Studies on the Mechanism of Ethylene-f

Enzymes and their Potential for Bioca

A thesis submitted to the Board of the Faculty of the Mathematica
Sciences Division for the degree of

Master of Science (Research) in **Organic Ch**

Yihong SUN

DECLARATION

I declare that the work described, and the figures presented in this thesis are entirely my own, except where I have acknowledged help and/or data from a named person or given reference to a published source.

Yihong Sun

St. Hilda’s College, Department of Chemistry, University of Oxford

Trinity Term 2025

ACKNOWLEDGEMENTS

I would like to thank all the people who have contributed in various ways to my study at Oxford.

I would like to express heartfelt gratitude to my supervisor Prof. Christopher Schofield for having me for my Master's degree and the fantastic project. I am also grateful to my mentors, Dr. Lennart Brewitz and Mr. Siddhant Dhingra for their supervision and support on the project, especially on so many late nights and weekends. I would also like to thank my collaborators, Dr. Zhihong Zhang, Dr. Xin Wen, and Dr. Mark D. Allen for their support and dedication to the progress of the project. My gratitude also extends to Dr. Coral Mycroft from the Chemistry NMR team and Prof. Christina Redfield from the Department of Biochemistry for their technical support. I would also like to thank my parents (Ms. Beibei Liang & Mr. Jingkang Li) and my brother (Mr. Adam Onyu Li) for their supports through one of the most important chapters in my life. Meanwhile, I would also like to extend my heartfelt gratitude to my best friend (Miss Jialin Li) for her unwavering support. Lastly, I would like to thank all my family members and friends for their unwavering support to my study. It would not have been possible without their understanding.

TABLE OF CONTENTS

Studies on the Mechanism of Ethylene-forming Enzymes and their Potential for Biocatalysis

ABSTRACT

The *Amborella trichopoda* 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO) and the *Pseudomonas syringae* ethylene-forming enzyme (PsEFE) are Fe(II)-dependent O_2 utilising enzymes that catalyse the formation of ethylene, an important industrial commodity and biological signalling molecule. This thesis describes biochemical and crystallographic studies on the mechanisms, structures, and substrate scopes of these two ethylene forming enzymes.

ACCO plays a crucial regulatory role in the growth and development of higher plants. It catalyses ethylene formation from ACC and O_2 to give CO_2, and HCN as by-products, and requires both bicarbonate and L-ascorbate for activity. *A. trichopoda* is the earliest plant to have evolved separately from the other flowing plants, making it

of interest in understanding the evolutionary pattern of flowering plants. ^{1}H NMR-based biochemical assays coupled with mutagenesis and substrate analogue studies were used to investigate the *A. trichopoda* ACCO (AtACCO) reaction. The results showed that K159, K290, S247, R245, Y91, and V185 are key residues for AtACCO catalysis, in accordance with reported mutagenesis studies on ACCOs of different organisms. Protein crystallography revealed that K159, K290, and Y91 jointly formed a network to assist the binding of bicarbonate, crucial for ACCO catalysis. Disrupting the K159, K290, and Y91 network hampered catalysis, reflecting the importance of bicarbonate for productive turnover.

Structure-guided protein engineering of AtACCO was employed to alter its selectivity, enabling efficient production of olefins other than ethylene from C-alkylated ACC derivatives, including propylene formation from the natural product norcoronamic acid. Notably, the S247G AtACCO variant catalysed propylene formation from norcoronamic acid 3-fold more efficiently than wildtype (wt) AtACCO. The results suggests that gaseous olefins other than ethylene may have roles in plant signalling.

Pseudomonas syringae PsEFE is an ethylene-forming enzyme that catalyses the formation of ethylene from 2-oxoglutarate (2OG) in a L-arginine-dependent bifurcating reaction. Previous work on the reactivity of PsEFE with 2-oxoglutarate derivatives revealed that PsEFE does not catalyse alkene formation from C3/C4-substituted 2OG derivatives; however, it remained unclear whether this observation

reflected different 2OG derivative binding modes. Efforts to crystallize PsEFE in complex with 2OG derivatives was unsuccessful. Therefore, as part of this thesis, a novel ethylene forming enyzme from 1449B *Pseudomonas syringae* species was produced, that differed in 41 amino acids from the previously used PsEFE, mostly locating on the 1449B PsEFE surface according to modelling studies. ^{1}H NMR studies revealed that 1449B PsEFE indeed catalyses ethylene formation from 2OG. Notably, co-crystallisation of 1449B PsEFE with C3/C4-substituted 2OG derivatives afforded high resolution structures that reveal an unusual 2OG/2OG derivative binding mode in the absence of L-arginine, i.e., the 2-oxoacid motif of 2OG did not chelate the metal co-factor but interacted with the guanidinium group of R277 which would typically binding to the 2OG C5 carboxylate. Such 2OG binding modes might contribute to atypical 2OG oxygenase reactivity.

The overall results provide new insight into the mechanisms and factors determining substrate and product selectivity by two mechanistically distinct but related ethylene forming enzymes. They will enable future studies aimed at optimising the biological production of ethylene and higher olefins.

ABBREVIATIONS

2OG	2-Oxoglutarate
ACC	1-Aminocyclopropane-1-carboxylic acid
ACCO	1-Aminocyclopropane-1-carboxylic acid oxidase
L-arg	L-arginine
DMSP	Dimethylsulphoniopropionate
DSBH	Double-stranded β-helix
EFE	Ethylene-forming enzyme
efe	Gene encoding for PsEFE
KMBA	2-Keto-4-methylthiobutyric acid
L-asc	L-ascorbate
P5C	*L*-4[1]-Pyrroline-5-carboxylic acid
PsEFE	*Pseudomonas syringae* pv. *phaseolicola* ethylene-forming enzyme
SAM	*S*-Adenosyl-*L*-methionine
TMSP-d_4	Trimethylsilyl propionic-2,2,3,3-d_4 acid sodium salt
WT	Wildtype
2xYT	2-Yeast Tryptone extract medium
NMR	Nuclear magnetic resonance
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
PDB	Protein data bank
SAM	S-Adenosylmethionine

MTA Methylthioadenosine

AMC 1-Aminomethylcyclopropanecarboxylic acid

AEC 1-Aminoethylcyclopropanecarboxylic acid

FPLC Fast protein liquid chromatography

rpm Revolutions per minute

RT Room temperature

IPTG Isopropyl ß-D-1-thiogalactopyranoside

MS Mass spectrometry

Tris Tris(hydroxymethyl)aminomethane

ESI. Electrospray Ionisation

GC Gas Chromatography

Amino acid residues are abbreviated in general by conventional 3 letter codes (Arg = arginine, etc.), and for PsEFE mutation sites by conventional 1 letter codes (F = phenylalanine, etc

Chapter 1. Introduction

1.1 Industrial gaseous olefins

Ethylene is a vital building block for the production of multiple chemicals that are vital to modern society, ranging from plastics to antifreeze and solvents[1]. The Organisation of Economic Cooperation and Development (OECD) has reported that the annual global ethylene production exceeds 100 million tons[2]. Currently, ethylene is mainly produced from cracking larger hydrocarbons from petroleum[1]. Petroleum cracking is also the sole source for production of gaseous olefins other than ethylene, including propene and 1-butene[3]. Propene and 1-butene are important precursors for the production of polypropylene, acrylonitrile, and butyl rubber[4,5].

Alternative sources for ethylene formation are being sought, primarily because of the emittance of greenhouse gases associated with its production from non-renewable sources. For example, ethylene can be produced by dehydration of ethanol, that can be obtained from fermentation of biomass[6,7]. However, this process is not yet economically viable.

1.2 The biological significance of ethylene

1.2.1 Ethylene in plants

Ethylene is not only industrially important, but also plays a crucial role in biology. It is a key growth and developmental regulator in higher plants, playing important roles in seed germination, fruit abscission, flower senescence, and other processes[8].

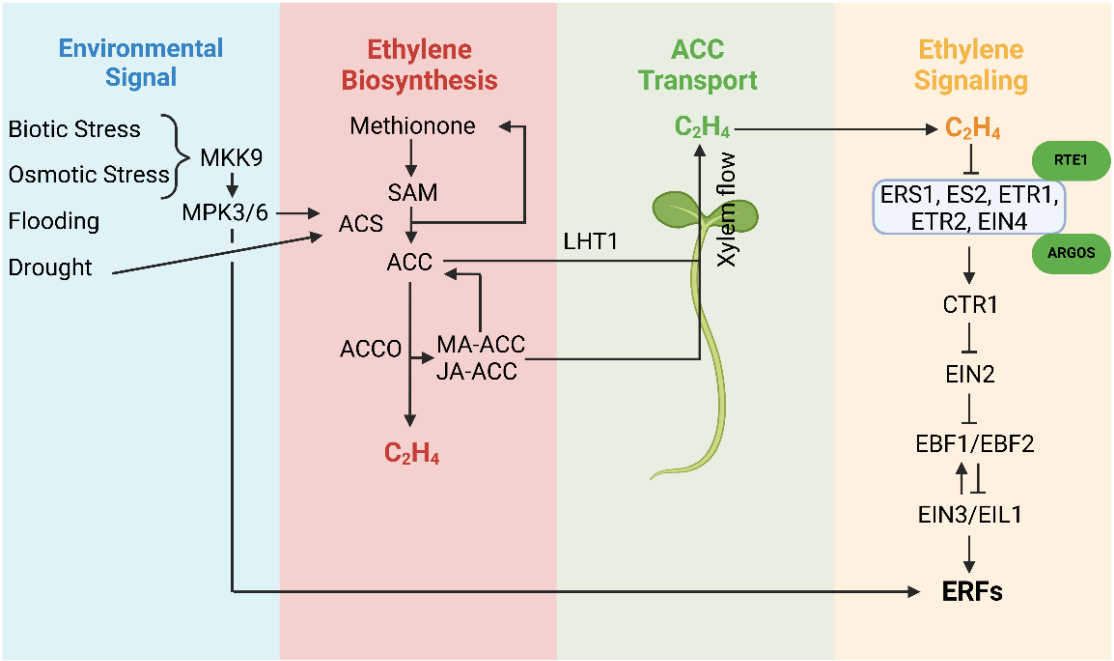


Figure 1.1 Summary of the processes involved in ethylene signaling pathways in higher plants. SAM: S-adenosyl methionine; ACS: ACC Synthase; LHT1: Lysine Histidine Transporter 1; ERS: ethylene response sensor; ETR: ethylene resistance; EIN: ethylene insensitive; RTE: reversion to ethylene sensitivity; CTR: constituitive triple response; ERF: ethylene response factor; MA-ACC: Malonyl-ACC; JA-ACC: jasmonoyl-ACC.

In seeding plants, ethylene is produced in response to both abiotic and biotic stresses,

including flooding, drought and shaded environment[8]. These environmental signals affect the expression levels of 1-aminocyclopropanecarboxylic acid (ACC) synthase (ACS) either directly or indirectly via MKK9 and MPK3/6[8]. ACS is responsible for catalysing the production of 1-aminocyclopropanecarboxylic acid (ACC); hence, ACS expression levels modulate ACC levels in plant cells. (Fig. 1.1) ACC is taken into cells by the amino acid transporter LHT1 (Lysine Histidine Transporter 1) and transported through the plant via xylem vessels[9]. In destination organs, ACC is converted into ethylene by ACC oxidase (ACCO). Beside enzymatic oxidation, ACC can also undergo conjugation to form malonyl-ACC (MA-ACC), γ-glutamyl-ACC (G-ACC) and jasmonoyl-ACC (JA-ACC) catalysed by ACC-*N*-malonyl transferase (AMT), γ-glutamyl-transpeptidase (GGT) and jasmonic acid resistance 1 (JAR1), respectively. These ACC conjugation processes contribute to regulating ACC levels and thus ethylene formation (Fig. 1.1).[10] Note that ACC is also reported to have signalling roles that are independent of ethylene[11].

Ethylene biosynthesis in seeding plants occurs in a three-step process from L-methionine as the primary metabolite precursor (Fig. 1.2-1.3a). Initially, L-methionine is converted into S-adenosyl methionine (SAM) by S-adenosyl methionine synthase (SAMS). 1-Aminocyclopropane-1-carboxylic acid synthase (ACS) then converts SAM into ACC giving methylthioadenosine (MTA) as a by-product. MTA can be recycled to methionine via the Yang cycle. (Fig. 1.2) As the levels of ACS are strictly regulated in cells, ACC synthesis is the rate-limiting step in ethylene biosynthesis.

Note that ACCO catalysis produces not only ethylene from ACC, but also CO_2 and cyanide, or cyanoformate ($NCCO_2H$) which rapidly decomposes into CO_2 and cyanide. Toxic cyanide is subsequently metabolised by ß-cyanoalanine synthase to give ß-cyanoalanine, which can be hydrolysed to asparagine [12].

Ethylene production triggers a signaling cascade in seeding plants, initiated by ethylene receptors localised to the endoplasmic reticulum (ER) membrane, including ETR1 (ethylene resistance 1), ERS1 (ethylene response sensor 1), ETR2, ERS2, and EIN4 (ethylene insensitive 4)) (Fig. 1.1). These receptors are active in the absence of ethylene by binding to and thus activating **CTR1** (constitutive triple response 1)[13], a kinase that represses the ER-located membrane protein EIN2. The presence of ethylene indirectly inhibits CTR1, inducing dephosphorylation of EIN2, which results in EIN2 cleavage to release a C-terminal fragment. This C-terminal EIN2 fragment is involved in gene-specific regulation of translation by binding to 3′-untranslated regions (3′-UTRs) of EBF1 and EBF2 mRNA transcripts and suppressing their translation[14]. EBF1 and EBF2 are two central F-box proteins that target the main ethylene-responsive transcription factors EIN3 and EIN3-LIKE1 (EIL1) responsible for protein degradation in the absence of ethylene[15,16]. In the presence of ethylene, EIN3 and EIN1 induce expression of ethylene-response factors (ERFs)[17]. (Fig. 1.1) The activity of some ERFs can also be induced by phosphorylation through the MPK3/6-cascade that also regulates ethylene biosynthesis, providing dual-level regulation of the ERF-mediated response[18,19] (Fig. 1.1).

Plant ethylene levels can also be regulated by external factors. For example, plant growth-promoting bacteria can reduce plant ACC levels by expressing ACC deaminase, which converts ACC to ammonia and α-ketobutyrate in the root environment. It is proposed that root exudates contain certain amounts of ACC which might attract these bacteria to establish the rhizophere interaction[20], with ammonia acting as a bacterial nitrogen supply[21].

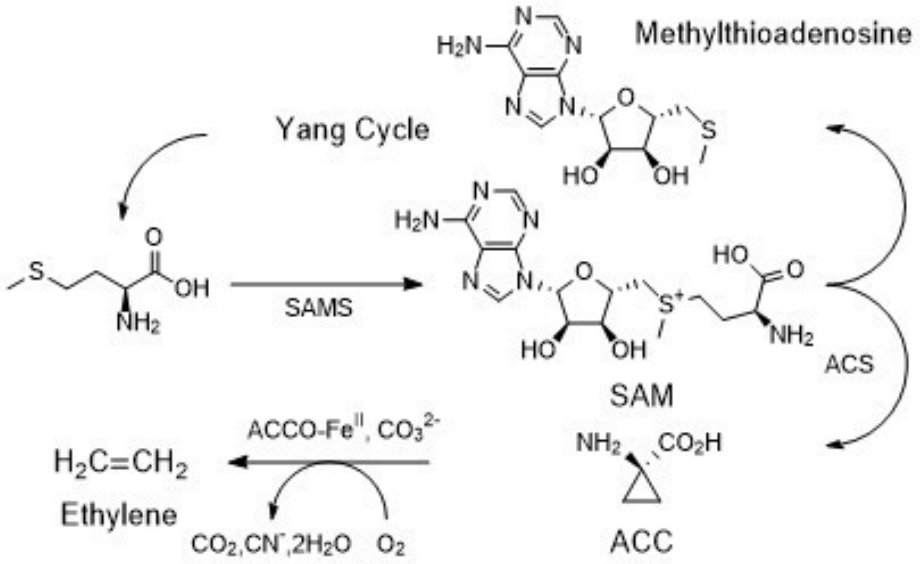


Figure 1.2 Biosynthetic process of ethylene in plants.

Despite the presence of ethylene receptors in lower plants, they do not appear to possess ACCOs, indicating the potential existence of alternative ethylene biosynthetic pathways[22]. In line with this proposal, it has been reported that ethylene can be produced from dimethylsulphoniopropionate (DMSP) in certain marine algae. (Fig.

1.3b) In this two-step pathway, DMSP is firstly turned over by DMSP lyase to form acrylate, which is then decarboxylated by acrylate decarboxylase to produce ethylene[23].

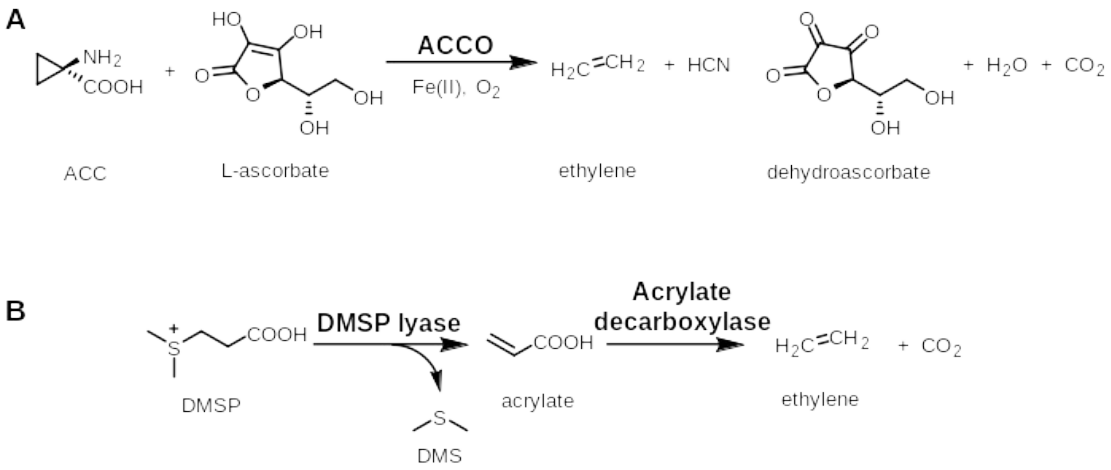


Figure 1.3 Biosynthesis of ethylene in A) seeding plants and B) marine algae[23].

1.2.2 Ethylene in Bacteria

Ethylene is produced in organisms other than plants and algae, including in bacteria and fungi[24]. However, the biological significance of ethylene signalling in bacteria is not well understood. Putative ethylene receptors have been discovered in many bacteria, including in the cyanobacterium *Synechocystis (Synechocystis sp. PCC 6803)*[25]. Ethylene is proposed to stimulate a putative plant-like ETR, slr1212, which contains a GAF domain and a histidine kinase domain similar to its counterpart in plant ETRs[26]. This process triggers the response regulator, slr1213, and subsequent

ethylene responses[26]. However, this pathway is not well understood. Ethylene production has been linked to bacteria-induced plant diseases.

Recent studies have identified plant ETR homologs in more than 100 proteobacteria and cyanobacteria[24,27]. The prokaryotic ETRs share similarities with the plant ETRs in terms of the domain arrangements, including the presence of a GAF domain and a histidine kinase-like domain[27]. Notably, a number of bacterial species show either a different domain configuration, or alternative output domains, suggesting that some bacterial ETRs may integrate multiple input signals and have diverse biochemical outputs. ETR homologs have been found in various non-plant eukaryotic clades where the genomes of a number of early diverging fungal species encodes proteins that resemble the ETRs of higher plants[28]. As these fungi are plant symbionts, it is possible that ethylene is involved in symbiotic communication.

In *Pseudomonas syringae* and *Penicillium digitatum*, ethylene is generated via a different pathway than in seeding plants and algae. It is produced from 2-oxoglutarate (2OG) by ethylene – forming enzymes (EFEs), which are, like ACCOs, Fe(II)-dependent cupin-fold oxygenases (Fig. 1.4a). On the other hand, in a number of aerobic bacteria including *E. coli* and *Cryptococcus albidus*, ethylene has been reported to be produced from the oxidation of 2-keto-4-methylthiobutyric acid (KMBA)[29], the transaminated derivative of methionine, possibly catalysed by a yet uncharacterized NADH: Fe(III) EDTA oxidoreductase[30,31]. (Fig. 1.4c)

A biological source of ethylene associated with anaerobic metabolism has also been reported[32]. A number of terrestrial and freshwater bacteria harness a methionine salvage pathway to recycle the amino acid from the toxic, SAM-derived 5´-methylthioadenosine. A nitrogenase-like reductase (MarBHDK) functions to reduce the (2-methylthio)ethanol intermediate to give methanethiol, ethylene, and water under stress[33]. (Fig. 1.4b)

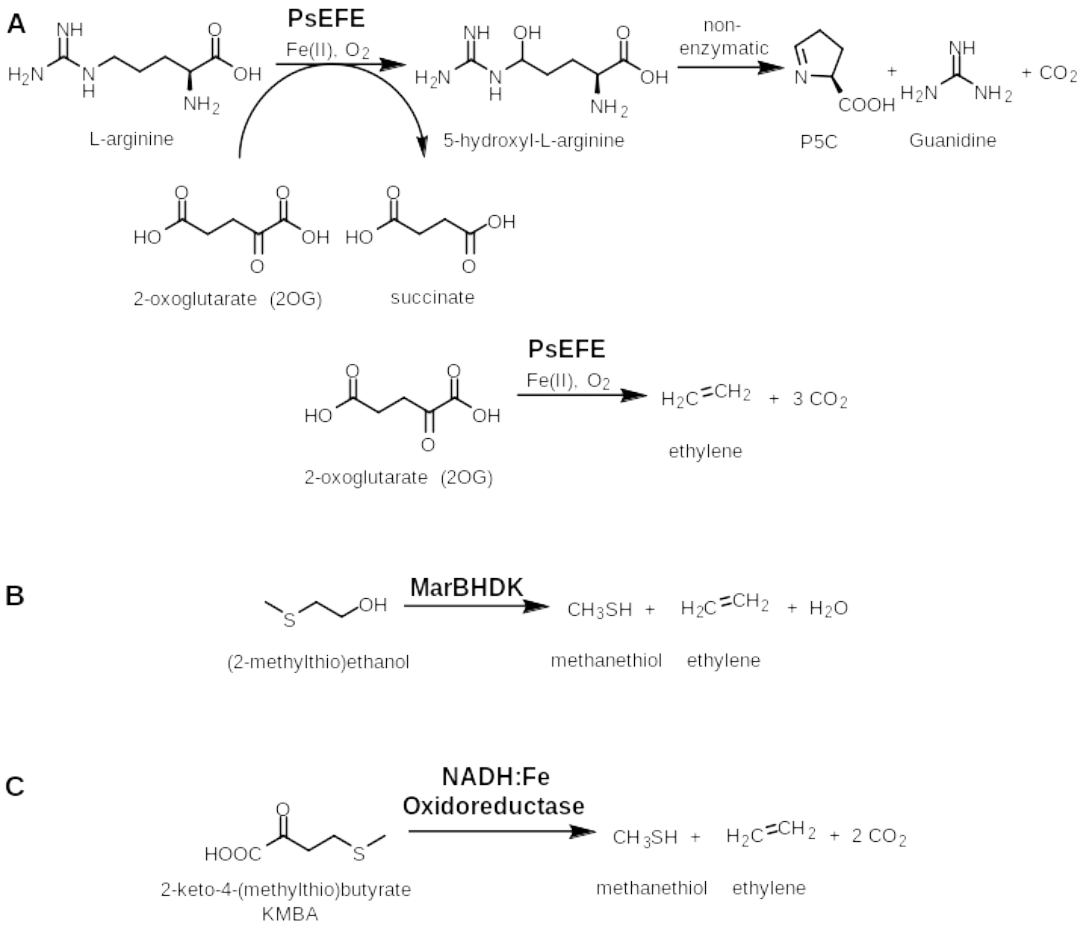


Figure 1.4 Ethylene biosynthesis in bacteria in a) *Pseudomonas syringae* and *Penicillium digitatum*; b) terrestrial and freshwater bacteria; c) anaerobic bacteria.[32,33]

1.3 Fe(II)-dependent cupin fold oxygenases

Fe(II)-dependent cupin fold oxygenases contain a distorted double-stranded ß-helix (DSBH) core fold[34]. The DSBH core fold is comprised of eight strands[35–39] (Fig. 1.5), arranged to form a barrel-like structure which supports the catalytic site. The accessible part of the barrel contains the Fe(II) and substrate binding residues. In many, but not all, of these oxygenases, Fe(II) is bound by three residues forming a highly conserved HX(D/E)…H triad. Many, but not all Fe(II)-dependent cupin fold oxygenases, employ 2OG as a co-substrate, which coordinates to Fe(II) in a bidentate manner via one of its C1 carboxylate and its ketone oxygen.[34] The 2OG C5 carboxylate is usually bound via an electrostatic interaction with a basic residue, frequently either a lysine or an arginine (as described in Chapter 1.4.2). From a structural perspective, the *Pseudomonas syringae* ethylene-forming enzyme is a representative example of an Fe(II) and 2OG dependent oxygenase though it has unusual features. On the other hand, some Fe(II)-dependent oxidising enzymes, such as ACCOs, do not employ 2OG as a co-substrate.

