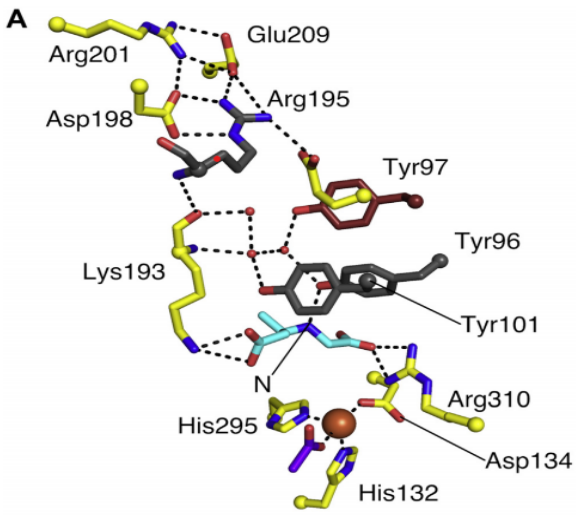
After first using a kit to detect H_2O_2 directly and using the oxygen probe on a coupled assay with an enzyme catalase (turns H_2O_2 into O_2), both produced negative results. This counteracts the

speculation of H_2O_2 being the product of the reduction of oxygen causing the speculation of H_2O to be more probable.

CABA binding site of ScoE

In the ScoE enzyme, one of the carboxylate moieties of substrate CABA is bound to the conserved arginine (Arg 310), while the other interacts with Lys 193. Immediately after CABA binding with Lys 193, a conformational change occurs at Lys 193, which flips the protein residue in and buries the CABA into the active site. In addition, the hydrogen bonding between Tyr 96 of ScoE and the secondary amine of the CABA molecule also helps to maintain the orientation and position of CABA in the binding site. The interaction between CABA and ScoE does not limit itself to the binding site but extends to the entire protein surface via a hydrogen bonding network.

Figure 2: Hydrogen bonding network in ScoE



This network involves hydrogen bonding among a series of amino acids at the surface of the protein, including Tyr 96, Tyr 97, Tyr 101, Lys 193, Arg 195, Asp 198, Glu 209, and Arg 201. The research has indicated that the substitution of these residues-- Tyr 96, Tyr 101, Arg 195, and Tyr 97-- will significantly compromise the hydrogen bonding network and cause no or reduced isonitrile formation. Thus, the ordered water network extending from the CABA binding site also plays an important catalytic role in isonitrile biosynthesis.

Fe (II) is bound at the conserved 2-His-1-Asp facial triad motif of the protein. During the interaction between CABA and Fe (II)=O species, the oxygen atom orientates away from the CABA binding site to form an off-line configuration stabilized by the hydrogen bonding with Arg 310. The off-line configuration provides protection for the highly reactive Fe (II)=O intermediate as the active site undergoes conformational changes required for reactions.

α-KG binding site of ScoE

Through the observation of electron density, the research has discovered that ScoE lacks a conserved arginine for the binding with α-KG, which gives α-KG binding site an off-site configuration. In addition, CABA binding sites and α-KG binding sites were observed to have inconsistent electron density, or different conformations to each other, indicating that these two binding sites are distinct and mutually exclusive.

The binding of α-KG involves two crucial residues: Arg 157 and His 299. A tartrate molecule was used for modeling α-KG binding. The tartrate did not interact with CABA binding site but induced conformational change in Arg 157 and His 299, which stabilized the binding. The replacement of tretrate with α-KG also generated the same observation. However, when tetrate was not present, conformational changes of Arg 157 and His 299 was not observed even in the

presence of CABA, suggesting that CABA did not induce response of this position. However, the research failed to visualize the Arg 157/His 299: α-KG interaction, so the assumption that these two residues are anchors for α-KG binding was unconfirmed. If the assumption is correct, then the substrate-enzyme interaction could be: first, one α-KG would bind to ScoE to form an offline, protected Fe(IV)=O intermediate; then, the binding between CABA and ScoE would cause conformational rearrangement and generate a bound CABA-hydroxylated intermediate, which would preclude further off-site α-KG binding due to steric hindrance; finally, Arg 157 and His 277 would move and form a new active site for the second α-KG binding.