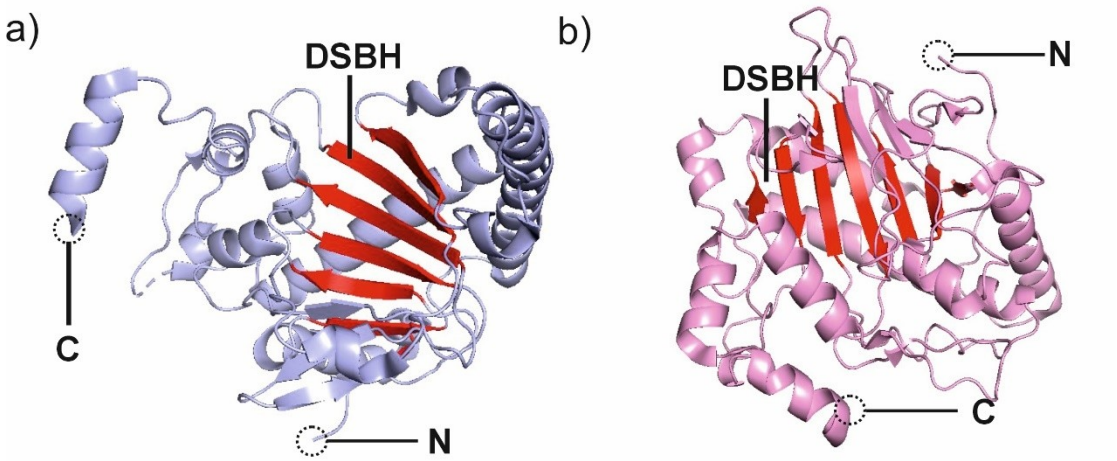


Figure 1.5 The double-stranded β-helix (DSBH) core fold of the 2OG oxygenase superfamily. Overall structural views of: (**A**) 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO) (PDB: 1W9Y; Zhang *et al.*, 2004)) and (**B**) *Pseudomonas syringae* ethylene-forming enzyme (PsEFE) (PDB ID: 5V2X; Martinez *et al*, 2017). The characteristic DSBH fold of the 2OG oxygenase superfamily to which ACCO and PsEFE belongs is highlighted in red. N-terminus (N) and C-terminus (C) are highlighted with dotted circles.

1.4 Ethylene-forming enzymes

1.4.1 The Plant Ethylene-forming enzyme

1-Aminocyclopropane-1-carboxylic acid (ACC) oxidase (ACCO) is responsible for ethylene formations in higher plant cells. The exact location of ACCO remains controversial. Some studies suggests that ACCO may locate to the cytosol whereas other work suggests ACCO may be located in the plasma membrane[40]. ACCO requires L-ascorbate for activity *in vitro*, possibly as a reductant which is oxidised into L-dehydroascorbate. However, the exact role of L-ascorbate in the ACCO catalysed ACC oxidation, and more generally in 2OG oxygenase catalysis, is unclear[41].

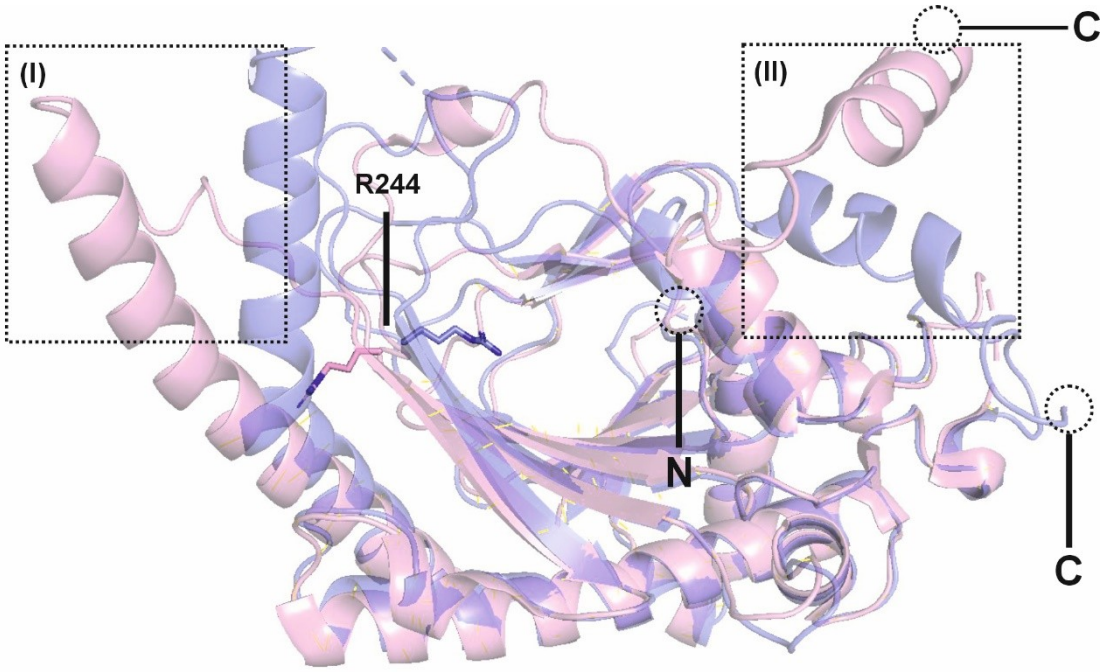


Figure 1.6 Structural superimposition of *Petunia* ACCO (PDB: 1WA6) with *Amborella* ACCO crystal structure derived views (unpublished data, Dr. Zhihong Zhang, University of Oxford), RMSD = 0.415 (Å). Structural differences are highlighted in black boxes. The N-terminus (N) and C-terminus (C-) are highlighted with circles. Structural divergences between the two ACCO orthologs are highlighted by the dotted boxes as (i) and (ii).

Previous structural research on ACCO has been centred on the *Arabidopsis thaliana* and *Petunia hybrida* ACCOs[42]. Recent unpublished work by Dr. Zhihong Zhang (University of Oxford) has focused on the *Amborella trichopoda* ACCO. *Amborella trichopoda* is a monotypic genus of the Amborellaceae family and the order Amborellales. *Amborella trichopoda* is the earliest group to have evolved separately from the other flowing plants, making it of interest in understanding the evolutionary pattern of flowering plants. As a result, the work discussed in the Chapter 2 and Chapter 3 focused on *Amborella trichopoda* ACCO (AtACCO).

The main chain of AtACCO contains eleven α helices and thirteen ß strands. (Fig. 1.5a) Eight of these strands form the distorted double-stranded ß helix (DSBH) core fold which is common to all identified members of the 2OG oxygenase structural superfamily[43] (Fig. 1.5a). This DSBH motif is also present in the broad cupin family of plant proteins and the 2OG-dependent JmjC (jumonji C) histone demethylases[44]. Comparing *Petunia hybrida* ACCO and AtACCO, their structures reveals that the two proteins share a high degree of structural similarity except a significant divergence for the C-terminal α-helix and a long α-helix on the surface of the DSBH (Fig. 1.6). Interestingly, R245 of *Petunia hybrida* ACCO is not present at the active site in the absence of ACC, but its side chain points away from the catalytic pocket of the ACCO binding site. However, in the AtACCO:ACC complex structure, the side chain of R245 faces into the catalytic pocket to assist in binding of ACC (Fig. 1.6)

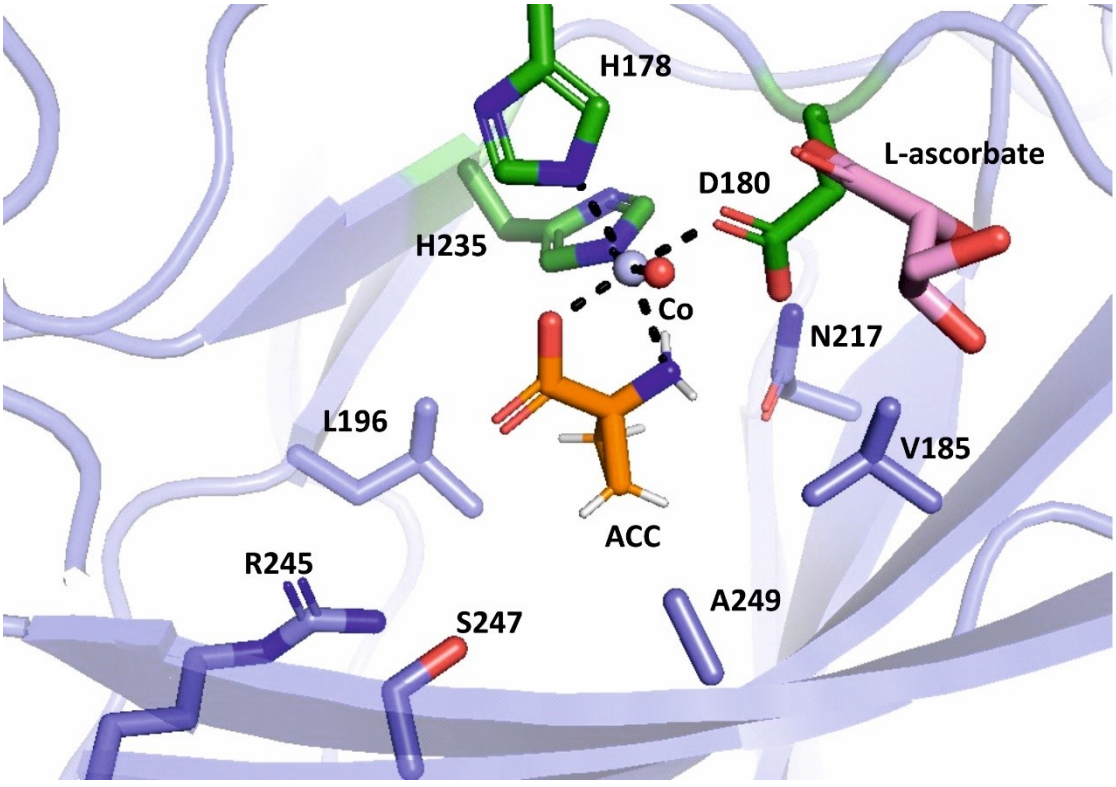


Figure 1.7 View of ACC binding at the AtACCO active site. The amino acids surrounding the substrate are shown in slate. The substrate is highlighted in orange. L-ascorbate is highlighted in pink. Residues coordinating to the metal are shown in green. The metal is shown as a light blue sphere. Water is shown in red spheres.

Crystal structures of AtACCO in complex with ACC, Co(II), and L-ascorbate or bicarbonate (unpublished data by Dr. Zhihong Zhang, University of Oxford) reveal that the metal cofactor is coordinated by a triad of residues (H178, D180 and H235)[42] (Fig. 1.7). The ACC cyclopropane ring is located in a pocket formed by S247, L196, N217 and V185 in which the 4 C-H bonds on the ACC ring are directed towards. The L-ascorbate is located at the entrance of the binding pocket.

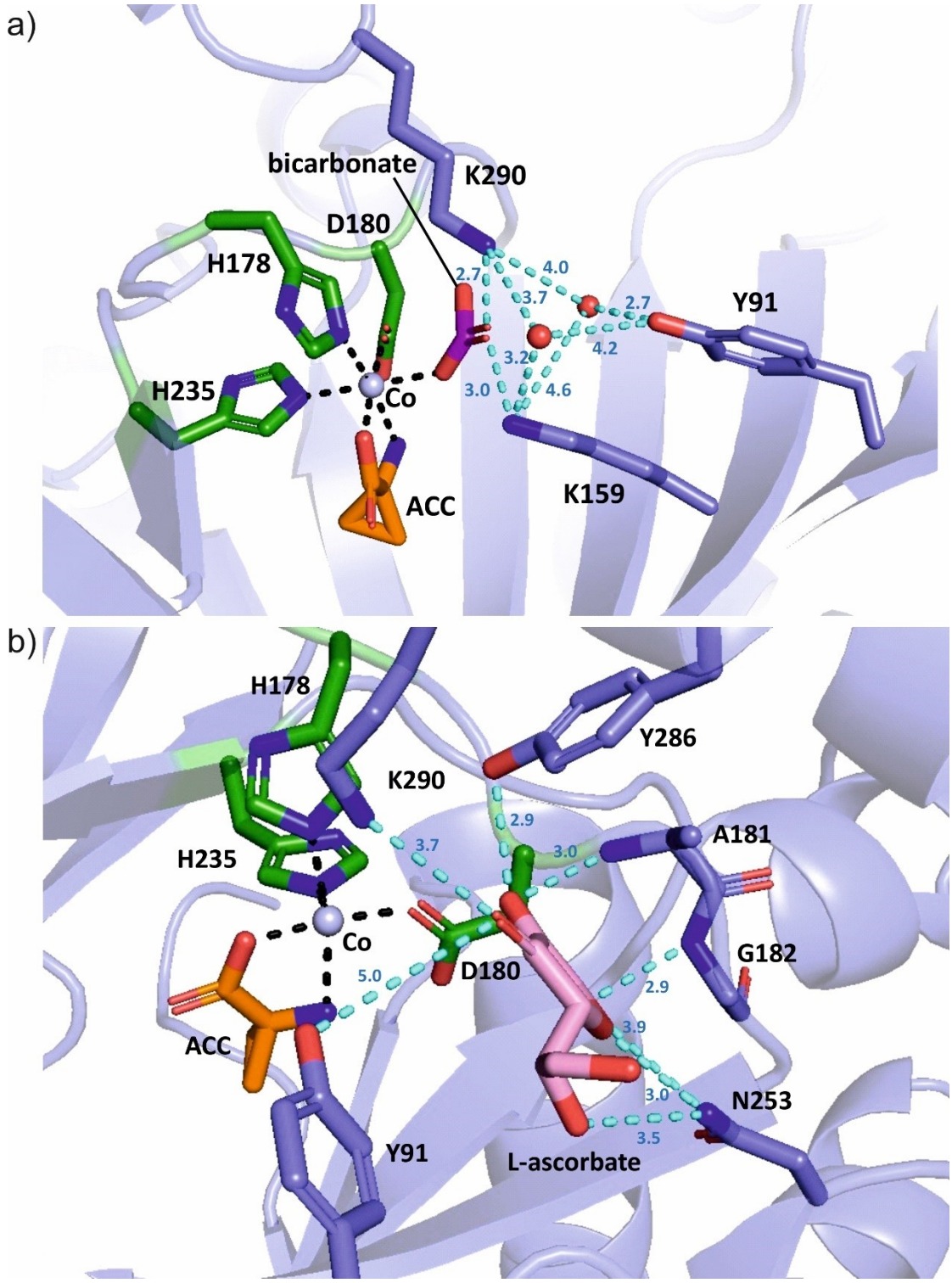


Figure 1.8 Mode of a) L-ascorbate and b) bicarbonate binding in the ACCO binding pocket. The hydrogen bonding interactions in coordinating the co-factors are highlighted in cyan dashed lines with the bonding distance shown in blue (unit: Å). Coordinating groups to the metal centre are shown in black dashed lines. Water is shown in red spheres.

The locations of L-ascorbate and bicarbonate binding to ACCO were identified with

the help of AtACCO crystal complex structures determined by Dr. Zhihong Zhang (University of Oxford). Bicarbonate binds to the sixth coordination site on the metal centre as shown in an ACCO crystal structure without L-ascorbate, also obtained by Dr. Zhihong Zhang (Fig. 1.8b). A co-crystal structure of AtACCO with Co(II), ACC and bicarbonate suggests that bicarbonate is coordinated by the side chains of K159 and K290. The conformations of K159 and K290 are apparently stabilised by a hydrogen bond network involving two water molecules and Y91. (Fig. 1.8a) L-ascorbate is positioned to interact with the side chains of N253 and K290, as well as the main chain nitrogens of A181 and G182 via hydrogen bonding interactions (Fig. 1.8b).

Previous mutagenesis studies on *Malus domestica* (apple) ACCO1 (MdACCO1) reported by Dilley and coworkers, investigated key MdACCO1 residues required for catalysis[45]. MdACCO1 residues T157, K158, R175, R244, and S246 (equivalent to T158, K159, R176, R245, and S247 in AtACCO) were shown to be important for activity. Variations in these residues differentially affected MdACCO1 catalytic activity: R175, R244 and S246 were determined to be important in affecting ACC binding. Residues bearing positive charges were hypothesised to potentially interact with bicarbonate. Mutagenesis study on these residues discovered that K158E, K199E, K230E, and K292E were the least active variants of those made and tested.

Although the mechanism of ACCO has not unequivocally been experimentally

validated, mechanistic proposals have been made based on reported biochemical and computational studies[46] (Fig. 1.9): The Fe(II) ion in ACCO is initially bound by three water molecules (A). The binding of ACC displaces two of these water molecules (B). ACC binding is followed by a monodentate binding of an O_2 molecule with the formation of an oxygen radical (superoxide) and oxidation of Fe(II) to Fe(III) (C). Formation of an Fe-linked peroxide (C-D) followed by its cleavage enables formation of an Fe(IV) species (D-E). The transfer of a hydrogen radical reduces Fe(IV) to Fe(III) and initiates ethylene formation (possibly via a radical pathway, F-G). The removal of ethylene from ACC leads to the formation of a cyanoformate, which remains bound to the metal centre (H). Cyanoformate then undergoes a decarboxylation to generate carbon dioxide and cyanide ion as the by-products which restores the unbound Fe (II), potentially due to the bond strain at the triply bonded carbon (A).

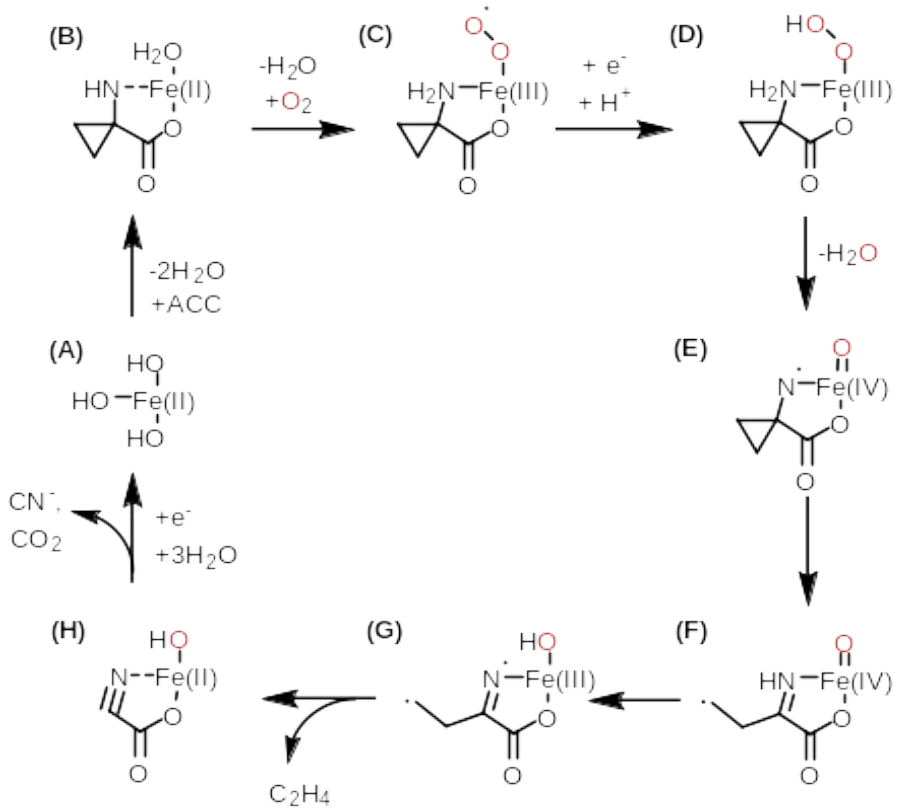


Figure 1.9 Proposed mechanism for ACCO catalysis. The oxygen atom from the atmospheric oxygen is highlighted in red.REF

1.4.2 Bacterial Ethylene-forming enzymes

Fe(II)-dependent cupinfold EFEs are also found in bacteria and fungi, including *Pseudomonas savastanoi* EFE (PsEFE) and *Penicillium digitatum*[47]. These ethylene forming enzymes require L-arginine and 2OG for catalysis. PsEFE has been found to possess a bifurcating reaction pathway (Fig. 1.10). The major reaction pathway utilises 2OG, L-arginine and O_2 to produce succinate and 5-hydroxyarginine via oxidative decarboxylation. 5-Hydroxyarginine then fragments to give guanidine and L-Δ^1-pyrroline-5-carboxylate (P5C). The concomitant pathway gives ethylene, CO_2,

and H_2O.

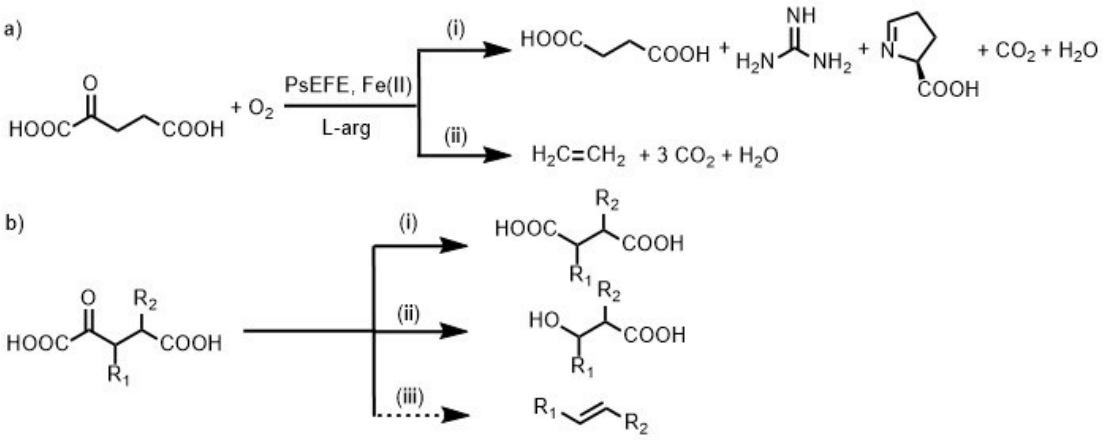


Figure 1.10 Reaction outcomes of PsEFE catalysis on a) 2OG and b) 2OG derivatives[48]

Previous studies have shown that PsEFE catalyses the conversion of some 2OG derivatives into alcohol (ß-hydroxypropionate) derivatives in addition to forming succinate (diacid) derivatives[48]. In contrast to the oxidation of 2OG, PsEFE catalysis with 2OG derivatives do not produce detectable levels of alkene products.[48] The steric bulk of the 2OG derivative substituents appears to dictate the reaction outcomes. Thus, sterically bulky substituents at the 2OG C3 and C4 position are proposed to favour the formation of the diacid product over the corresponding alcohol product[48]. Currently, there is no structural information to substantiate this proposal.

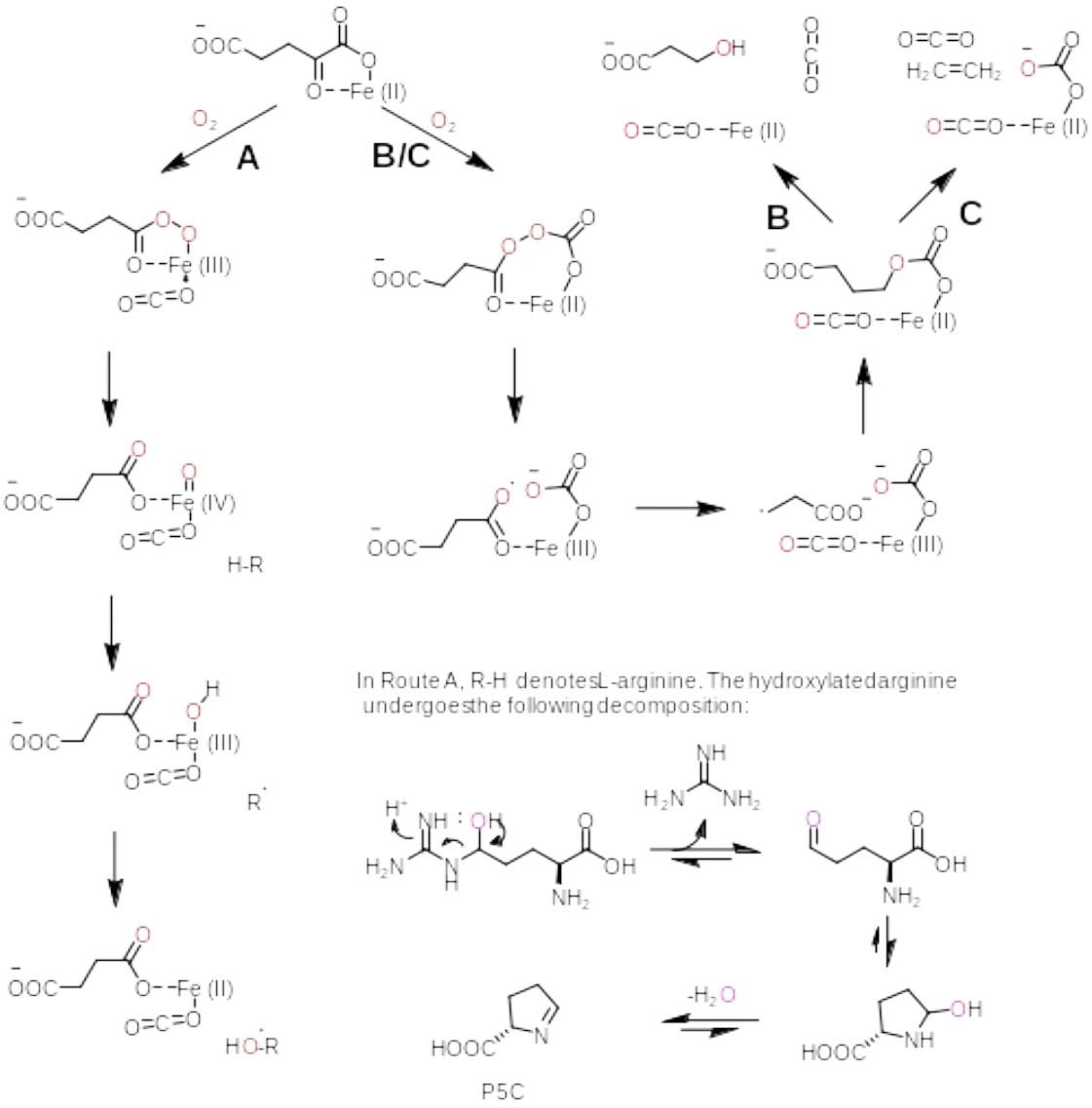


Figure 1.11 Proposed outline mechanism for PsEFE catalysis. A: Succinate Formation Route; B: Alcohol Formation route; C: Ethylene Formation Route.

An outline PsEFE mechanism has been proposed based on experimental studies and computational simulations[46] (Fig. 1.11). In the oxidative decarboxylation pathway, the Fe(II)-coordinated 2OG undergoes a reaction with O_2, forming an Fe(III)-superoxo complex, followed by decarboxylation and formation of an Fe(IV)=O intermediate. The latter abstracts a C5 hydrogen from L-arginine to form an Fe(III)-OH

intermediate. Radical rebound then occurs to form C-5 hydroxylated L-arginine, together with formation of succinate and the reduction of Fe(III) to Fe(II). The C-5 hydroxylated L-arginine is an unstable species that readily decomposes to give into guanidine and P5C[46]. In the ethylene biosynthesis pathway, O_2 is formally inserted in between the 2OG C1 and C2 bonds. The cleavage of the O-O bond releases a radical species that undergoes further decarboxylation to produce ethylene or an alcohol. (Fig.1.11)

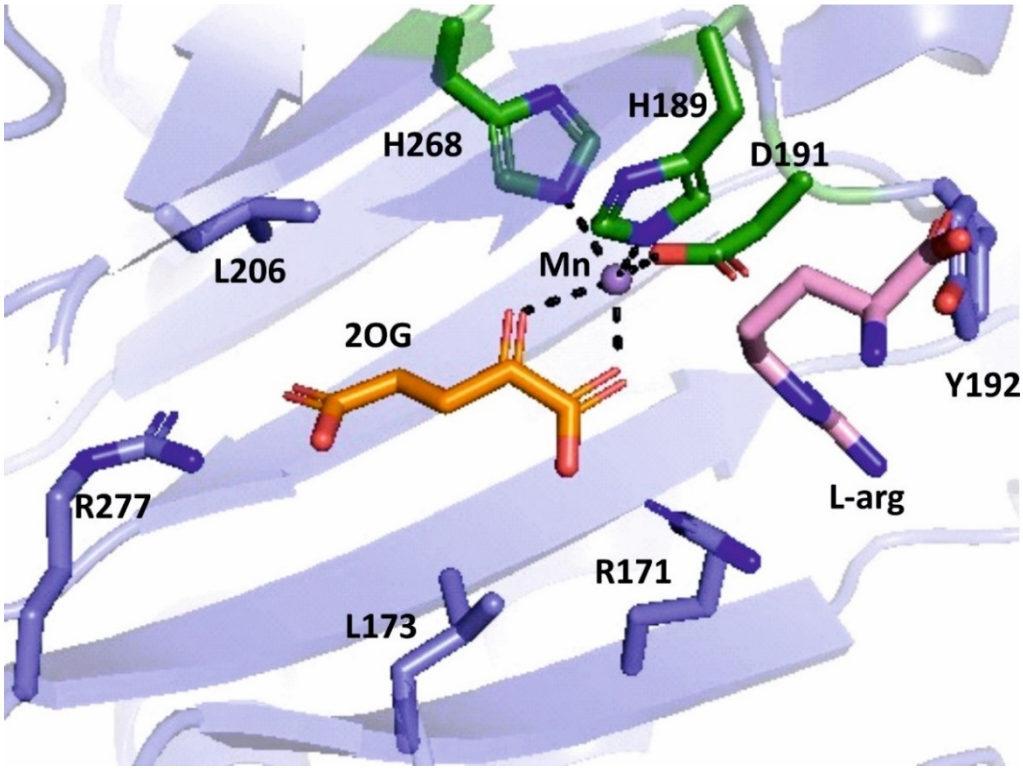


Figure 1.12 View from a Structure of PsEFE co-crystallised with 2OG, Mn(II) and L-arginine (PDB: 5V2Y). The amino acids in the catalytic pocket are shown in slate. L-arginine is shown in pink. 2OG is highlighted in orange. The residues coordinating to the metal centre are shown in green. The metal centre is shown in light blue sphere. The coordinations to the metal centre are shown in black dashed lines.

Crystal structures of the archetype ethylene-forming enzyme (EFE) of *Pseudomonas*

syringae pv. *Phaseolicola* PK2 reveal that PsEFE employs a triad of H189, D191 and H268 residues to coordinate the Fe(II) at the active site, in a similar manner as occurs with ACCOs[46] (PDB: 5V2Y) (Fig. 1.12). In the presence of L-arginine, 2OG binds to Mn(II) (substituting for Fe(II)) in a bidentate manner with the 2OG C5 carboxylate interacting with R277 through a hydrogen bond interaction, while the oxygens on the 2OG C1 carboxyl group and C2 keto group are coordinated to the metal centre.

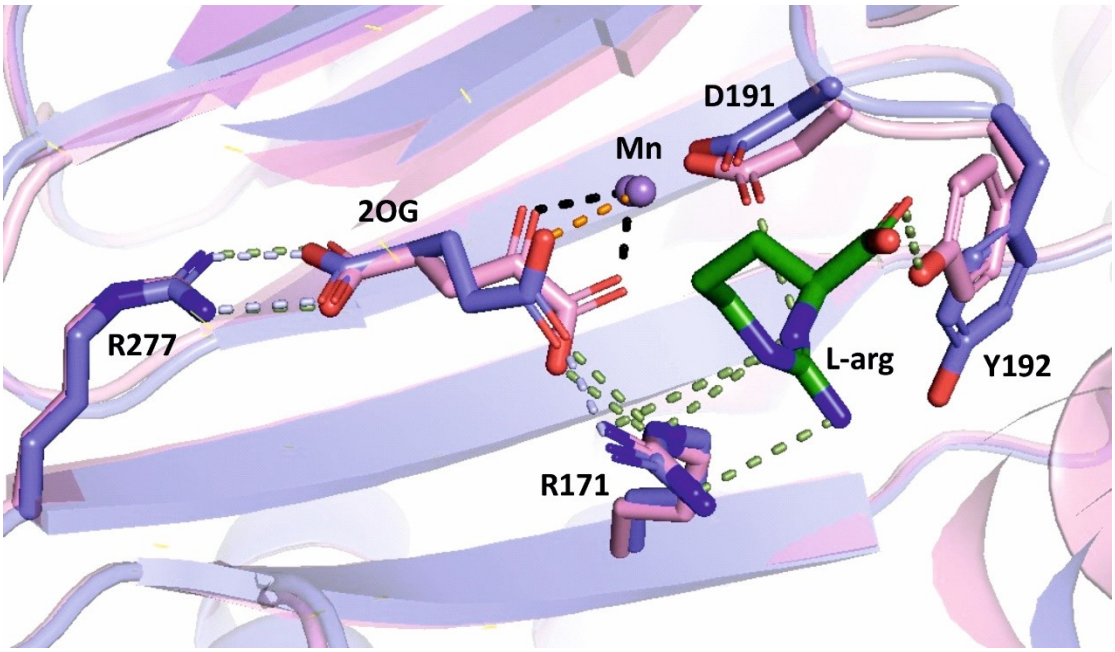


Figure 1.13 Structural superposition view of PsEFE co-crystallised with 2OG, Mn(II) with or without L-arginine (PDB: 5V2X & 5V2Y). The structural view without arginine (PDB: 5V2X) is displayed in slate, with hydrogen bonding interactions shown in light blue dash lines and 2OG coordination to metal centre shown in pink dash lines. The structural view with arginine (PDB: 5V2Y) is displayed in pink, with hydrogen bonding interaction shown in green and 2OG coordination to the metal centre shown in cyan. RMSD = 0.469Å.

The binding of L-arginine induces conformational changes in the PsEFE active site: 2OG shifts from a monodentate manner to a bidentate coordination to Mn (II), (Fig. 1.13) [49]. L-arginine binding also induces a shift in the orientation of D191, in a manner to facilitate the hydrogen bonding interaction with the NH1 of the L-arginine;

a significant positional change for Y192 also occurs to form hydrogen bonding interaction of its phenolic OH group with the doubly bonded oxygen in the carboxylic acid group of L-arginine (Fig. 1.13).

1.5 Aims of the work described in this thesis

Despite the fundamental roles of ethylene-forming enzymes in plant development and their potential for biocatalysis, their mechanisms are not yet fully understood. The aims of the results described in this thesis were: (i) to investigate the mechanisms of ACCO and PsEFE, including the roles of bicarbonate and L-ascorbate during ACCO catalysis; and (ii) to apply the mechanistic knowledge to enable the design of ethylene-forming enzymes that catalyse formation of gaseous olefins other than ethylene, including propylene.

Chapter 2 describes: (i) the optimisation of a ^{1}H NMR-based assay to monitor ACCO catalysis by determining product formation and substrate depletion to investigate the *Amborella trichopoda* 1-aminocyclopropanecarboxylate acid oxidase (AtACCO) substrate scope, and (ii) the biochemical and structural characterisation of AtACCO variants to probe its mechanism. Chapter 3 describes the potential of engineering ACCO to enable the biocatalytic production of propylene and 1-butene. Chapter 4 describes structural studies on a newly characterized bacterial ethylene-forming enzyme, providing insights into its co-substrate binding modes. Chapter 5 summarises the results and provides possible future research directions.

Chapter 2. Mechanistic studies on ACCO Catalysis

2.1 Introduction

The work described in this chapter focused on mutational studies to investigate the effects of AtACCO residues interacting with ACC, bicarbonate, and L-ascorbate on catalysis. Protein crystallography was employed to complement the mechanistic insights gained into the roles of these residues in AtACCO catalysis.

2.2 The expression and purification of wt *Amborella trichopoda* ACCO and ACCO variants

The pET28a(+)-SUMO plasmid encoding for the *Amborella trichopoda* ACCO (AtACCO) was provided by Dr. Zhihong Zhang (University of Oxford). This plasmid encodes for AtACCO with an N-terminal small ubiquitin-like modifier (SUMO) domain which is located between the N-terminal His_6-tag and the AtACCO gene sequence (Fig. 2.1).

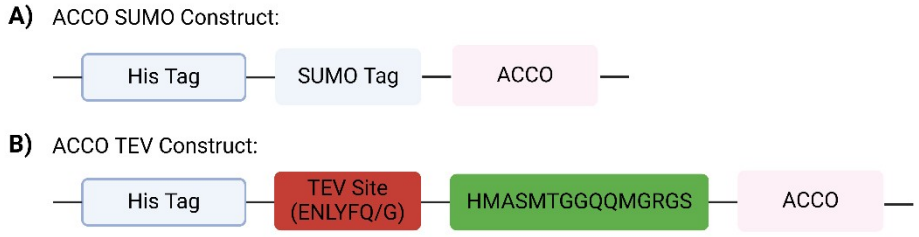
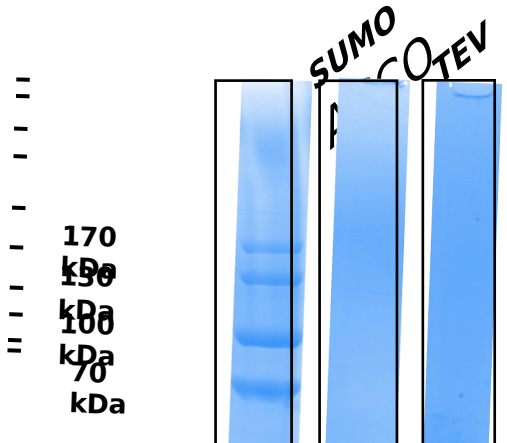


Figure 2.1 ACCO plasmid Constructs. A) Illustration of ACCO encoded in a pET28a(+)-SUMO vector. B) Illustration of ACCO encoded in a pET28a(+)-TEV vector.

Recombinant AtACCO was produced in chemically competent *E. coli* cells. The resulting cell lysates were first purified with Ni(II)-affinity column using reported procedures. (as described in Chapter 6.2.3) In this process, the recombinant proteins with a *His*$_6$-tag attached remain bound to the Ni(II) on the resin, whereas other proteins from the bacterial cells should not bind the resin and hence flow through the column. The recombinant AtACCOs were subjected to overnight protease cleavage to remove their N-terminal *His*$_6$-tags. The pET28a(+)-SUMO plasmid contains a SUMO-specific protease 2 (SenP2) cleavage site which can used to cleave the SUMO domain from AtACCO at the C-teminus of the SUMO domain. *His*$_6$-tag cleavage is followed by a reverse Ni(II)-affinity column purification to remove the protease. In the course of this purification, the cleaved recombinant proteins flow through the column whereas the protease, which contains a *His*$_6$-tag, is retained on the Ni(II)-affinity column. The recombinant ACCOs were subsequently subjected to size exclusion chromatography for further purification. Mass spectrometry was used to confirm both the identity and purity of the purified AtACCO (yield: ~8.78 mg/L media) (Fig. 2.2-2.3, S1-21).



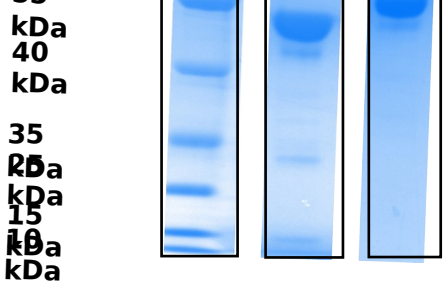


Figure 2.2 SDS-PAGE gel of AtACCO production from a pET28a(+)-SUMO vector or a pET28a(+)-TEV vector.

An alternative plasmid encoding the AtACCO was purchased from Genscript with the AtACCO being encoded on a pet28a(+)-TEV backbone (Fig. 2.1b). This construct was obtained because the pET28a(+)-SUMO ACCO plasmid was contaminated, potentially with bacteriophage. Expression of this plasmid in chemically competent cells and subsequent purification with Ni(II) affinity chromatography and size exclusion chromatography afforded AtACCO (yield: ~ 4.32 mg/L media) (Fig. 2.2, 2.4, S22-25).

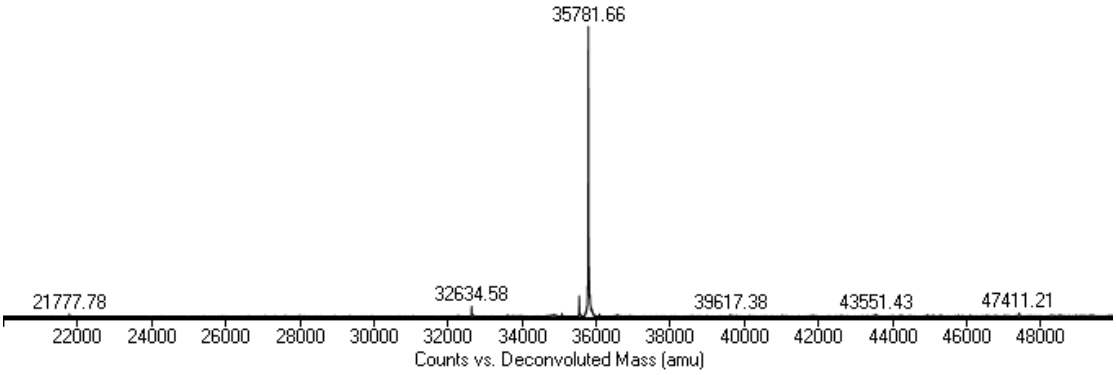


Figure 2.3. Representative deconvoluted mass spectrum of ACCO encoded on pET28a(+)-SUMO

vector. Deconvoluted mass spectrum wt AtACCO. Calculated mass: 35781.95; Observed mass: 35781.66 Da.

The pET28a(+)-TEV plasmid encodes for proteins containing a TEV protease recognition site (ENLYFQ/G; cleavage between Q and G) which is located between the N-terminal His_6-tag and the AtACCO gene sequence, enabling the removal of the His_6-tag during the purification process (Fig. 2.1b). 15 additional amino acids (that is: HMASMTGGQQMGRGS) were inserted between the TEV recognition site and the restriction site used for cloning. Along with the residual glycine from the TEV cleavage site, a total of 16 additional acids will be present at the N-terminus of the AtACCO construct purified. This 15-amino-acid sequence was inserted to improve the exposure of the TEV cleavage site to the TEV protease and enable cloning of the protein sequence into the commercial pET28a-TEV vector where the protein sequence will not contain an identical restriction site to that for cloning. The sequence was determined by the manufacturer, Genscript. The mass of the AtACCO was thus 1592 Da greater than the predicted protein mass of AtACCO without the additional N-terminal amino acids. Mass spectrometry analysis confirmed the identity and purity of the purified AtACCO (Fig. 2.4, S22-25).

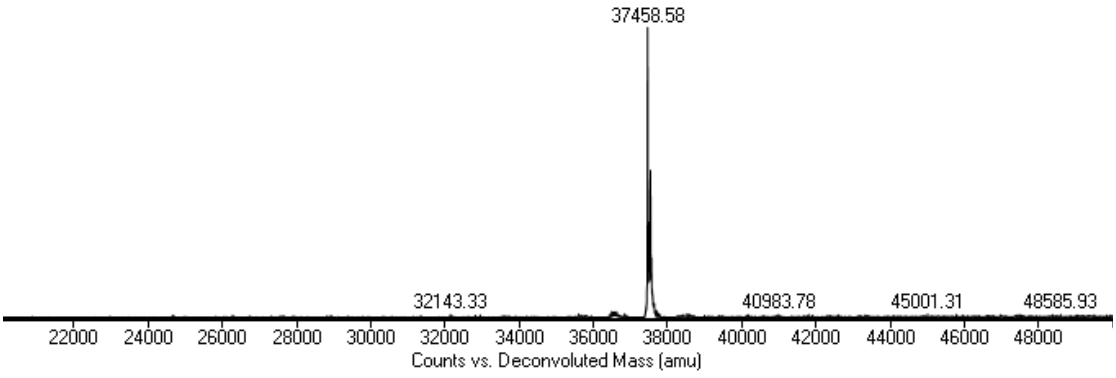


Figure 2.4. Representative deconvoluted mass spectrum of ACCO encoded on pET28a(+)-TEV vector. Deconvoluted mass spectrum of the S247G AtACCO variant. Calculated mass: 37457.88 Da; Observed mass: 37458.58 Da. With the additional 16 residues, the masses are approximately 2k Da higher than the counterparts encoded on the pET28a(+)-SUMO vector.

2.3 Optimisation of an NMR assay for ACCO activity

An ^{1}H NMR-based assay was developed to monitor AtACCO catalysis and enable quantification of product formation to investigate the substrate scope of ACCO (Fig. 2.5). The NMR assays complement reported assays for ACCO catalysis that employ gas chromatography[50,51]. Although the ethylene signal was observed by ^{1}H NMR (~ 5.44 ppm), given the gaseous nature of ethylene and its relatively poor aqueous solubility, substrate depletion was used as a mean to quantify turnover with inert 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP-d_4) as the internal standard. This workflow assumed that AtACCO does not catalyse reactions other than the fragmentation of ACC into ethylene, for which no evidence was obtained by NMR analyses.