Discussion

Previous studies have revealed a new isonitrile biosynthetic pathway catalyzed by mononuclear Fe(II) α-ketoglutarate-dependent dioxygenases, an enzyme superfamily previously not associated with isonitrile formation. Given the lack of knowledge about this pathway, it is important to investigate the general aspects of this enzyme-catalyzed reaction, including stoichiometry and crystal structures regarding the enzyme-substrate interactions. This research chooses the enzyme ScoE, a member of the mononuclear Fe(II) α-ketoglutarate-dependent dioxygenases superfamily, as research subject to study the mechanism of this newly discovered pathway.

An important missing data was the stoichiometry of the ScoE-catalyzed reaction. The result has shown a 1:1 ratio of substrate CABA/ co-substrate α-KG consumption versus INBA/ succinate formation, respectively (*Table 1*). Notably, the 1:2 ratio was observed for CABA versus α-KG consumption in general reaction (*Table 1*). This leads to the assumption that there is an alternative α-KG binding site on ScoE.

Then, the research turns to the structural analysis to understand the mechanism of ScoE-substrate interaction. For CABA binding site, the conserved arginine residue as well as the conserved 2-His-1-Asp facial triad motif for Fe (II) binding were observed (*Figure 2*). The research has also discovered a distinct α-KG binding site, which confirmed the assumption inspired by the stoichiometry analysis. Two residues-- Arg 157 and His 299-- were observed to play crucial roles in this active site formation.

In conclusion, the research studies both stoichiometry and crystal structure of the ScoE-catalyzed isonitrile biosynthesis. It provides knowledge to understand isonitrile production catalyzed by mononuclear Fe(II)/α-ketoglutarate (α-KG) dependent dioxygenase and insight to potential cure for pathogens that rely on this unprecedented pathway.

References

(1) Jonnalagadda, R.; Flores, A. D. R.; Cai, W.; Mehmood, R; Narayanamoorthy, M; Ren, C; Zaragoza, J. P. T.; Kulik, H. J.; Zhang, W; Drennan, C. L. Biochemical and crystallographic investigations into isonitrile formation by a nonheme iron-dependent oxidase/decarboxylase. *Elsevier Inc., ASBMB.* **2021**, 296.

(2) Martinez, S. Catalytic Mechanisms of Fe(II)- and 2-oxoglutarate-dependent oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702-20711.

(3) Huang, Y.; Cai, W.; Flores A. D. R.; Twigg, F.; Zhang, W. Facile Discovery of Isonitrile natural products *via* tetrazine based click reactions. *Anal. Chem.* **2019**, 92, 1, 599–602.

(4) Anandan, R.; Dharumadurai, D.; Manogaran, G. P.; An Introduction to Actinobacteria. *Actinobacteria - Basics and Biotechnological Applications*; Dhanasekaran, D.; Jiang, Y. IntechOpen; London, UK, 2016.

(5) Harris, N. C.; Born, D. A.; Cai, W.; Huang, Y.; Martin, J.; Khalaf, J.; Drennan, C. L.; Zhang, W. Isonitrile Formation by a Non-Heme Iron(II)-Dependent Oxidase/Decarboxylase. *Angewandte Chemie; International ed. in English.* **2018** 57, 31, 9707-9710.

(6) Oakley, B. R. *Aspergillus nidulans. Brenner's Encyclopedia of Genetics. Elsevier Inc.* 2016

(7) Harris, N. C.; Sato, M.; Herman, N. A.; Twigg, F.: Cai, W; Liu, J.; Zhu X.; Downey, J.; Khalaf, R; Martin, J.; Koshino, H.; Zhang, W. Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria. *PNAS.* **2017**, 114, 27, 7025-7030.

(8) Knauer, S. H.; Hartl-Spiegelhauer, O.; Schwarzinger, S.; Hänzelmann, P.; Dobbek, H. The Fe(II)/α‑ketoglutarate‑dependent taurine dioxygenases from *Pseudomonas putida* and *Escherichia coli* are tetramers. *The FEBS Journal.* **2012**, 279, 5, 816-831.