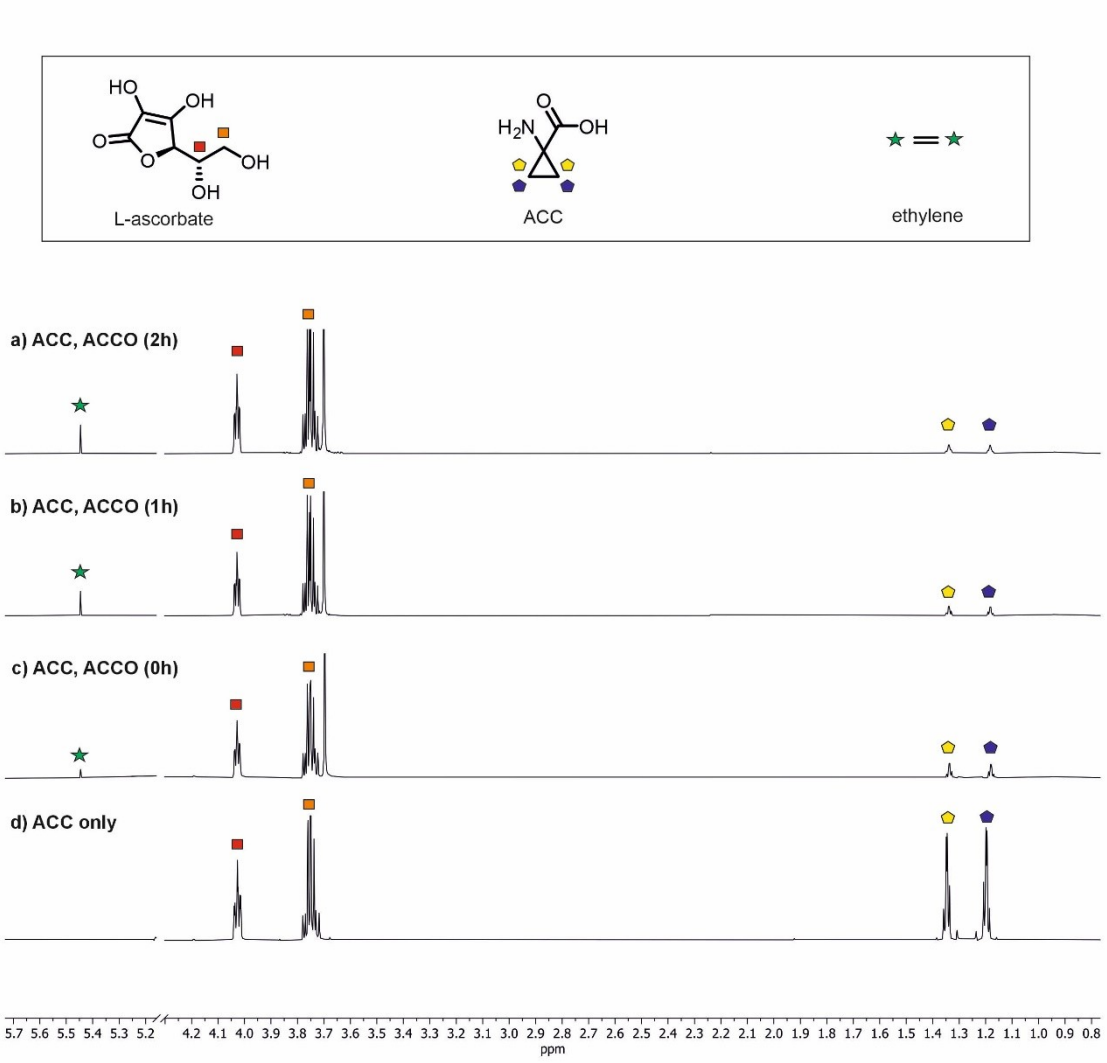


Figure 2.5. ACCO catalysis with ACC as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of ACCO with ACC after: **a)** 2h, **b)** 1h and **c)** 0 h. **d)** ^{1}H NMR of ACC only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-*d*$_4$ in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

As a starting point for an optimised assay, the following conditions were used that were developed by Samuel E. Digby (University of Oxford) on AtACCO orthologs: 800 µM ACC, 40 µM ACCO, 2 mM $NaHCO_3$, 1.6 mM L-ascorbate, 50 mM Fe (II). The ACC concentration was first varied to enable clear comparison between conditions. The ACC concentration was increased gradually while the AtACCO

concentration was kept at 40 µM. It was observed that the percentage substrate depletion decreases as the ACC concentration increases (Fig. 2.6a). As AtACCO addition was found to negatively impact the signal-to-noise ratio, 200 µM ACC was employed instead of 800 µM ACC to give better levels of substrate depletion, while retaining an acceptable substrate-to-noise ratio.

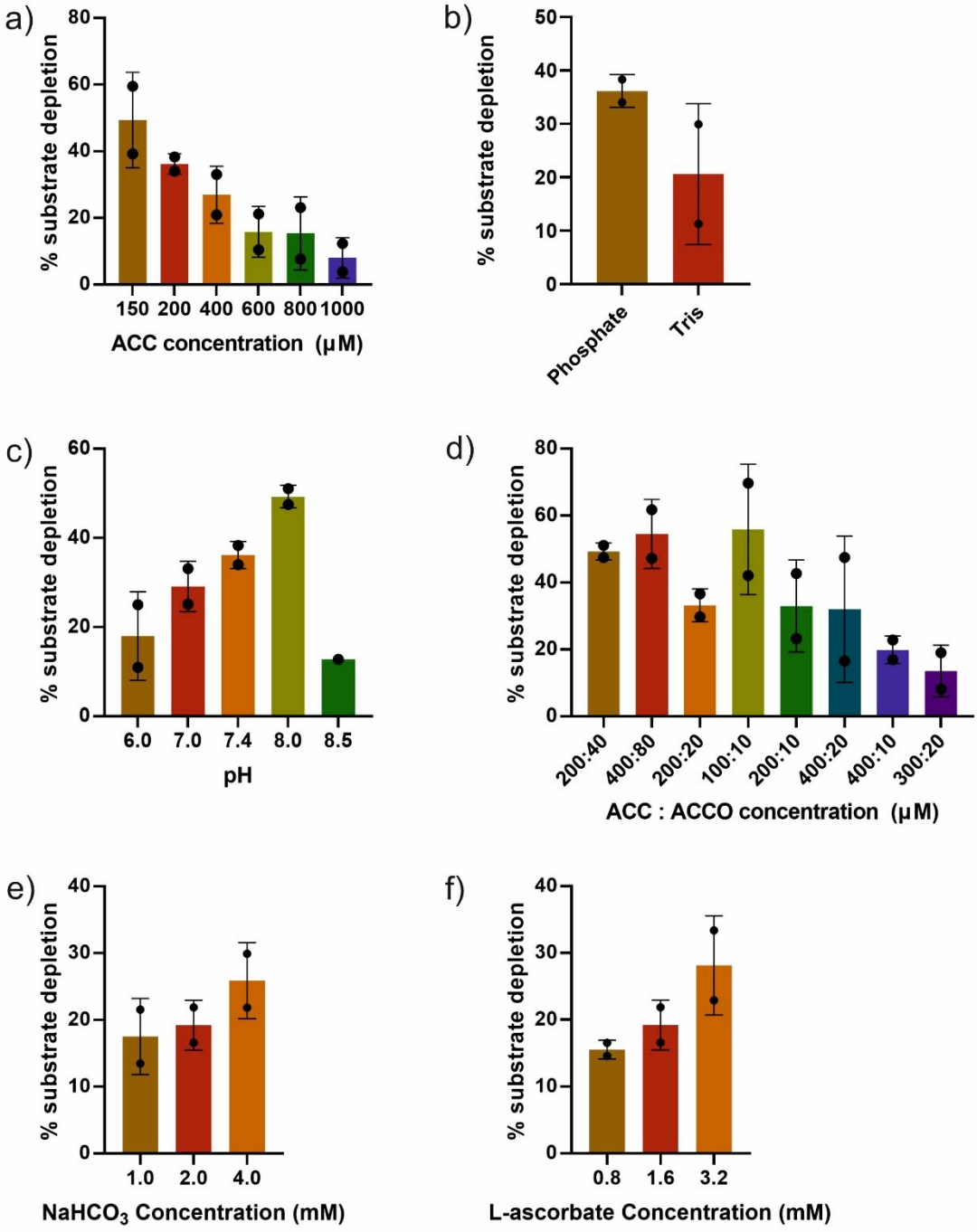


Figure 2.6 ^{1}H NMR assay optimisation. a) Effect of varying ACC concentration on ACCO turnover. 200 µM ACC was chosen, because of the high level of turnover and adequate signal-to-noise ratio; b) Buffer screening: Phosphate buffer was more efficient than Tris buffer; c) pH screening with 50 mM phosphate buffer. pH 8 was most efficient; d) Effect of the enzyme:substrate ratio: 400 µM ACC:20 µM ACCO was gave best turnover and signal-to-noise; e) Bicarbonate concentration. 4 mM $NaHCO_3$ was most efficinet; f) L-ascorbate concentration: 3.2 mM L-ascorbate was most efficient. (n=2, experiment replicate, error bar = standard error)

The effect of the buffer on AtACCO catalysis was tested with Tris-d_{11} and phosphate buffers (Fig. 2.6b). The results revealed that AtACCO catalysis was more efficient in phosphate buffer than in the previously used Tris-d_{11} buffer (by Samuel E. Digby, University of Oxford, on AtACCO orthologs). As ACCO is reported to localise in the cytoplasm where the pH is in the range of 7.1-7.5, the effect of the pH on substrate turnover was investigated in the range of pH 6-8.5 (Fig. 2.6c). The results revealed substantially improved turnover at pH 8, enabling further optimisation of the enzyme:substrate ratio. Due to the relatively high noise levels from ACCO at pH 8, a compromise had to be reached between turnover level and worsened signal to-noise ratio (Fig. 2.6d). 400 µM ACC:20 µM AtACCO was thus employed although this ratio did not give the highest substrate depletion.

Lastly, the effects of $NaHCO_3$ and L-ascorbate on AtACCO catalysis were examined by varying their concentrations. The increasing concentrations of both $NaHCO_3$ and L-ascorbate resulted in increased ACC substrate depletion (Fig. 2.6e-f). The simultaneous increase of these co-factor concentrations was anticipated to further boost the enzymatic activity. Hence, the concentrations of the two cofactors were doubled, resulting in the following partially optimised ^{1}H NMR ACCO assay conditions: 400 µM substrate, 20 µM AtACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). The optimised assay condition increased the turnover by >2-fold compared to the previous conditions (that is: 800 µM ACC, 40 µM ACCO, 2 mM $NaHCO_3$, 1.6

mM L-ascorbate, 50 mM Fe(II) in Tris-d_{11}, pH 8).

The low ACC turnover by AtACCO observed was proposed to be a result of increasing cyanide concentration during catalysis. Cyanide is a strong metal ligand which hinders oxygen binding to protein complexed Fe(II), for example in the case of haemoglobin.[52] Thus, the AtACCO catalysed conversion of ACC into ethylene was trialled in the presence of variable cyanide concentrations. The ^{1}H NMR assay results revealed that increasing cyanide concentrations reduced ACC turnover and ethylene formation. (Fig. 2.7)

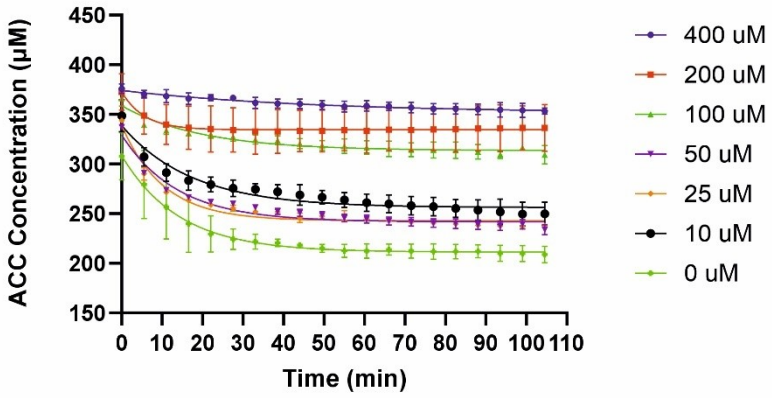


Figure 2.7
AtACCO catalysed conversion of ACC into ethylene in the presence of variable cyanide concentrations. (n=2, experiment replicate, error bar = standard error**)**

It was also investigated whether ions other than bicarbonate can sustain AtACCO catalysis (Fig. 2.8). The results revealed that carbonate was able to sustain AtACCO catalysis to a similar extent as bicarbonate (~90% of activity by bicarbonate), whereas all other ions tested were mostly inactive with only trace amounts of ethylene being observed, an observation which may reflect that the buffer contained low levels of

atmospheric CO_2 (Fig. 2.8). The combined results suggest that bicarbonate (or carbonate) is highly important for ACCO catalysis, consistent with pioneering work.[53]

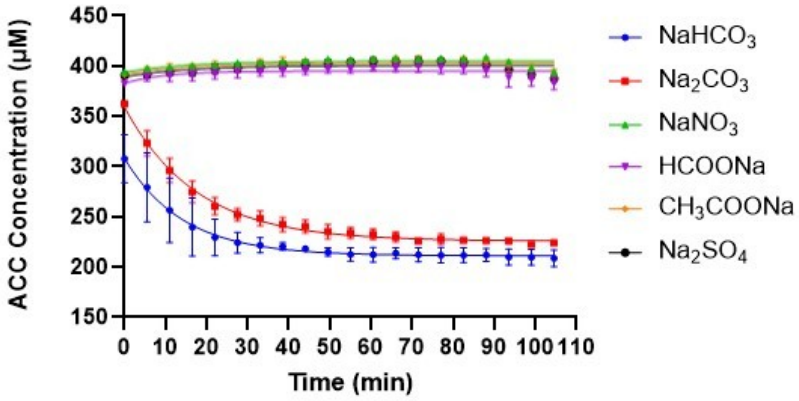


Figure 2.8 Kinetic profiles of ions on AtACCO catalysis. (n=2, experiment replicate, error bar = standard error**)**

2.4 Mutagenesis studies of AtACCO

Mutagenesis studies were conducted to gain insights into the roles of specific active site residues in AtACCO catalysis by crystallography. The focus lay initially on the residues observed to directly interact with the substrate (V185, N217, R245, S247, Y163, L196; Fig. 2.9a). Based on the wildtype (wt) AtACCO crystal structure obtained in complex with Fe(II) and ACC by Dr. Zhihong Zhang, the effects of polar residues observed to interact with the bicarbonate (K290, K159, Y91) were assessed

(Fig. 2.9b). AtACCO variants were obtained by site-directed mutagenesis followed by protein production and purification as described for wt AtACCO (as described in Chapter 2.2 & 6).

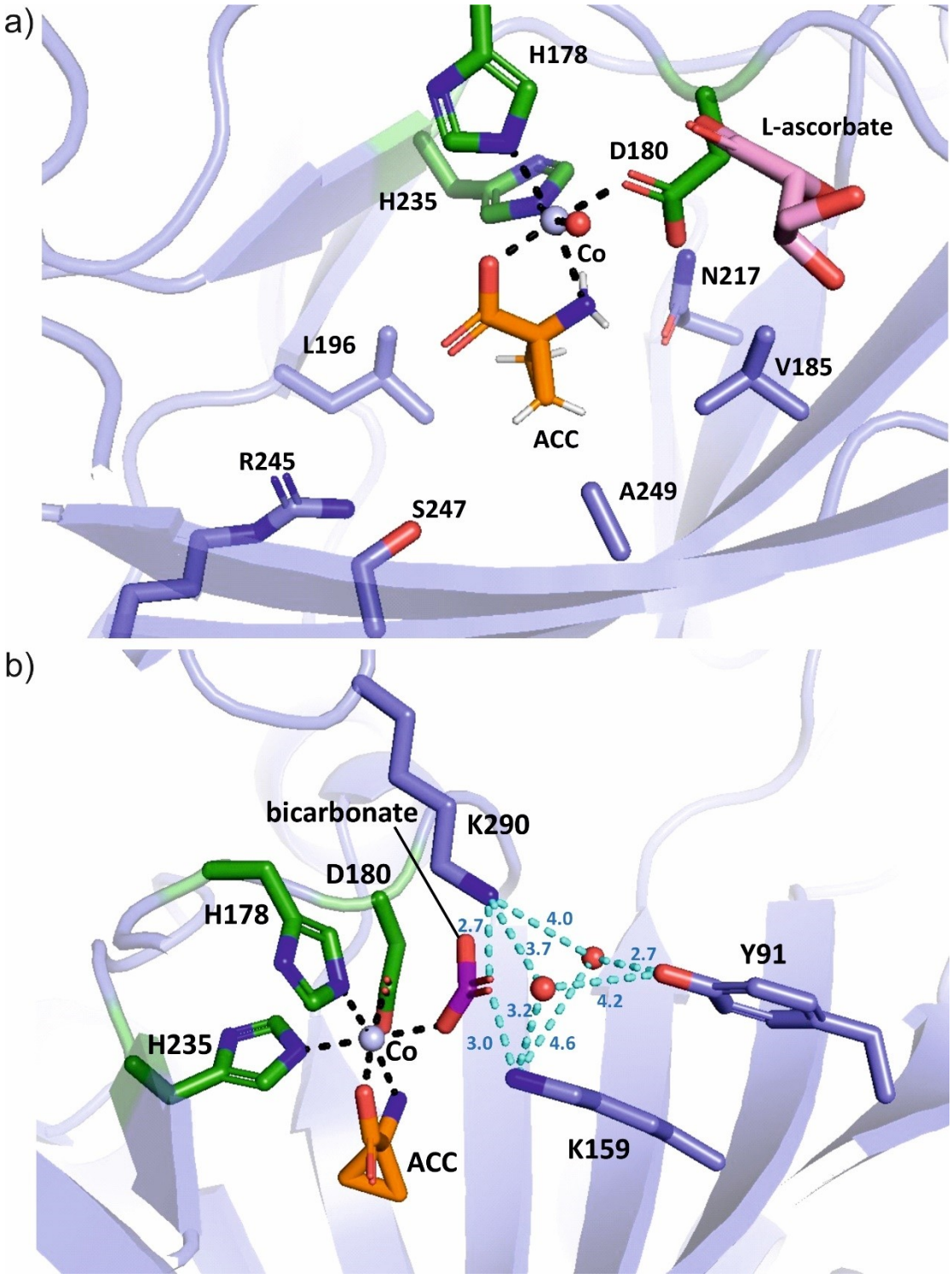


Figure 2.9. Residues interacting with ACC and bicarbonate at the AtACCO active site as

observed by crystallography. ACC is shown in orange. Co is shown as a light blue sphere. Metal binding residues are shown in green. Coordinations to the metal centre are shown with black dashed lines. a) Residues interacting with ACC are shown in slate. L-ascorbate in shown in pink; b) Bicarbonate is shown in purple. Selected water molecules are shown in red spheres. Hydrogen bonding interactions involved in the formation of water network are shown in cyan dashed lines with distance in *Å*.

Together with S247, R245 forms the highly conserved RXS motif commonly found in AtACCO orthologs[45]; these two residues coordinate to the carboxylic acid group of ACC through a water network along with Y163 (Fig. 2.10). The network consists of two water molecules bridging the interactions between the carboxylic acid group of ACC and three polar residues (Y163, R245 & S247). The R245Q AtACCO variant was designed to assess the impact of the R245 guanidinium group on the formation of the water network and on stabilizing ACC binding. ^{1}H NMR-based analysis R245Q ACCO variant revealed formation of only trace amounts of ethylene with no observable substrate depletion (Fig. 2.11a), supporting on the important role of R245 in ACC binding.

In contrast to R245Q AtACCO, ^{1}H NMR-based assays revealed that S247A AtACCO only manifested ~15% activity of wt AtACCO catalysis. This observation supports a reported mutagenesis study on *Malus domestica* ACCO1 where the catalytic activity of the equivalent ACCO variant was reduced by 66%.[45] Notably, S247G ACCO was more active than S247A ACCO; its catalytic activity was 64% that of wt ACCO (Fig. 2.11a). This effect may reflect the greater void created by the S247G variation, likely permitting partial restoration of the water network.

Y163H and Y163F AtACCO were made to assess the role of hydrogen bonding in the water network involving Y163. (Fig. 2.10) Unexpectedly, Y163F AtACCO induced a lower reduction in the catalytic activity (~21% of wt AtACCO catalytic activity) than Y163H ACCO (~43% of wt AtACCO activity), despite Y163H ACCO being able to participate in hydrogen bonding (Fig. 2.11a). This observation may reflect the polar interaction induced by Y163H and the disruption of the water network of the RXS motif by the polar side chain of Y163H which impacts on the binding of the ACC substrate (Fig. 2.10).

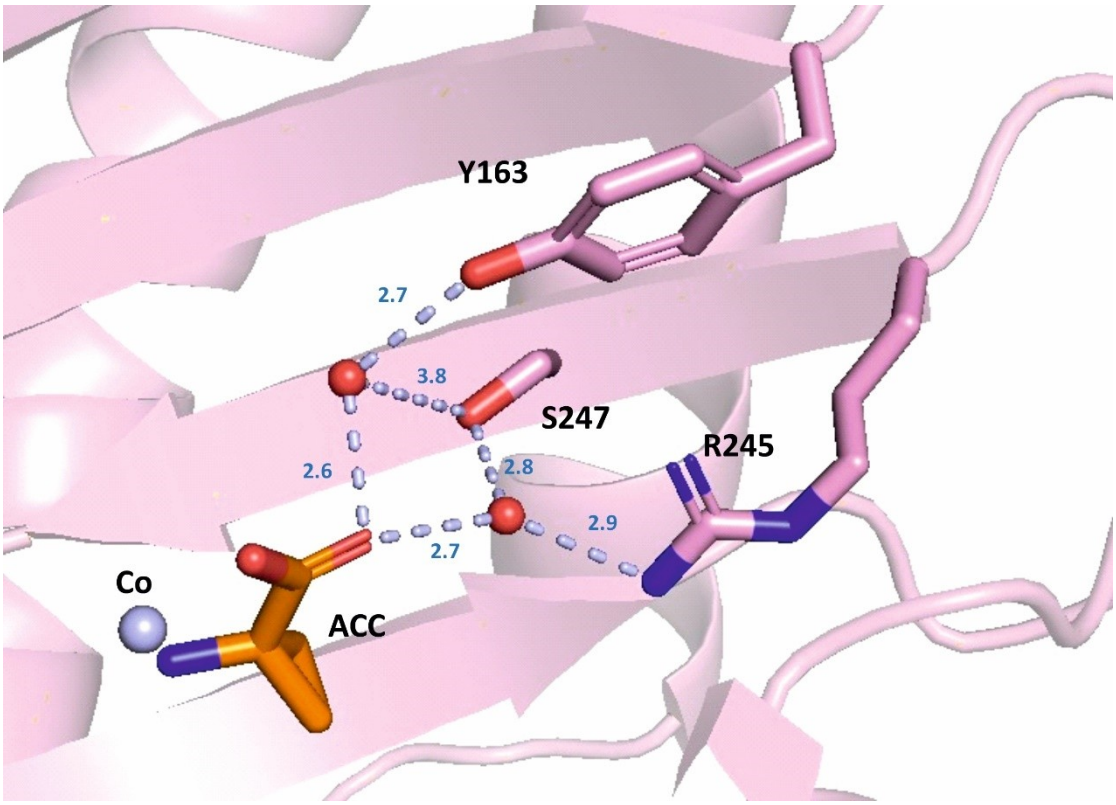


Figure 2.10. The interaction of the conserved RXS motif interacting with ACC via a water network (slate dashed line). Co(II) is shown in light blue sphere. ACC is shown in orange. Residues are shown in pink.

Based on the analysis of the AtACCO:Co(II):ACC complex crystal structure, the side chain of N217 is positioned to interact with the amino group of ACC through hydrogen bonding (Fig. 2.9a). Thus, the N217Q AtACCO variant was made to assess whether the elongation of the N217 side chain would disrupt the binding of the ACC amino group to Fe(II). N217Q ACCO, like R245Q ACCO, catalysed formation of only trace amounts of ethylene (Fig. 2.11b), potentially because the longer glutamine side chain competed with the amino group of ACC for space in the active site and/or binding to Fe(II). A previous mutagenesis study with *Malus domestica* ACCO1 led to the observation that N217F (equicalent to N217 in AtACCO) mutation manifested 43% of the wt AtACCO activity[45]. This proposal is supported by the observation that the catalytic activity of N217G AtACCO was ~72% that of wt AtACCO.

The roles of V185 and L196 on ACCO catalysis were also assessed as these two residues likely stabilise ACC binding through hydrophobic interactions of their side chains with the ACC cyclopropane ring (Fig. 2.9a). L196A AtACCO retained only ~14% of the catalytic activity of wt AtACCO (Fig. 2.11b), implying that L196 plays an important role in stabilising the binding of the ACC which is crucial for catalysis despite not being directly involved in the ring-opening process of ACC.

In contrast to the L196A AtACCO variant, the V185G AtACCO variant did not catalyse ethylene formation (Fig. 2.11b). Our study suggested that the access to L-

ascorbate is essential for ACC enzymatic oxidation. ACCO crystallography data suggests the possibility of a secondary ACC binding site at the sixth coordination site of Fe(II) (unpublished data, Dr. Zhihong Zhang, University of Oxford). V185G could have enabled the distortion of the primary ACC binding site and hence the binding of a secondary ACC, blocking the access of the primary ACC to L-ascorbate and resulting in a complete loss of ethylene formation. Alternatively, the V185G variation might have created a void in the binding pocket that significantly abolishes hydrophobic interaction with the cyclopropane ring on ACC.

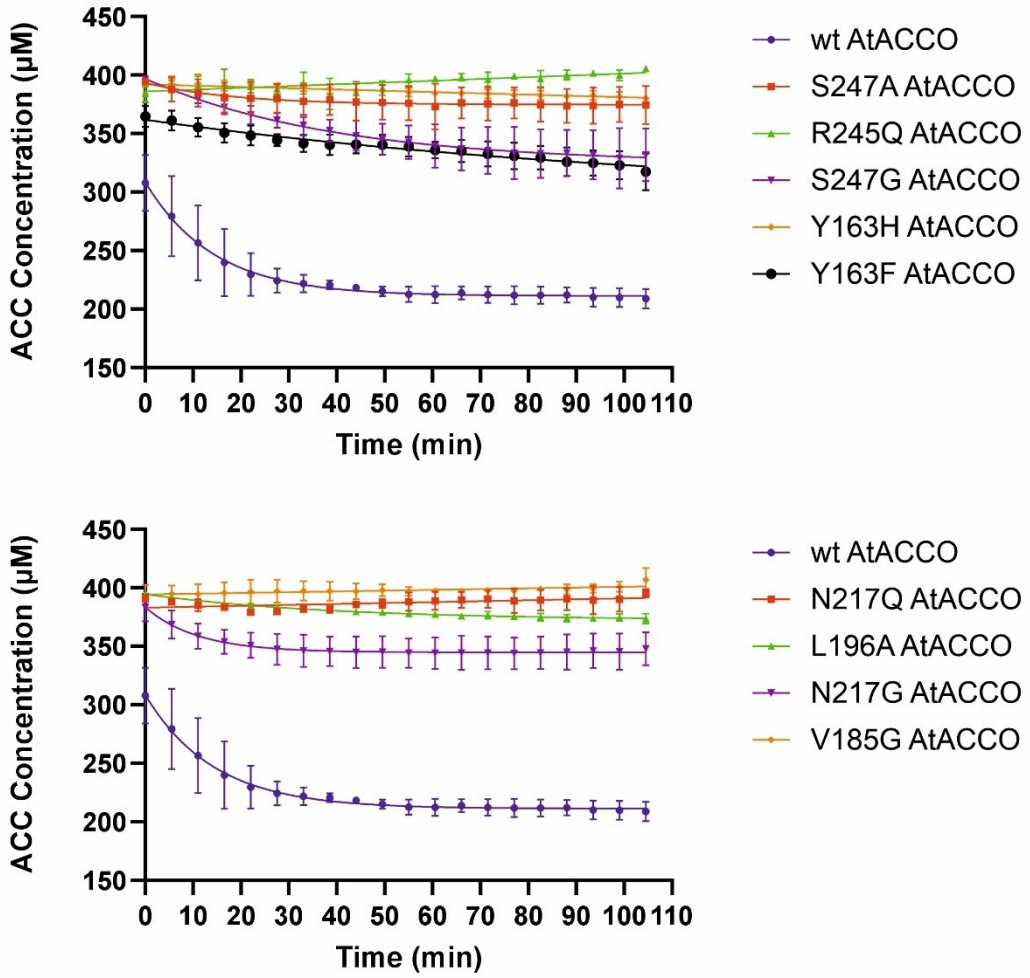


Figure 2.11 Effects of the RXS motif and residues stabilizing the hydrogen bonding network on

ACCO catalysis. a) Residues involved in the hydrogen bonding network by the RXS motif for ACC binding (R245, S247, Y163); b) Other residues involved in ACC binding (N217, L196, V185). **(**n=3, experiment replicate, error bar = standard error**)**

The role of AtACCO residues involved in co-factor binding was also explored. Analysis of the AtACCO:Co(II):ACC:HCO_3^- complex crystal structure suggests the presence of a triad of residues, Y91, K159 and K290 at the potential bicarbonate ion binding site (Fig. 2.9b). Thus, AtACCO variants were produced to affect the hydrogen bonding capabilities of these residues, including Y91F, Y91H, K159A, K159E, K159R, K290A, K290E, and K290R AtACCO, to assess the effect of completely or partially removing the polar interaction in the triad formed by Y91, K159 and K290.

The Y91H AtACCO variant was only slightly less active (~93% of wt AtACCO activity), whereas the Y91F AtACCO variant, which is unable to engage in hydrogen bonding interactions involving its side chain, showed only trace levels of activity, with substrate depletion reaching the detection limit of the assay. Nevertheless, the presence of an ethylene peak was still observable using the Y91F AtACCO variant.

The two lysine residues in the triad (K159 and K290) were mutated to arginine, alanine, and glutamic acid residues respectively. It was hypothesised that the mutation of lysine to arginine or alanine residues could partially or completely remove the polar interactions between lysine and the bicarbonate ion, while mutations to glutamic acid could change the side chain charge from positive to negative and consequently hinder binding of the bicarbonate ion. Previous work on *Malus domestica* ACCO1 demonstrated that the K158R, K158L and K158E MdACCO1 variants only retained

7.2%, 2.0%, and <1% activity of wt *Malus domestica* ACCO1[45]. With AtACCO, all six mutations led to significant loss of AtACCO activity. Notably, the K159E variation completely abolishes AtACCO activity (Fig. 2.12a), likely reflecting the importance of the positive charge on the K159 side chain to stabilise binding of the negatively charged bicarbonate anion. Similarly, the K159A, K290A and K290E AtACCO variants also demonstrated only trace level activity. By contrast, K159R and K290R AtACCO retained relatively higher levels of catalytic activity contrasting with the K159A, K290E and K290A AtACCO variants, possibly reflecting that the side chains of K159R/K290R are able to engage in polar interactions with bicarbonate (Fig. 2.12a-b). Collectively, the biochemical work in this section supports the assignment of the bicarbonate binding site based on crystallography.

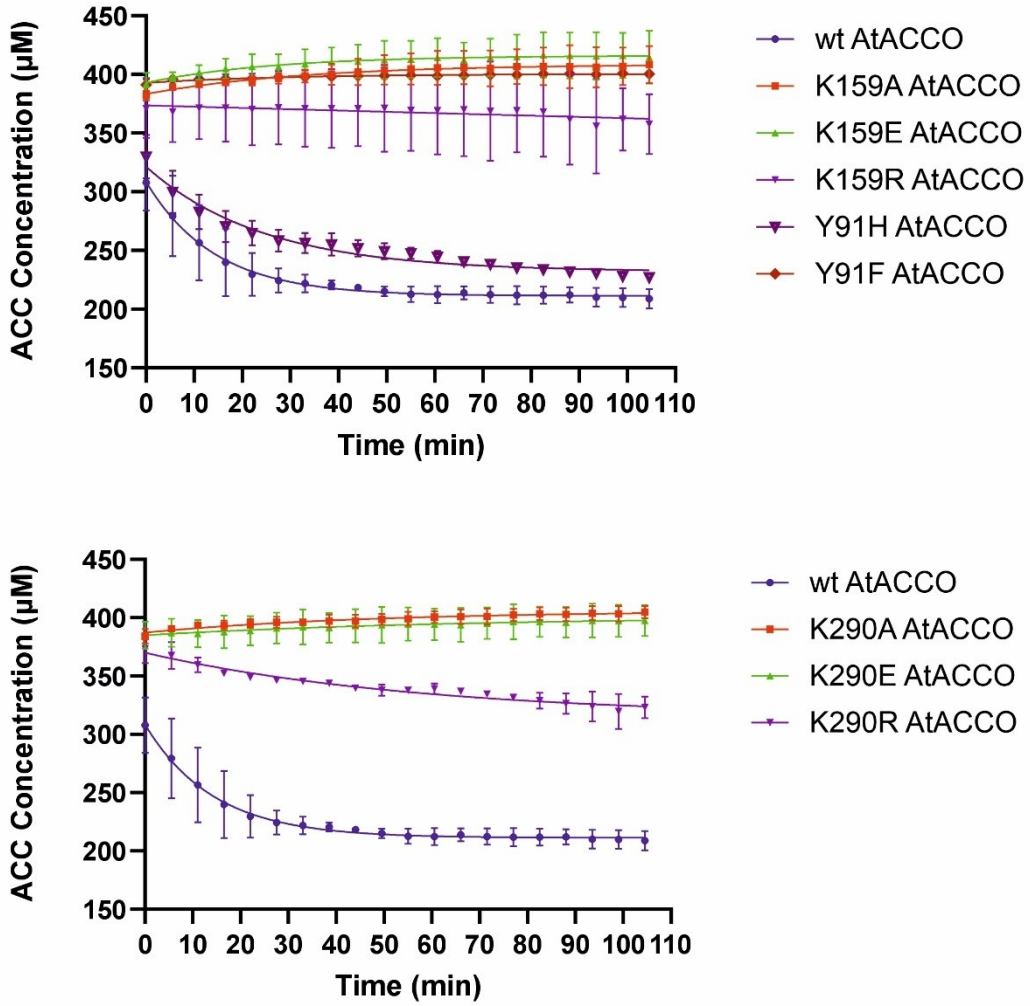


Figure 2.12 Effect of ACCO active site residues potentially engaged in bicarbonate or L-ascorbate binding on catalysis. The modified residues interact with a) bicarbonate (K159, Y91) and b) both bicarbonate and L-ascorbate (K290). (n=3, experiment replicate, error bar = standard error)

The AtACCO crystallographic studies imply that N253 and Y286 are key residues in binding L-ascorbate at the AtACCO active site via hydrogen bonding interactions (as described in Chapter 1.4.1). The N253A and Y286A AtACCO variants were designed to abrogate the hydrogen bonding interactions with L-ascorbate at the AtACCO active site. As expected, N253A and Y286A AtACCO only catalysed the formation of a trace amount of ethylene as revealed by ^{1}H NMR-based assays (Fig. 2.13). N253W AtACCO was designed to increase the steric bulk of the side chain, by partially

occupying the pocket for L-ascorbate. N253W AtACCO also catalysed the formation of a trace amount of ethylene. The results collectively suggest that L-ascorbate binds in the active site metal to enable the enzymatic oxidation process through direct coordination to the metal centre.

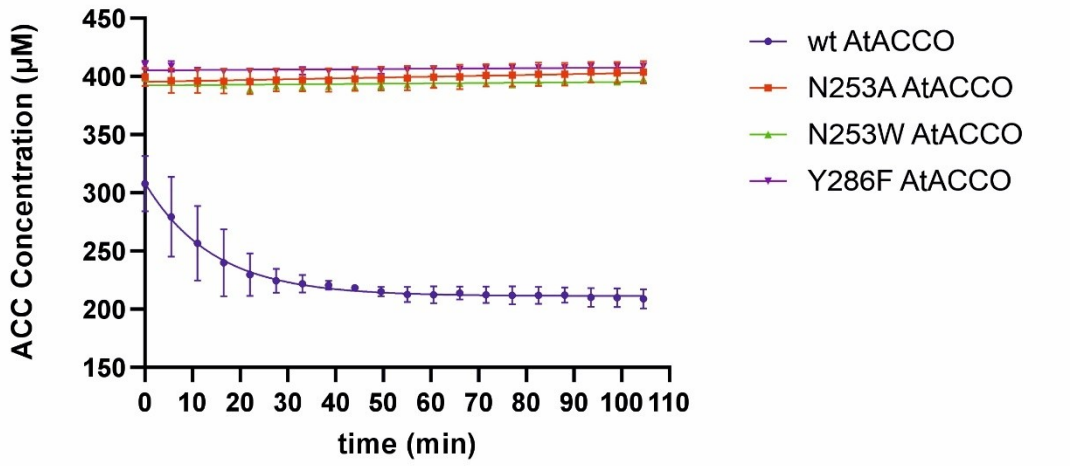


Figure 2.13 Effect of ACCO active site residues potentially engaged in L-ascorbate binding on catalysis. The modified residues interact with bicarbonate and L-ascorbate (N253 & Y286). (n=3, experiment replicate, error bar = standard error)

2.5 Investigations on the substrate scope of ACCO

The substrate scope of AtACCO was investigated to inform on whether derivatisations on ACC will affect ACCO catalysis and to probe potential alternative reaction pathways. Thus, the reaction of ACC derivatives with wt AtACCO was investigated under optimised conditions using the ^{1}H NMR-based turnover assay.

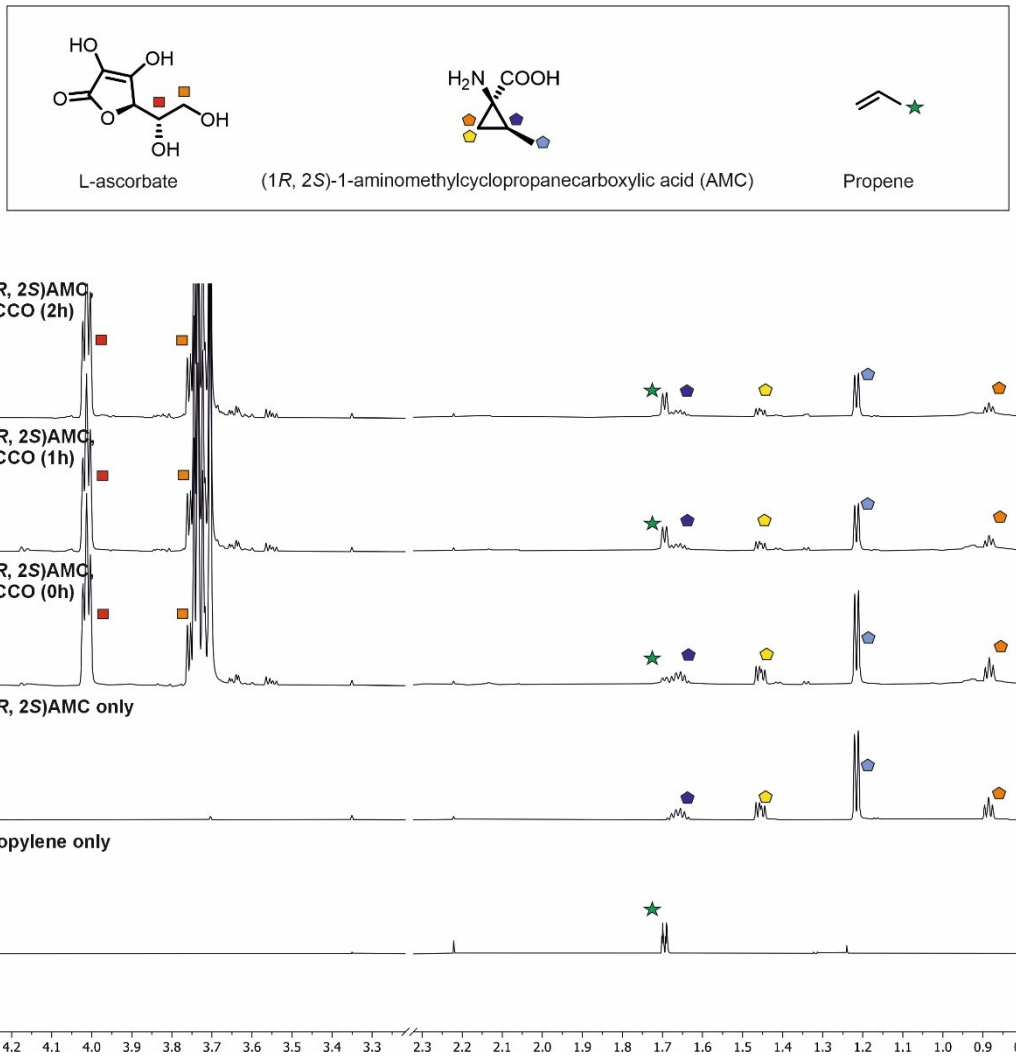


Figure 2.14. **ACCO catalysis with 1-aminomethylcyclopropanecarboxylic acids as a potential substrate.** 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with (1*R*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid after **a)** 2 h, **b)** 1 h and **c)** 0 h. ^{1}H NMR spectrum of **d)** (1*R*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid alone and **e)** propene alone in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

The results with the four possible stereoisomers of 1-aminomethylcyclopropane-1-carboxylic acid (Fig. 2.14-15, prepared by Dr. Xin Wen, University of Oxford) revealed that AtACCO catalysed propylene production from all the four possible stereoisomers, as per comparison with an authentic propylene sample (Fig. 2.14). Despite the observation of propylene formation, (1*S*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid and (1*S*, 2*S*)-1-

aminomethylcyclopropanecarboxylic acid showed no significant substrate depletion, possibly reflecting the weaker binding of these two ACC derivatives to the ACCO active site due to steric interactions between the methyl group and the core ß-sheet structure (Fig. 2.15). Notably, (1*S*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid, norcoronamic acid, is a natural product intermediate reportedly found in *Pseudomonas*.[54] In contrast to the (1*S*, 2*R*) and (1*S*, 2*S*) isomers, AtACCO catalysed (1*R*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid with a turnover of ~15% of that for the natural substrate ACC and similar levels of turnover of (1*R*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid to that for ACC (Fig. 2.15).

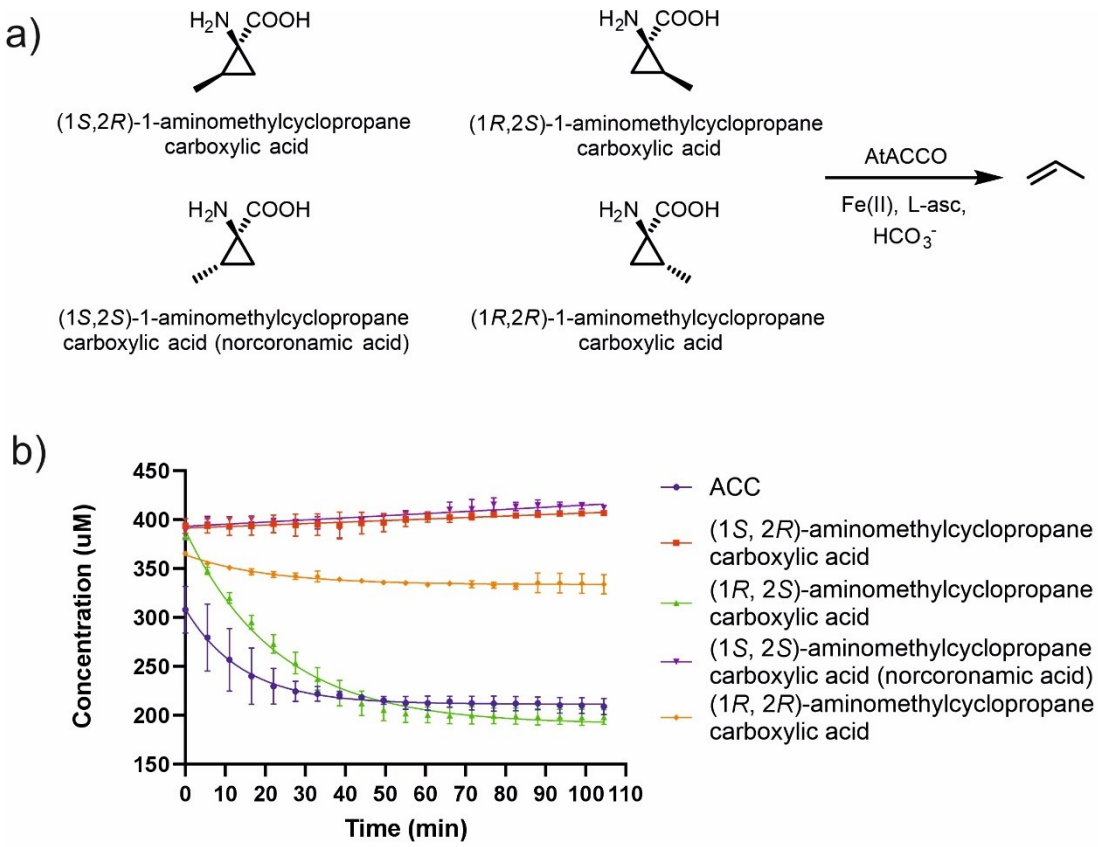


Figure 2.15 Investigation of AtACCO catalysed propylene formation from stereoisomers of 1-aminomethylcyclopropanecarboxylic acids. a) Four stereoisomers of 1-

aminomethylcyclopropanecarboxylic acid being converted to propylene by wt AtACCO. b) Turnover profiles of the four stereoisomers converted by wt AtACCO into propylene. (n=3, experiment replicate, error bar = standard error)

The reactions of AtACCO with two isomers of 1-aminoethylcyclopropane-1-carboxylic acid, (1*S*, 2*S*)-1-aminoethylcyclopropanecarboxylic acid and (1*R*, 2*R*)-1-aminoethylcyclopropanecarboxylic acid, were investigated. Notably, (1*S*, 2*S*)-1-aminoethylcyclopropanecarboxylic acid, coronamic acid, is also a natural product intermediate reportedly found in *Pseudomonas*. Although cell lysates containing ACCO have been reported to give trace amounts of 1-butene from these derivatives[55], the ^{1}H NMR assay results with isolated AtACCO reported here revealed no evidence for formation of 1-butene, possibly reflecting effective enzyme concentrations and/or subtle differences in the binding pockets of ACCO enzymes from different sources.

The screening of other ACC derivatives with AtACCO has revealed an alternative oxidation pathway[56]: AtACCO catalysed ketone formation from ACC derivatives with expanded rings (Fig. 2.16a, c), with 1-aminocyclobutane-1-carboxylic acid being converted into cyclobutanone (Fig. 2.16a, 17). However, AtACCO catalysis with ACC derivatives bearing larger rings, such as 1-aminocyclopentane-1-carboxylic acid, did not result in observable levels of ketone formation, possibly due to the compact size of the AtACCO substrate binding pocket (Fig. 2.16c). This observation contrasts with the reported reactivity of AtACCO homologues with ring-expanded ACC analogues (previously reported by Samuel Digby, University of Oxford).

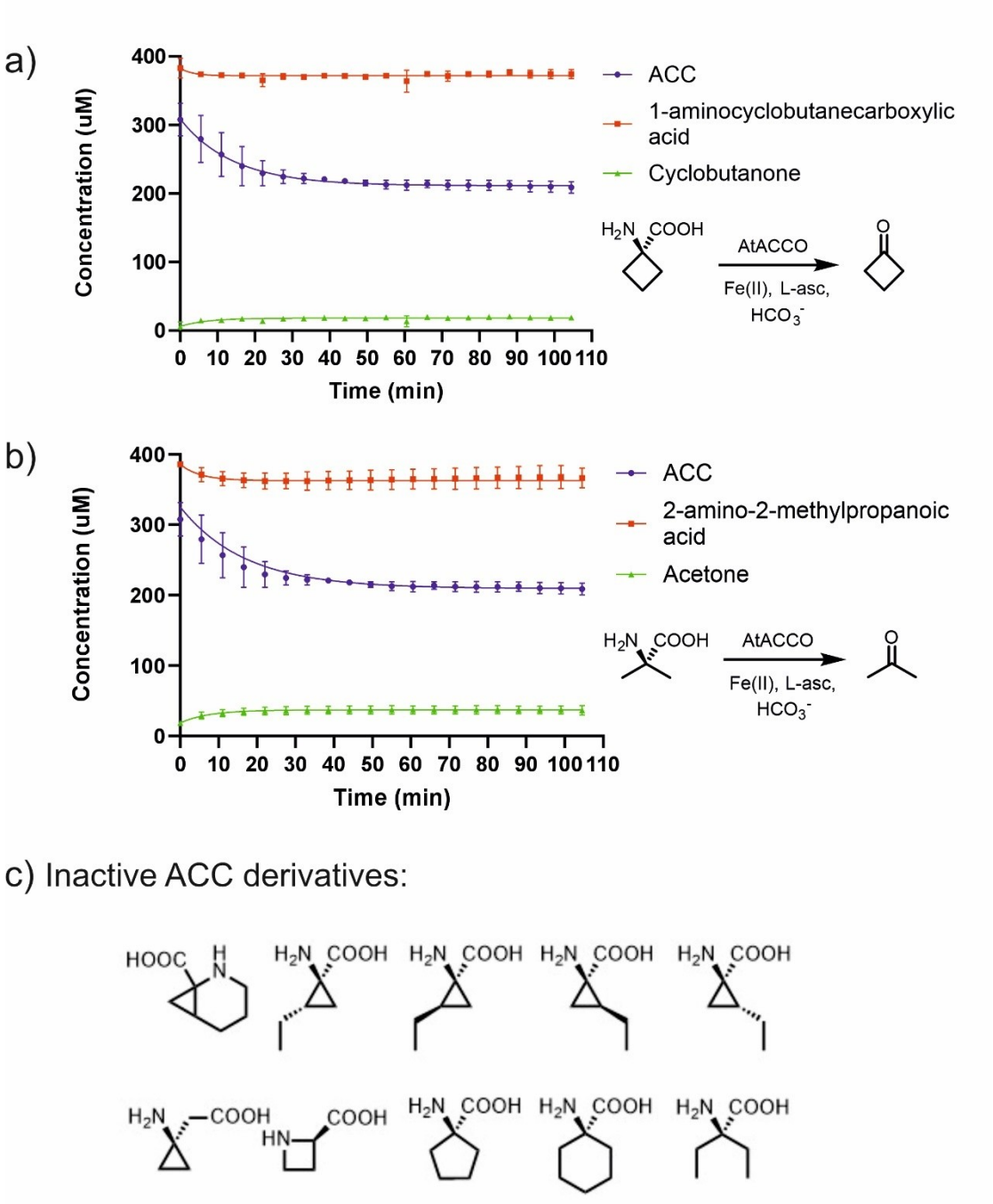
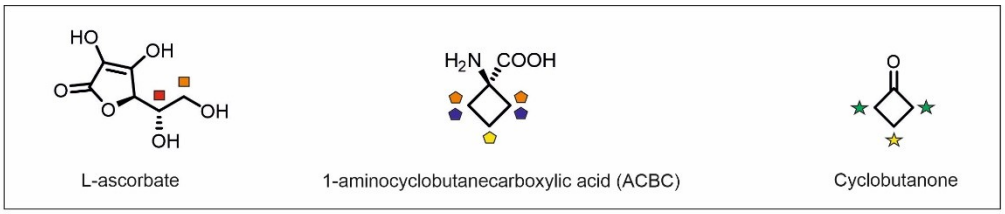
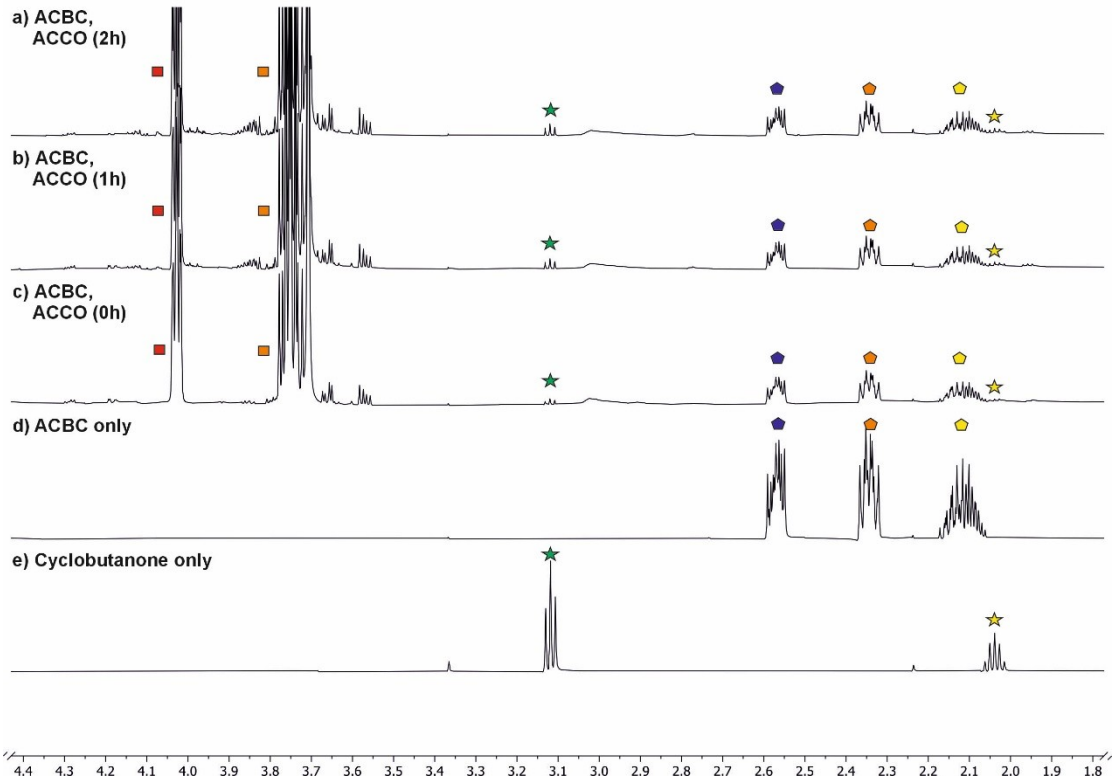


Figure 2.16 AtACCO catalyzes ketone formation reactions. AtACCO catalyses formation of a ketone product from: a) 1-aminocyclobutane-1-carboxylic acid and b) 2-amino-2-methylpropanoic acid. c) Structures of inactive ACC derivatives tested. (n=3, experiment replicate, error bar = standard error)

To probe whether the ketone formation pathway is specific to cyclic ACC derivatives,

the reaction of AtACCO with 2-amino-2-methylpropanoic acid and 2-amino-2-ethylbutanoic acid was tested. Consistent with the reported results (unpublished data by Samuel Digby, University of Oxford), 2-amino-2-dimethylpropanoic acid was oxidised to form acetone (Fig. 2.16b, 2.18), whereas 2-amino-2-diethylbutanoic acid was not oxidised by AtACCO under the tested conditions (Fig. 2.16c). Note that three other ACC derivatives were also inactive in the ACCO ^{1}H NMR-based assay: that is, 2-(1-aminocyclopropyl)acetic acid, L-azetidine-2-carboxylic acid, and (1*R*, 2*R*, 3*R*)-1-





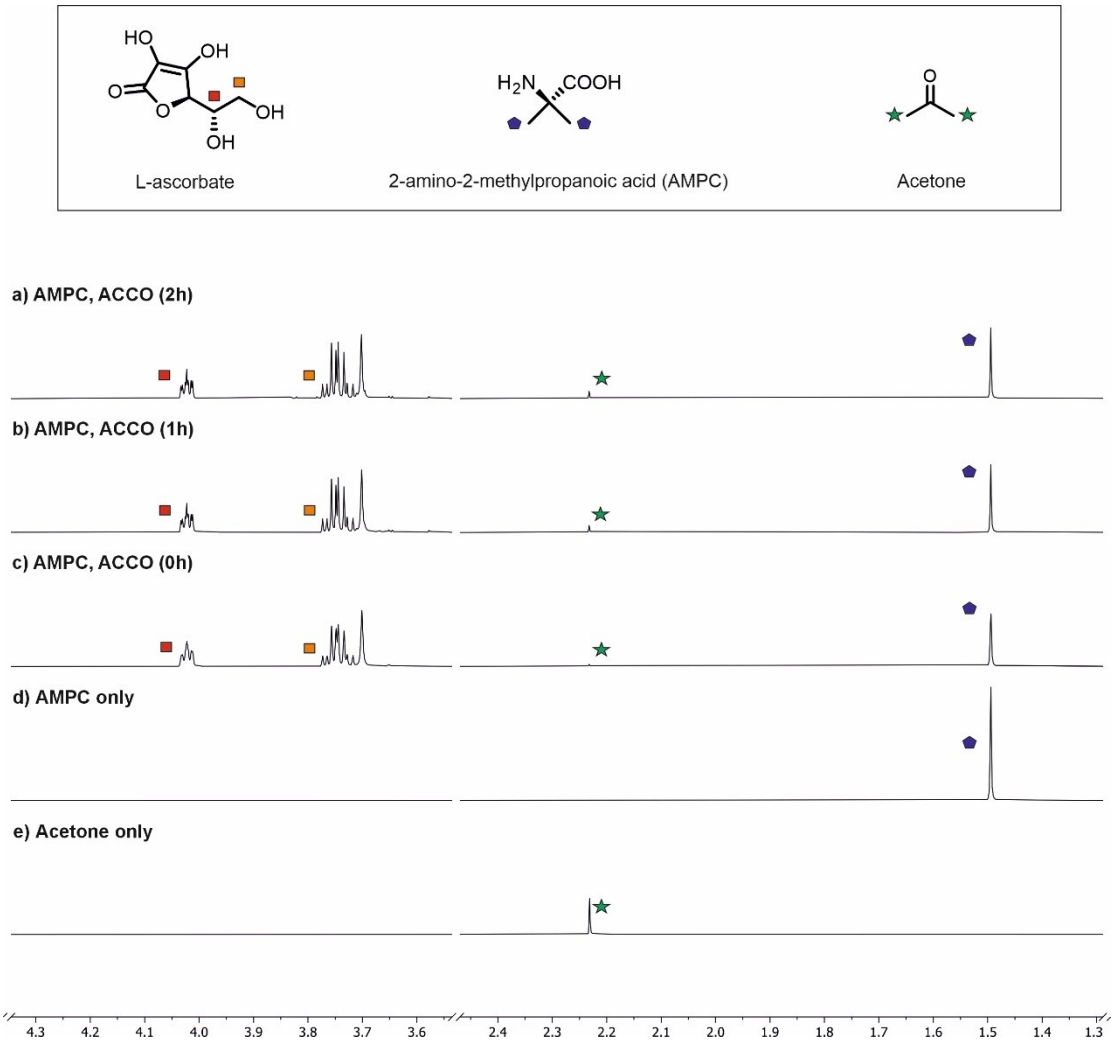


Figure 2.18. Analysis of ACCO catalysis with 2-amino-2-methylpropanoic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of 2-amino-2-methylpropanoic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

The combined observations suggest that AtACCO requires its substrate to contain an amino acid-like scaffold, that is the presence of both a free amino group and a free carboxylic acid group, for productive turnover; the relative arrangement of the amino and carboxylic acid groups determined the reaction outcome. A potential catalytic mechanism was proposed for a ACCO catalysed oxidation of α-dialkylated α-amino acid derivatives (Fig. 2.19):

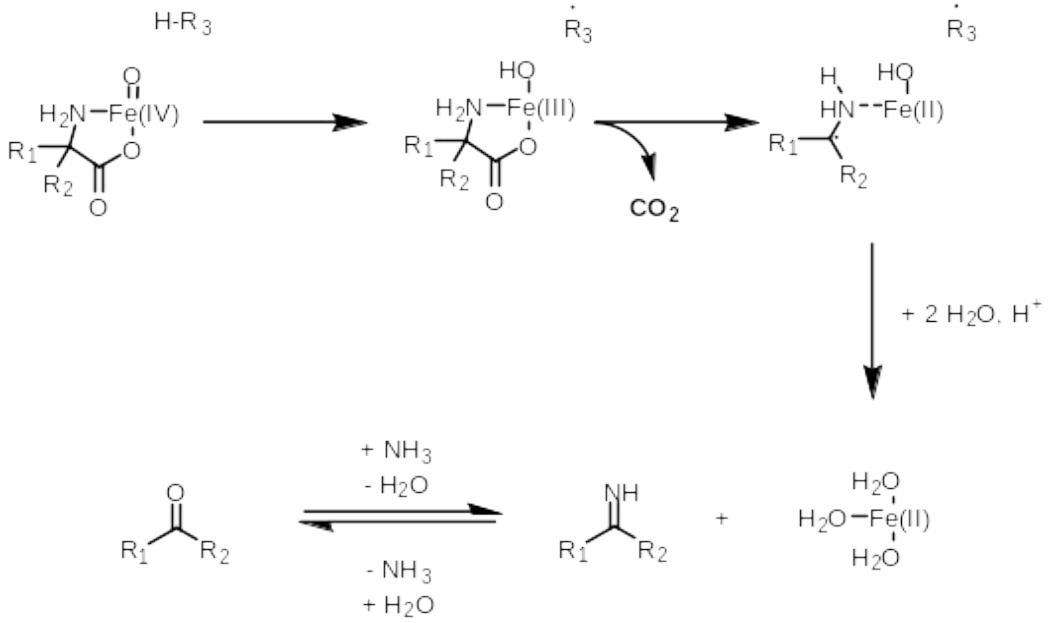


Figure 2.19 Proposed outline mechanism for the ACCO catalysed ketone forming reaction via an α-radical intermediate.

In the AtACCO catalysed ketone-forming reaction, it is proposed that the substrate binds with the amino group and the carboxylic acid coordinating to the AtACCO bound metal. Similar to the mechanism of ACCO catalysing the oxidation of ACC, this leads to the formation of a Fe(IV)-oxo species. (Fig. 2.19) A proton-donating species, potentially the L-ascorbate, would reduce the complex to an Fe(III) species and subsequently trigger the decarboxylation process to form a tertiary carbon radical and reduce the Fe(III) species to an Fe(II) species. The initially formed imine product is then hydrolysed to the corresponding ketone product.

2.6 Structural Studies of ACCO mutants

AtACCO variants were co-crystallised with ACC to investigate the effect of the active site variations on AtACCO conformation, substrate/(co)substrate binding and the water bonding network (Fig. 2.20-23), and to correlate the assay results with current mechanistic proposals. ACCO co-crystallisation was conducted with a commercial broad screen (Morpheus Complete HT-96; AtACCO: 18 mg/mL; $CoCl_2$: 3 mM; ACC: 20 mM). Crystals of four AtACCO variants were obtained, which diffracted leading to K290A, Y91H, Y163F and N217Q AtACCO:Co:ACC complex structures (K290A AtACCO:Co:ACC complex: Space group: P 21212121, resolution: 1.65 Å, Morpheus F5; Y91H AtACCO:Co:ACC complex: Space group: P 21212121, resolution: 1.70 Å, Morpheus G12; Y163F AtACCO:Co:ACC complex: Space group: P 21212121, resolution: 1.75 Å, Morpheus B5; N217Q AtACCO:Co:ACC complex: Space group: P 21212121, resolution: 1.55 Å, Morpheus G8).

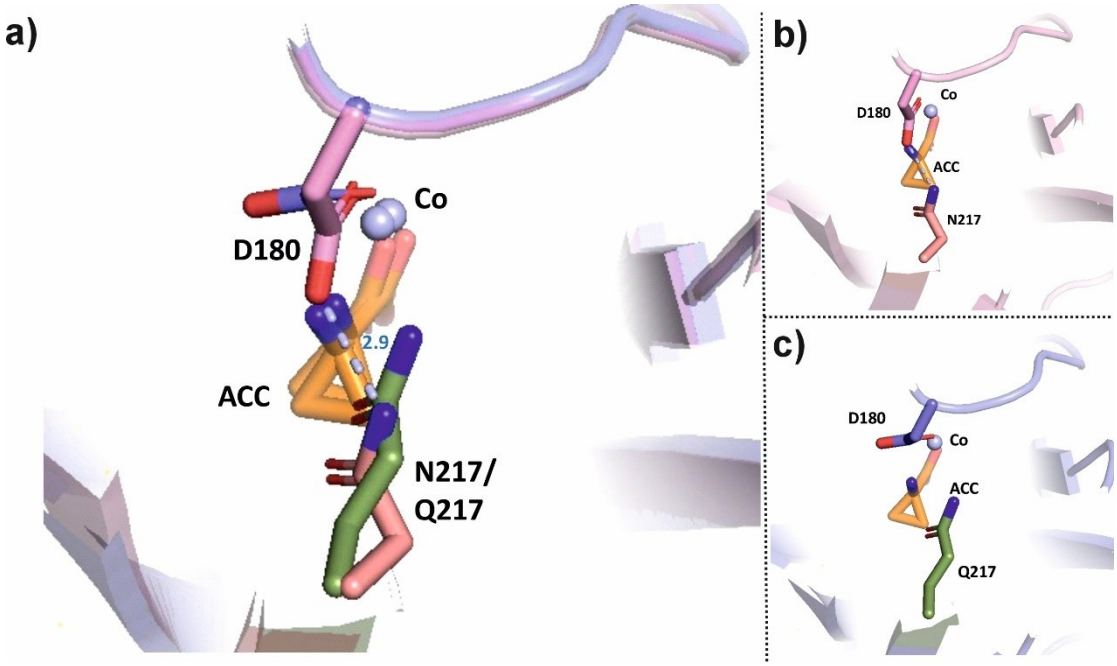


Figure 2.20 Views from co-crystals of wt AtACCO and N217Q AtACCO variants bound to Co (II) and ACC. a) Superposition views from co-crystals of wt AtACCO (pink) and N217Q AtACCO variant (slate) bound to Co (II) and ACC. b) Views from co-crystals of wt AtACCO bound to Co (II) and ACC. c) Views from co-crystals of N217Q AtACCO bound to Co(II) and ACC. Co(II) is in light

blue spheres and ACC is in orange. N217 is shown in salmon and Q217 is shown in green. Hydrogen bonding interactions in complex structure are shown as dashed lines (RMSD = 0.130 Å).

Comparison of the N217Q AtACCO:Co(II):ACC and AtACCO:Co(II):ACC:HCO_3^- complex structures (the latter obtained by Dr. Zhihong Zhang, University of Oxford) reveals that the N217Q variation results in a greater steric hindrance near the metal-binding D180 side chain, resulting in the side chain of D180 occupying an alternative conformation compared to that observed in the wildtype structure (Fig. 2.20). Consequently, the hydrogen bonding interaction between the side chains of N217 and D180 in wtACCO (distance: 2.9Å) is absent in the N217Q AtACCO:Co:ACC complex structure, which may destabilise the Fe(II) binding triad (H178, D180 and H235) resulting in substantially reduced catalytic activity.

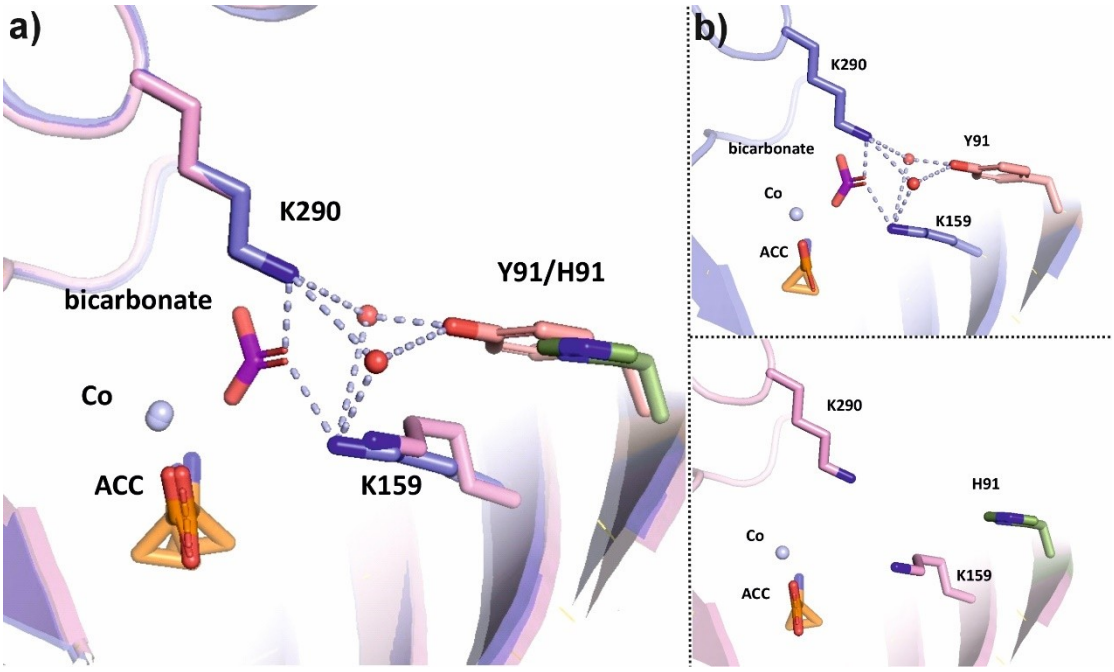


Figure 2.21 Views from co-crystals of wt AtACCO and Y91H AtACCO variants bound to Co (II) and ACC. a) Superposition views from co-crystals of wt AtACCO (slate) and Y91H AtACCO variant (pink) bound to Co(II) and ACC. b) Views from co-crystals of wt AtACCO bound to Co (II) and ACC. c) Views from co-crystals of Y91H AtACCO bound to Co(II) and ACC. Co(II) is in light blue spheres and ACC is in orange. Y91 is shown in salmon and H91 is shown in green. Hydrogen bonding

interactions in complex structure are shown as dashed lines (RMSD = 0.317 Å).

The K290A, Y91H and Y163F AtACCO:Co(II):ACC complex structures did not reveal major conformational changes in the active site compared to the wtAtACCO:Co(II):ACC:HCO_3^- complex structure (Fig. 2.21-23). Nevertheless, these variations induce substantial changes in the water network in the AtACCO binding pocket which may affect catalysis and substrate/co-factor binding affinities. Comparing the Y91H AtACCO:Co(II):ACC and AtACCO:Co(II):ACC:HCO_3^- complex structures reveals that the two water molecules coordinating K159 and K290 to Y91 (described in Chapter 1.4) were absent in the Y91H AtACCO:Co(II):ACC complex structure (Fig. 2.21). By contrast, comparing the K290A AtACCO:Co(II):ACC and AtACCO:Co(II):ACC:HCO_3^- complex structures reveals that only the water molecule coordinating K290A to Y91 was absent in the K290A AtACCO:Co(II):ACC complex structure, whereas the water molecule coordinating K159 to Y91 remains conserved (Fig. 2.22). Comparing the Y163F AtACCO:Co(II):ACC and AtACCO:Co(II):ACC:HCO_3^- complex structures reveals that only one of the two water molecules in the network (described in Chapter 2.4) is conserved, whereas the water molecule coordinating S247 and Y163 to the ACC is not present (Fig. 2.23).

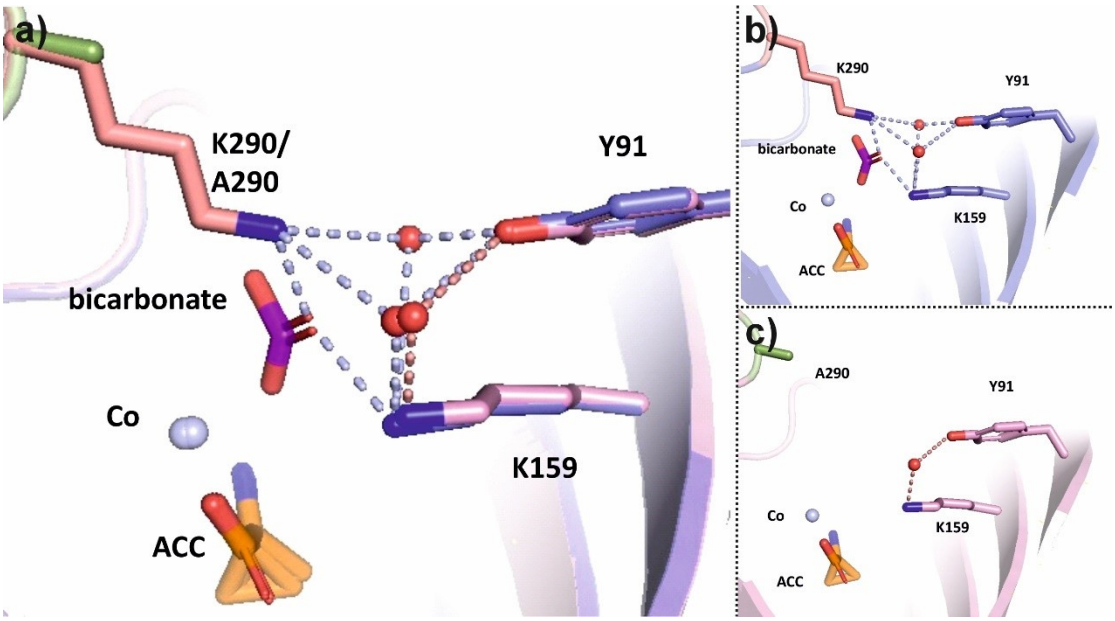


Figure 2.22 Views from co-crystal structures of wt AtACCO and K290A AtACCO variants bound to Co(II) and ACC. a) Superposition views from co-crystals of wt AtACCO (slate) and K290A AtACCO variant (pink) bound to Co(II) and ACC. b) Views from co-crystals of wt AtACCO bound to Co(II) and ACC. c) Views from co-crystals of K290A AtACCO bound to Co(II) and ACC. Co(II) is in light blue spheres and ACC is in orange. K290 is shown in salmon and A290 is shown in green. Hydrogen bonding interactions in complex structure are shown as dashed lines (RMSD = 0.179 Å).

The crystallography data on the AtACCO variants helps to rationalise the AtACCO variant turnover data with ACC (Fig. 2.11a, 2.12). K290A ablates the ability of AtACCO to form a hydrogen bonding interaction with bicarbonate, consistent with the observation that the K290A AtACCO variant only formed trace amounts of ethylene (Fig. 2.12b, described in Chapter 2.4). In contrast, the Y91H AtACCO:Co(II):ACC complex structure suggests that Y91H, though potentially weakening the stability of the triad formed with K159 and K290, does not induce observable conformational changes on K159 and K290, hence retaining the ability of the lysine side chains in assisting bicarbonate binding. This is consistent with the observation that Y91H AtACCO displayed a substantially higher turnover than

K290A ACCO. Similarly, as Y163F AtACCO retains the water network at the RXS motif partially, the turnover for Y163F ACCO is significantly higher than that of K290A ACCO (described in Chapter 2.4).

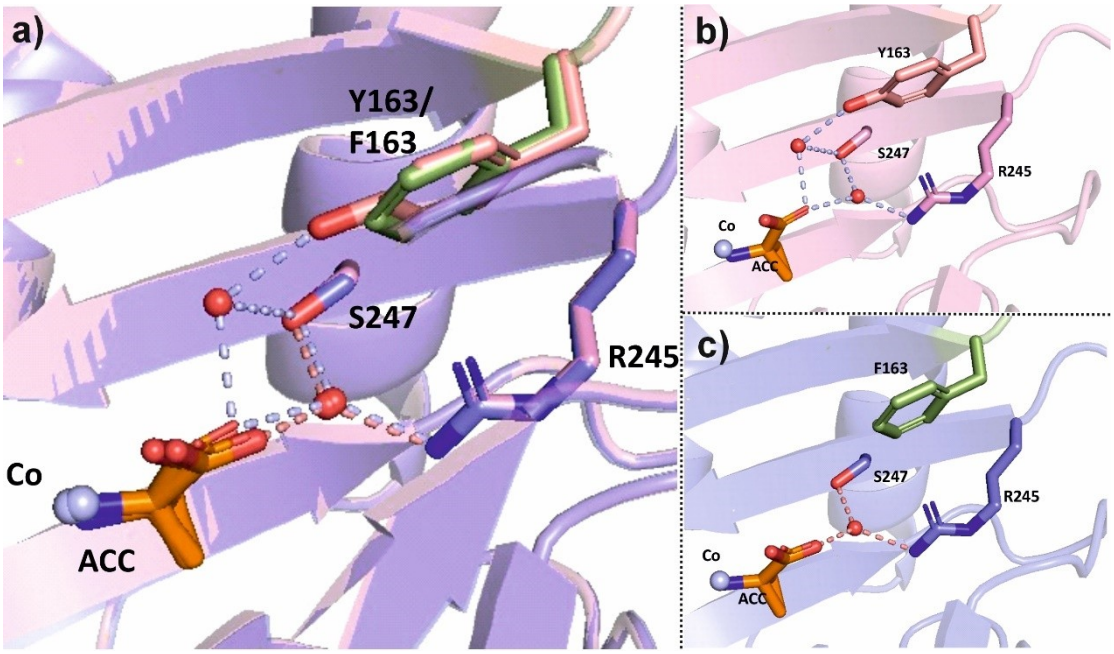


Figure 2.23 Views from co-crystals of wt AtACCO and Y163F AtACCO variants bound to Co (II) and ACC. a) Superposition views from co-crystals of wt AtACCO (pink) and Y163F AtACCO variant (slate) bound to Co(II) and ACC. b) Views from co-crystals of wt AtACCO bound to Co(II) and ACC. c) Views from co-crystals of Y163F AtACCO bound to Co(II) and ACC. Co(II) is a light blue sphere and ACC is in orange. Y163 is shown in salmon and F163 is shown in green. Hydrogen bonding interactions are shown as dashed lines (RMSD = 0.144 Å).

Chapter 3. Broadening ACCO Bio-catalysis

3.1 Introduction

Olefins other than ethylene are important building blocks in industry (described in Chapter 1.1), however, their biological roles are mostly unclear. Development of biocatalytic methods to access these olefins could enable their production in a sustainable manner. So far, no biosynthesis pathway for the production of olefins, such as propylene and butene, has been reported. The work described in Chapter 2 showed that wt AtACCO can catalyse propylene formation from C-methylated ACC derivatives (1-aminomethylcyclopropanecarboxylic acid), albeit not very efficiently (Fig. 2.15), highlighting the potential of engineering ACCO to catalyse the oxidation of C-alkylated ACC derivatives more efficiently.

Among these 1-aminomethylcyclopropane-1-carboxylic acids, (1*S*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid and (1*S*, 2*S*)-1-aminoethylcarboxylic acid, norcoronamic acid and coronamic acid, respectively, are found in *Pseudomonas syringae* as intermediates of the coronatine and norcoronatine biosynthesis pathway,

respectively[57]. (Fig. 3.1) Coronatine is a toxin produced by the bacterium *Pseudomonas syringae*, involved in stomatal reopening after they close in response to pathogen-associated molecular patterns and interference with the responses mediated by salicylic acid upon infection[58]. Norcoronatine was isolated and identified from the liquid cultures of *Pseudomonas syringae*. Norcoronatine reportedly employs a similar biosynthetic pathway as the coronatine[57]. Due to their high degree of structural similarity, coronatine and norcoronatine are hypothesised to be made via a similar enzyme system where L-leucine was reported to be the precursor of a two-step bio-synthesis of coronamic acid. L-leucine undergoes an isomerisation process to alloisoleucine and a subsequent ring-closure process to yield coronamic acid[57]. Similarly, norcoronamic acid was reportedly formed through the same process from L-valine[57].

3.2 Structure-guided protein engineering for bio-catalysis of higher olefins

AtACCO residues facing the four ACC C-H bonds were selected for site-directed mutagenesis (V185, N217, S247, and L196, Fig. 2.9a). V185, N217 and S247 reside on the core ß-sheet while L196 is located on a flexible loop. It was anticipated that by substituting these residues to sterically smaller ones, the AtACCO substrate binding pocket may be extended to accommodate the relatively bulky ACC C-alkyl group.

Initially the L196A AtACCO variant was produced and purified to test this proposal. In comparison to wt AtACCO, L196A AtACCO catalysed turnover of ACC occurred with only ~14% the efficiency of the wildtype ACCO, likely due to the reduced hydrophobic interactions between the binding site and ACC (Fig. 3.1a). In contrast, L196A AtACCO catalysed conversion of (1*R*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid to propene ~51% more efficiently than wt AtACCO and also catalysed the conversion of (1*R*, 2*R*)-1-aminoethylcyclopropane-1-carboxylic acid to 1-butene with ~10% substrate depletion, highlighting the feasibility of the engineering approach. (Fig. 3.2)

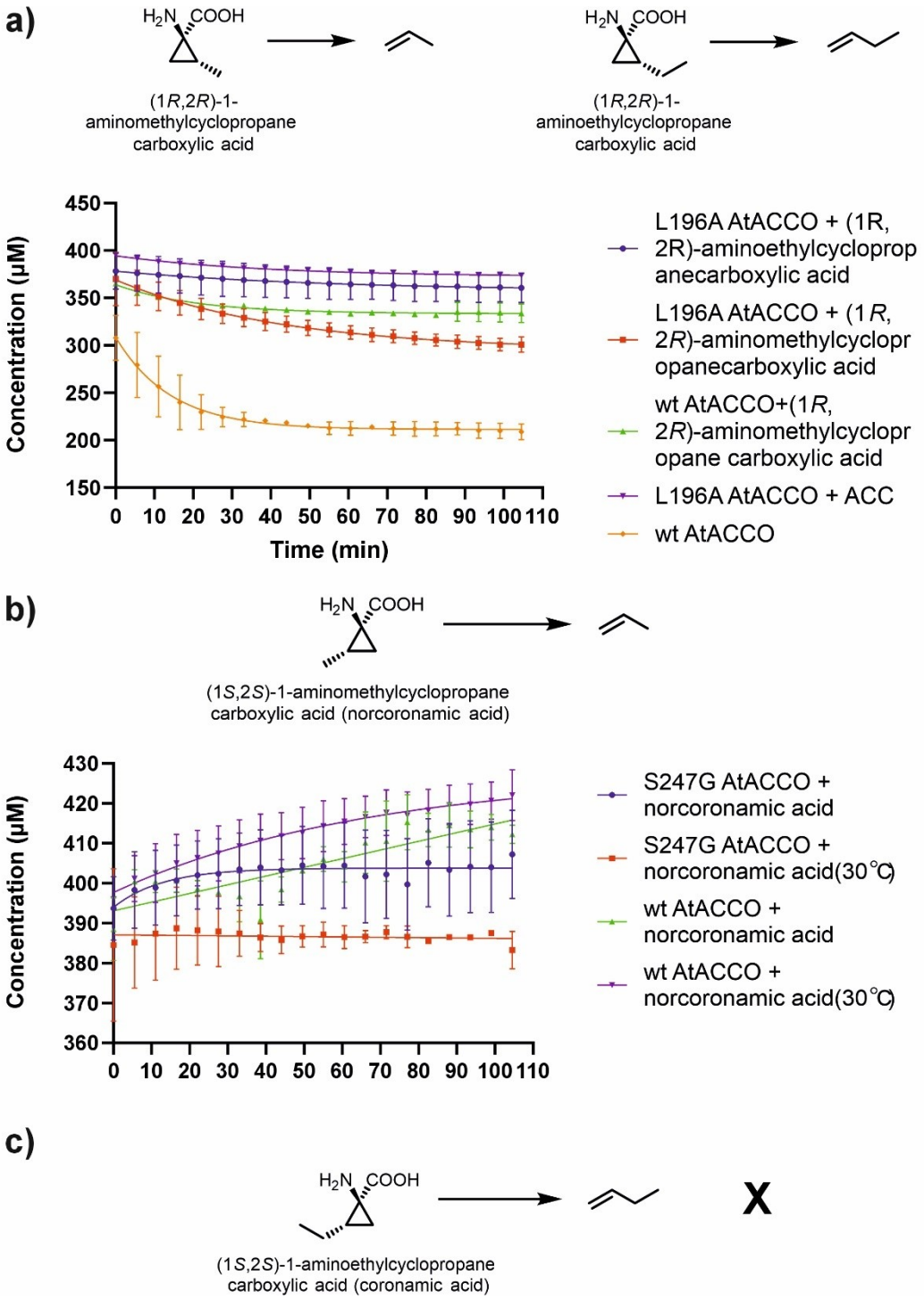


Figure 3.1 a) wt AtACCO and L196A AtACCO L196A with ACC or ACC derivatives ((1*R*, 2*R*)-aminomethylcyclopropanecarboxylic acid and (1*R*, 2*R*)-aminoethylaminocyclopropanecarboxylic acid) (Error: n = 2); b) wt AtACCO or AtACCO S247G catalysis of (1*S*, 2*R*)-aminomethylcyclopropanecarboxylic acid (norcoronamic acid); (Error: n = 2); c) (1*S*, 2*R*)-aminoethylcyclopropanecarboxylic acid (coronamic acid) cannot be catalysed by wt AtACCO or S247G AtACCO into 1-butene. (n=3, experiment replicate, error bar = standard error)

To increase turnover, it was investigated whether decreasing the steric bulk of L196 further than achieved with the L196A AtACCO variant would be productive. However, the L196G AtACCO variant was not expressed well, possibly due to fold destabilisation through generation of a hydrophobic void.

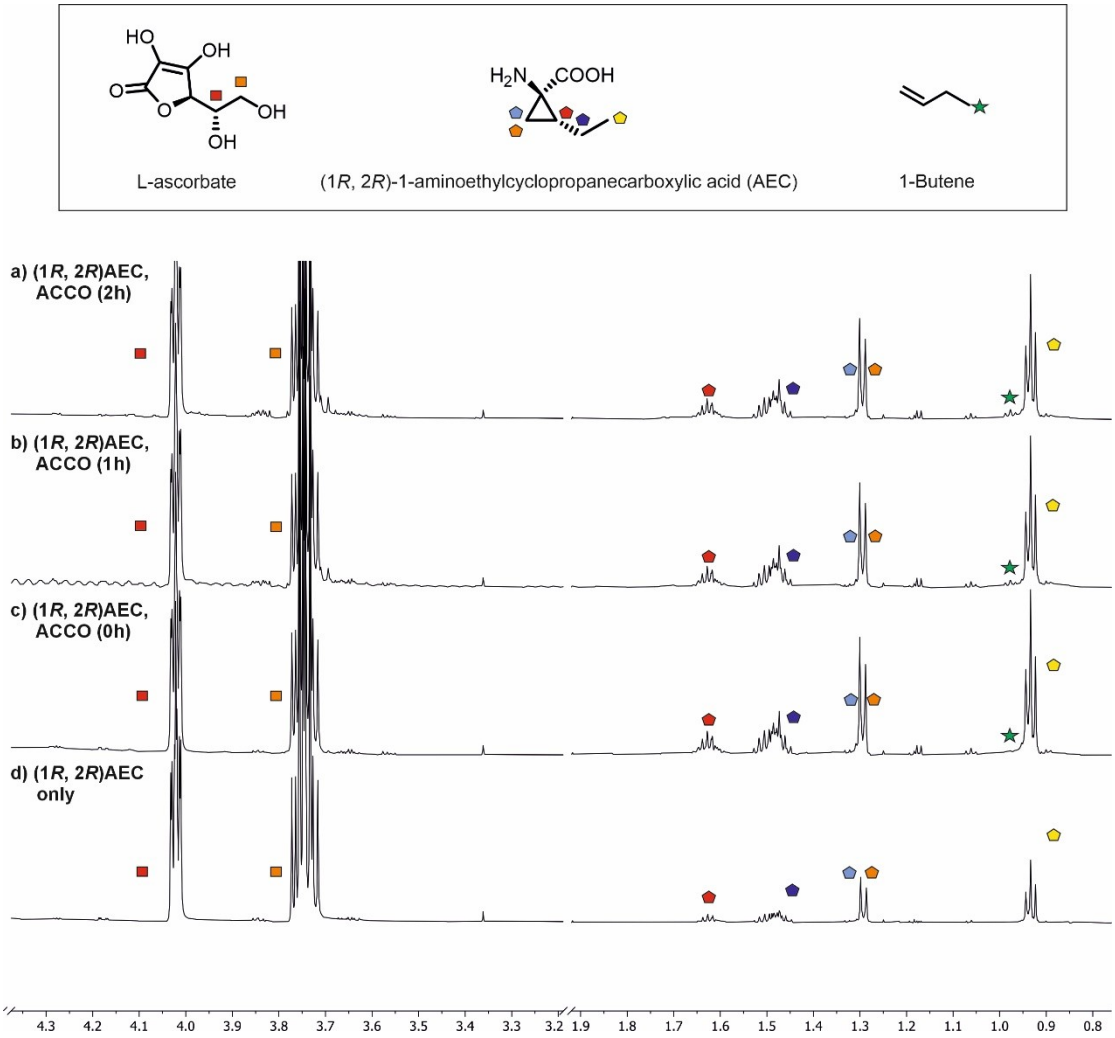


Figure 3.2 Analysis of L196A ACCO catalysis with (1*R*, 2*R*)-1-aminoethylcyclopropanecarboxylic acid as a substrate to generate 1-butene. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of (1*R*, 2*R*)-1-aminoethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

Although the mutagenesis studies implied an important role for S247 in AtACCO catalysis (Fig. 2.11a), S247G AtACCO variant was produced to create hydrophobic void and with the aim of enabling the more efficient binding of an ACC C-alkyl group. ^{1}H NMR assays revealed that the S247G AtACCO was substantially less efficient in catalysing ACC turnover to ethylene than wt AtACCO (described in Chapter 2.4). By contrast, both the S247G AtACCO variant and wt AtACCO showed similar activities with norcoronamic acid (Fig. 3.1b). Given *Amborella trichopoda* originated from tropical region in the Pacific Ocean, the reaction temperature was increased by 5 degrees Celsius to 30 °C to mimic physiological conditions. At elevated temperature, the S247G AtACCO variant catalysed propene formation from norcoronamic acid ~3-fold more efficiently than wt AtACCO (Fig. 3.1b). However, it did not employ the coronamic acid as a substrate, possibly reflecting the insufficient hydrophobic void created by mutation S247G.

To investigate turnover of coronamic acid into 1-butene, ACCO double variants were engineered based on the AtACCO:Co(II):ACC complex crystal structure. The side chains of L187 and R245 are in close proximity to that of S247; L187 was selected to investigate variations aiming at reducing its size on catalysis due to the substantially reduced activity of the R245Q AtACCO variant (as described in Chapter 2.4). The L187G, L187A/S247G and L187A/S247A AtACCO variants were produced and their ability to accept coronamic acid as a substrate was tested using ^{1}H NMR assays. The results revealed that none of these AtACCO variants produced detectable amounts of

1-butene. This lack of activity may be a result of disrupting the water network by substituting the S247 side chain. Therefore, the L187S/S247G AtACCO variant was designed to restore the water network. However, L187S/S247G AtACCO did also not produce 1-butene.

Chapter 4. Structural Studies on PsEFE

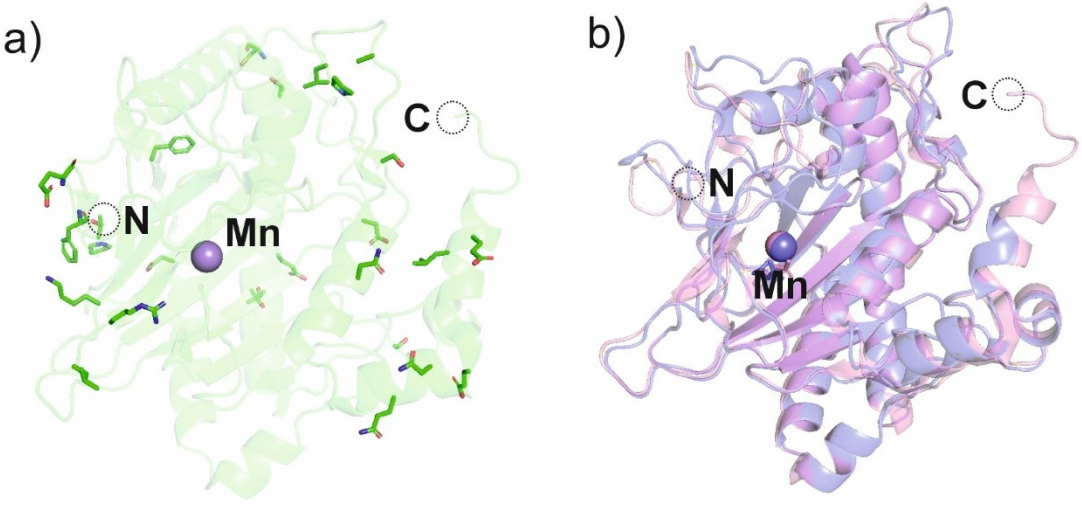
4.1 Introduction

It was of interest to assess the ability of PsEFE to enable biocatalytic propylene formation, given the relatively high levels of propylene formation from C-methylated ACC derivatives (Fig. 2.14-15, described in Chapter 2.5). PsEFE catalyses the bifurcating oxidation of 2OG to generate variable ratios of succinate and ethylene. This is in a sharp contrast to the ethylene-forming enzyme in plants, ACCO, which gives ethylene together with cyanide. During the course of this work, a study showing that PsEFE active site variants catalyse formation of propylene from 4-methyl-2OG based on the reported ability of PsEFE to accept C3- and/or C4-methyl substituted 2OG derivatives as substrates was reported[48]. Due to the absence of structural information, it was, however, unclear how C3- and/or C4-methyl substituted 2OG derivatives bind to PsEFE. The versatile reported 2OG binding modes of PsEFE complicated prediction on the binding of 2OG derivatives to the active site, hence,

binding of the 2OG derivatives to the PsEFE active site was investigated using protein crystallography.

Figure 4.1 Sequence alignment between PK2 PsEFE, glycinea PsEFE and 1449B PsEFE.

A crystal structure of the characterised *P. savastanoi pv. phaseolicola* PK2 PsEFE (abbreviated as PK2 PsEFE) in complex with 2OG derivatives could not be obtained. Therefore, bioinformatic analysis were conducted by S. Dhingra (Schofield group, University of Oxford) to identify PsEFE orthologs that crystallise more robustly. The results revealed that *P. savastanoi pv. glycinea (*abbreviated as glycinea PsEFE) *and P. savastanoi pv. phaseolicola* 1449B (abbreviated as 1449B PsEFE) have ~99.7% and ~88.3% sequence identity, respectively, with the well characterised *P. savastanoi pv. phaseolicola* PK2 (Fig. 4.1). In comparison with PK2 PsEFE, glycinea PsEFE differs by only one residue, that is the R326S substitution (Fig. 4.1). 1449B PsEFE, on the other hand, differs in 41 amino acids from PK2 PsEFE; given that these residues are predicted to localise mostly on the protein surface, they could alter the physical



properties (including crystallisation properties) (Fig. 4.2).

Figure 4.2 Structural differences between PK2 PsEFE (PDB: 5MOF) and 1449B PsEFE (unpublished data) as observed by crystallography. a) Distribution of the differing mutants between PK2 PsEFE and 1449B PsEFE. b) Structural superimposition between PK2 PsEFE and 1449B PsEFE. The Mn is displaced in purple spheres. N-terminus (N) and C-terminus (C) are highlighted with circles.

4.2 Expression and Purification of PsEFE

A plasmid encoding for the glycinea PsEFE was generated by site-directed mutagenesis from the pET28a(+)-SUMO plasmid encoding for PK2 PsEFE. Recombinant protein production and purification were performed as described for AtACCO in Chapter 2.2. The pET28a(+)-SUMO plasmid contains a small ubiquitin-like modifier (SUMO) domain which is located between the *His*$_6$-tag and the PsEFE gene sequence. A SUMO-specific protease 2 (SenP2) was used to cleave SUMO domain from the protein of interest at the C-terminus of the SUMO domain. SDS-page and mass spectrometry were used to confirm the identity and purity of the purified PsEFE.

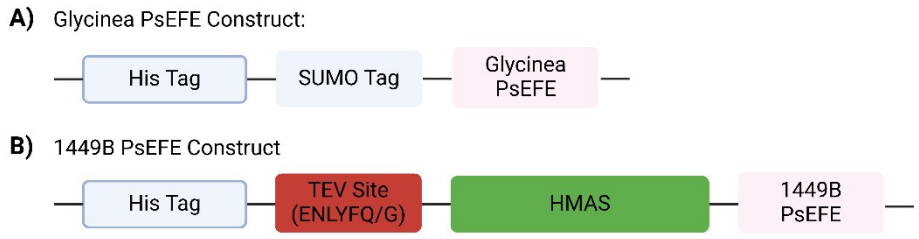
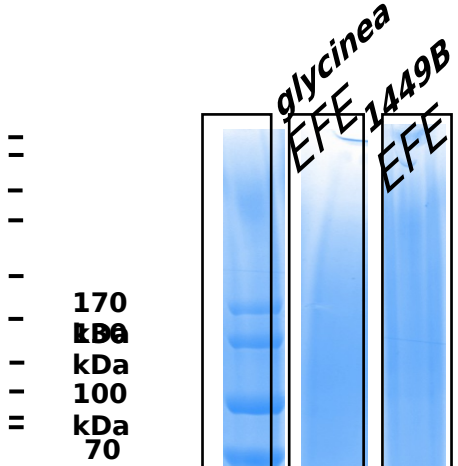


Figure 4.3 Constructs for plasmid PsEFE from a) Glycinea and b) 1449B

Plasmids encoding for the 1449B PsEFE were purchased from Genscript where the protein of interest is encoded on a pET28a(+)-TEV vector. Recombinant protein production and purification were performed as described in Chapter 2.2. (Fig 4.3) The pET28a(+)-TEV plasmid encodes a TEV protease recognition site (ENLYFQ/G; cleavage between Q and G) which is located between the His_6-tag and the N-terminus of 1449B PsEFE. This allows the removal of the *His_6*-tag during the purification process. Four additional amino acids (HMAS) were inserted between the TEV recognition site and the NheI restriction site used for cloning. This 4-amino-acid sequence was inserted to improve the exposure of the TEV cleavage site to the TEV protease and enable cloning of the protein sequence into the commercial pET28a-TEV vector where the protein sequence will not contain an identical restriction site to that for cloning. The sequence was determined by the manufacturer, Genscript. Along with the glycine in the TEV cleavage site, a total of 5 additional amino acids are present at the N-terminus. The protein mass acquired was thus 501 Da greater than the theoretical protein mass matching the 1449B PsEFE sequence alone. Mass spectrometry analysis confirmed the identity and purity of the purified 1449B PsEFE (Fig. 4.4-4.6).



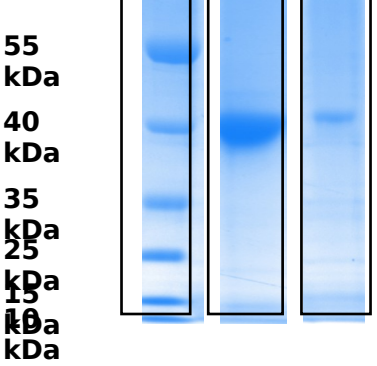


Figure 4.4 SDS-PAGE gel display of protein expression for glycinea EFE and 1449B EFE.

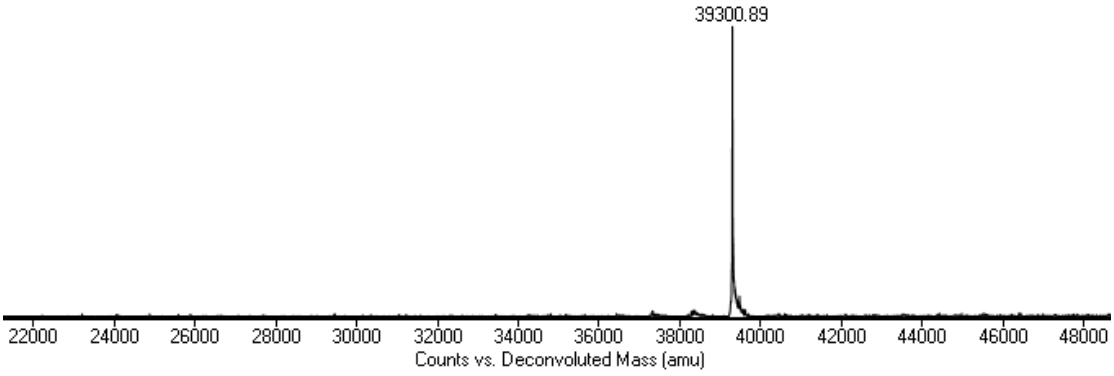


Figure 4.5. Deconvoluted mass spectrum of Glycinea PsEFE. Calculated mass: 39375.52 Da; Observed mass: 39300.89 Da.

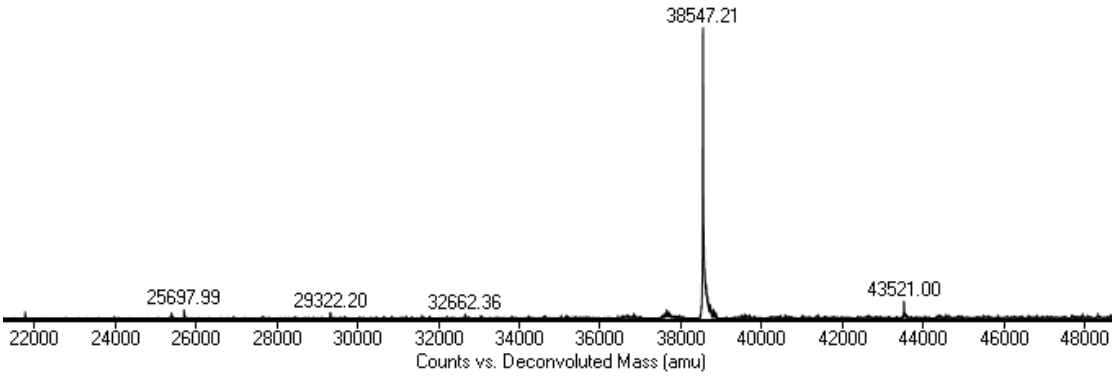


Figure 4.6. Deconvoluted mass spectrum of 1449B PsEFE. Calculated mass: 38547.82 Da; Observed mass: 38547.21 Da.

4.3 Binding modes of 2OG/2OG-derivatives to PsEFE

Activity tests revealed that both glycinea and 1449B PsEFEs catalysed conversion of 2OG into ethylene (Fig. S36-37). Similar to PK2 PsEFE, which did not crystallise reliably under the reported conditions (11 mg/mL; 2OG: 10 mM; Mn: 3 mM) and all the other conditions additionally tested, glycinea PsEFE, which is highly similar to PK2 PsEFE, did not crystallise under the trialled conditions using commercial broad screens. By contrast, the crystallisation trials using the tested commercial broad screens revealed that 1449B PsEFE crystallised using a condition from the PEG/Ion screen kit (H4; buffer condition: 0.03 M citric acid, 0.07 M bis-tris propane (BTP) pH7.6, 20% $_{w/v}$ PEG3350; crystallisation: 1449B PsEFE: 11 mg/mL; 2OG: 10 mM; Mn: 3mM). This condition was further optimised, varying the PEG3350 concentration

from 18% $_{w/v}$ to 23% $_{w/v}$. Single crystals suitable for X-ray diffraction studies were obtained at each tested PEG3350 concentration.

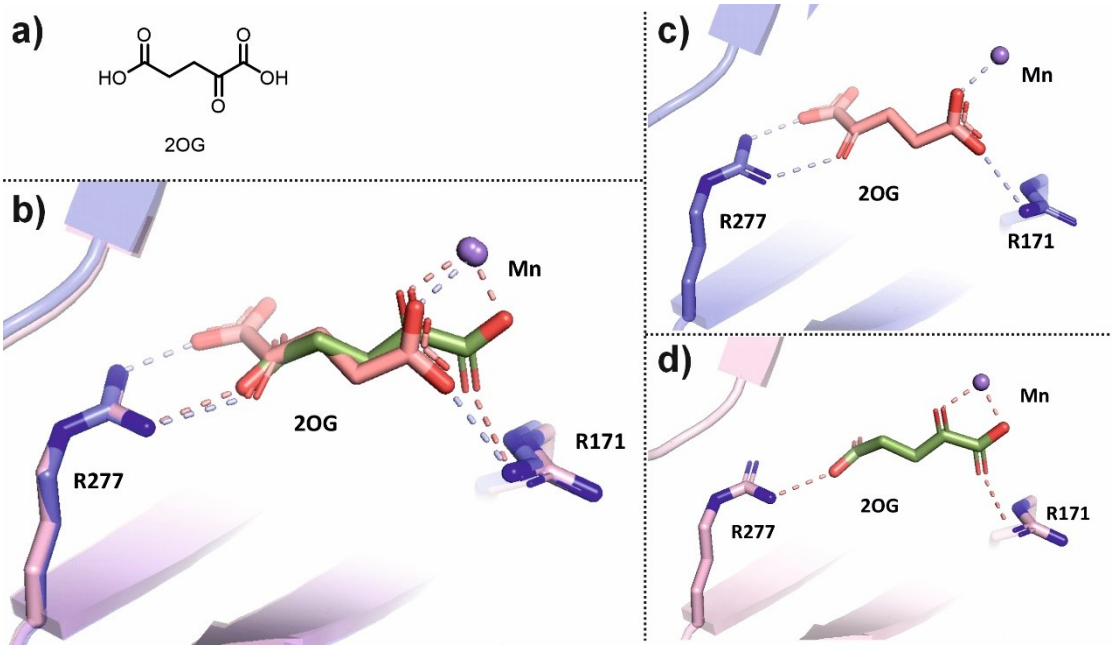


Figure 4.7 Structures of PsEFE co-crystallised with 2OG with/without L-arginine. a) 2OG structure; b) Structural superposition of PK2 PsEFE:2OG:Mn(II):L-arg complex structure (PDB: 5MOF, pink) with 1449B PsEFE:2OG:Mn (slate). (RMSD = 0.415 Å); c) Views of 2OG binding mode in 1449B PsEFE:2OG:Mn(II); d) Views of 2OG binding mode in PK2 PsEFE:2OG:Mn(II):L-arg complex structure. Mn(II) is shown in a light blue sphere.

A 1449B PsEFE crystal structure was obtained using X-ray diffraction. The crystal structure was solved (space group: P 43; resolution: 1.65 Å) using molecular replacement based on the protein structure predicted by AlphaFold3[59] (conducted by Dr. Mark D. Allen, University of Oxford). The resultant 1449B PsEFE structure was similar to that of the PK2 PsEFE structure (RMSD = 0.415 Å). The structure of the 1449B PsEFE active site pocket is apparently well conserved. (Fig. 4.7) The 41 residues differing between PK2 PsEFE and 1449B PsEFE are all located on the

protein surface (Fig. 4.3a), potentially helping to improve the crystallisation properties. Consistent with the reported PK2 PsEFE:Mn(II):2OG complex structures, analysis of the 1449B PsEFE:Mn(II):2OG complex structure suggested that the 2OG substrate adopted an inverted monodentate binding mode. In this mode, the 2OG C5 carboxylic acid was bound to the metal centre and its binding was stabilised by the side chain interactions with R171, while the 2OG C1 carboxylate and C2 keto group were bound to the R277 guanidinium group through electrostatic interaction (Fig. 4.7). The comparisons of binding modes of these 2OG derivatives suggested the distorting effects of substituents on the substrate binding. Note that no evidence was accrued that 2OG occupies an alternative conformation in the 1449B PsEFE:Mn(II):2OG complex structure, whereas it occupies two conformations, with 2OG being both monodentate and bidentate binding to the metal centre, in the reported PK2 PsEFE:Mn(II):2OG complex structure[49].

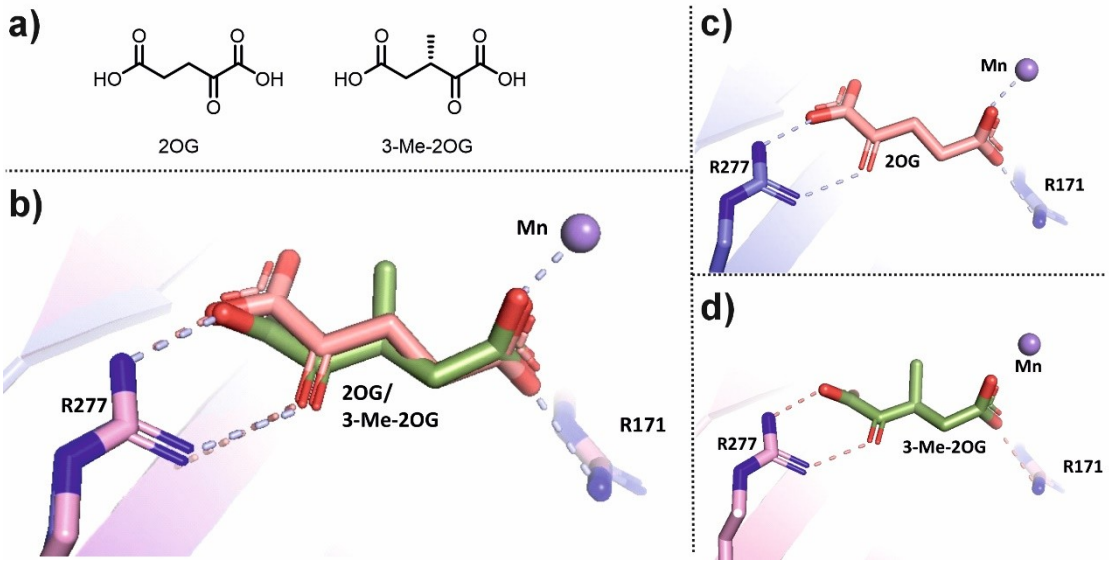


Figure 4.8 Structures of PsEFE co-crystallised with 2OG or 3-methyl-2OG. a) Structures of 2OG and 3-methyl-2OG (3-Me-2OG); b) Structural superposition 1449B PsEFE:2OG:Mn(II) (slate) with

1449B PsEFE:3-Me-2OG:Mn(II) (pink).(RMSD = 0.106 Å); c)Views of 2OG binding mode in 1449B PsEFE:2OG:Mn(II); d) Views of 2OG binding mode in 1449B PsEFE:3-Me-2OG:Mn(II) complex structure. 2OG is shown in salmon and 3Me2OG is shown in green. Mn(II) is shown in a light blue sphere.

1449B PsEFE:Mn(II):3-methyl-2OG (3-Me-2OG) complex structure (space group: P 43, resolution: 2.01 Å, Fig. 4.8) suggested the methyl group on C3 did not cause significant distortion on 2OG backbone compared to 2OG where the backbones of 2OG and 3-Me-2OG are mostly coherent (Fig. 4.8b). The C1 carboxyl and C2 keto both interact with R277 through hydrogen bonds while the C5 carboxylate binds to the metal in a monodentate manner. However, an increased steric bulk of the substituent on the C3 causes an increasing distortion on the binding mode of the substrate. The 3-ethyl-2OG (3-Et-2OG):Mn(II):1449B PsEFE complex structure (space group: P 43, resolution: 1.79 Å, Fig. 4.9) revealed an altered conformation of the co-substrate C1-C2 bond, apparently leading to the loss of hydrogen bonding between the 3-Et-2OG C2 keto group and the R277 side chain causing the C2 keto oxygen to be displaced by 2.8 Å. This flip in the substrate binding mode enabled formation of a salt-bridge of the co-substrate C1 carboxylate with the R277 guanidinium group (Fig. 4.9).

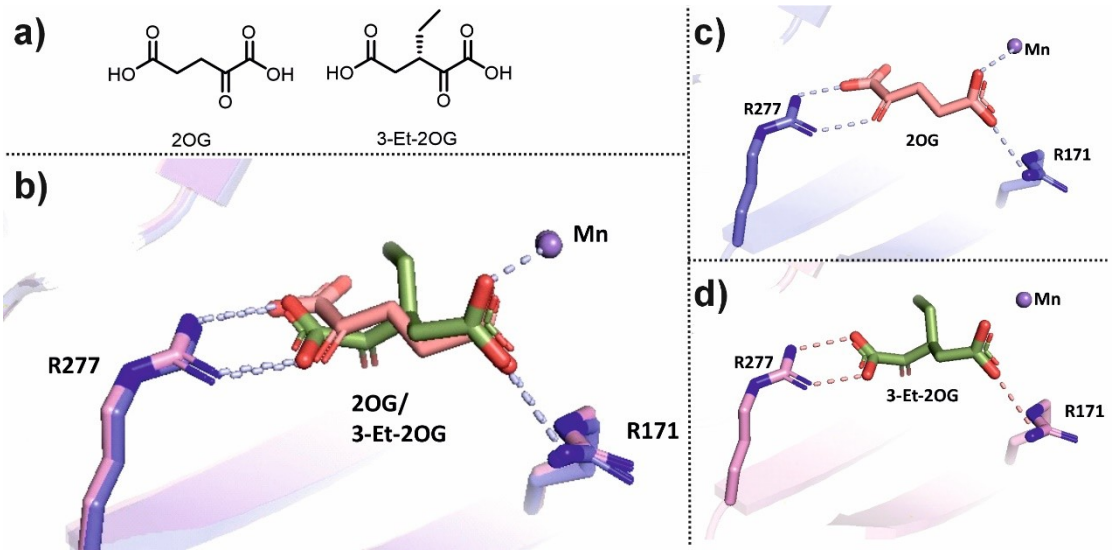


Figure 4.9 Structures of PsEFE co-crystallised with 2OG or 3-ethyl-2OG. a) Structures of 2OG and 3-ethyl-2OG (3-Et-2OG); b) Structural superposition 1449B PsEFE:2OG:Mn(II) (slate) with 1449B PsEFE:3-Et-2OG:Mn(II) (pink).(RMSD = 0.260 Å); c)Views of 2OG binding mode in 1449B PsEFE:2OG:Mn; d) Views of 2OG binding mode in 1449B PsEFE:3-Et-2OG:Mn complex structure. 2OG is shown in salmon and 3Et2OG is shown in green. Mn(II) is shown in a light blue sphere.

The addition of a methyl group on the 2OG C4 position manifests apparent distorting effects on the conformation of the carbon scaffold. The 1449B PsEFE:Mn(II):4-methyl-2OG (4-Me-2OG) complex structure (space group: P 43, resolution: 1.44 Å, Fig. 4.10) shows that the distance between the R277 side chain and the C2 keto group is increased from 3.1 Å to 3.5 Å. Hence the hydrogen bonding interaction between imine nitrogen of the guanidinium of R277 and the C2 keto oxygen on 2OG is weakened. The substitution on C4 results in a significant displacement of C4 carbon by 1.5 Å. (Fig. 4.10b) Nevertheless, the monodentate binding mode of the C5 carboxyl was similar to that observed for 2OG in the 1449B PsEFE:Mn(II):2OG complex structure (Fig. 4.10c-d). The crystallography data collectively suggest that the increasing bulkiness of the substituents on the 2OG scaffold plays an important

role in perturbing the binding mode of 2OG.

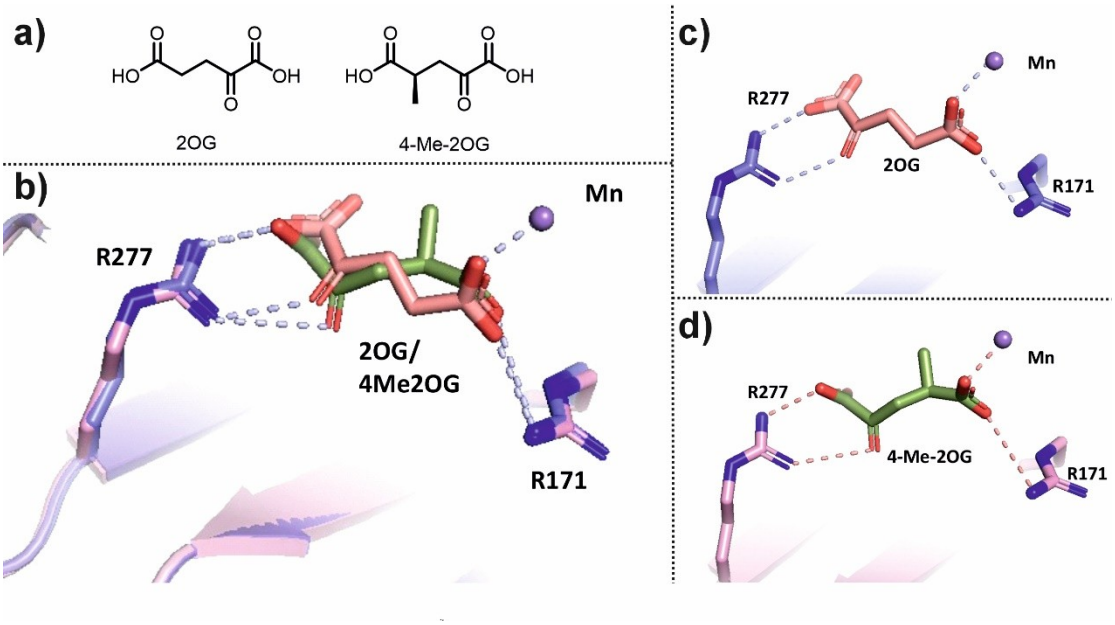


Figure 4.10 Structures of PsEFE co-crystallised with 2OG or 4-methyl-2OG. a) Structures of 2OG and 4-methyl-2OG (4-Me-2OG); b) Structural superposition 1449B PsEFE:2OG:Mn(II) (slate) with 1449B PsEFE:4-Me-2OG:Mn(II) (pink).(RMSD = 0.176 Å); c)Views of 2OG binding mode in 1449B PsEFE:2OG:Mn(II); d) Views of 2OG binding mode in 1449B PsEFE:4-Me-2OG:Mn(II) complex structure. 2OG is shown in salmon and 4Me2OG is shown in green. Mn(II) is shown in a light blue sphere.

It is not yet clear to what extent this unusual 2OG (derivative) binding mode reflects the reactivity of PsEFE and its ability to produce ethylene. However, it is reported that crystallisation in the presence of the Arg substrate alters the 2OG conformation in complex with PK2 PsEFE[49]. Further work needs to investigate how substituting Mn (II) by Fe(II) affects the conformation of 2OG in complex with PsEFE.

Chapter 5. Summary and Future Prospects

The biochemical mutagenesis and crystallographic studies described here provide insight into how key ACCO residues enable the binding of ACC, Fe(II), bicarbonate and L-ascorbate at the active site. Note up to the stage of the O_2 reaction, all binding steps are reversible, hence the order of binding of ACC and, in particular, bicarbonate and L-ascorbate may differ. Importantly, the results show that the L-ascorbate and bicarbonate likely bind to the AtACCO active site at the position predicted based on recent AtACCO:L-ascorbate/bicarbonate complex crystal structures (Fig. 1.8). This is an important observation, because it is under active scientific debate as to whether L-ascorbate directly binds to ACCO to enable ethylene formation from ACC or whether it assists ACCO catalysis in vitro by merely serving as a reducing agent in solution to maintain the Fe (II) oxidation state under the catalytic conditions. The results thus suggest the possibility that L-ascorbate binds to the active site of cupin oxygenases other than ACCO that require it for activity, to directly modulate catalysis inter alia of human prolyl residue hydroxylases shown to require L-ascorbate, at least *in vitro*.

The results are consistent with the proposed roles of L-ascorbate both in promoting reaction and in completing ACCO reaction cycles[42]. They also support the proposed binding mode of bicarbonate involving Fe-coordination, two lysines (K159 and K290), and hydrogen bond network formed by two waters and Y91. (Fig. 2.21, as discussed in Chapter 2.6) Although variations are possible, the overall results support

a preferred mechanism in which ACC binds first to the active site come on followed by bicarbonate and/or L-ascorbate given the binding positions of bicarbonate and L-ascorbate obtained from the crystallography data. The role of bicarbonate maybe to help order the active site preparing it for O_2 coordination which occurs with displacement of bicarbonate. The work thus helps to improve the current understanding on the ACCO mechanism, providing a basis for future work to investigate the exact sequences of metal, co-factors, and ACC binding, together with the conformational changes associated with these binding events.

Investigations on the AtACCO substrate scope revealed an alternative reaction outcome with ketones being formed from some, but not all, ACC derivatives bearing larger cycloalkyl scaffolds than cyclopropane. (discussed in Chapter 2.5) This observation reveals the potential of ACCO for biocatalysis applications beyond ethylene formation, including some of the tested linear α-amino acids were transformed into ketones and ammonia, the formation of which from atmospheric nitrogen accounts for a substantial amount of global energy consumption[60,61].

Notably, AtACCO also catalysed propene formation from C-methylated ACC derivatives, some of which are natural products (norcoronamic acid and coronamic acid, Fig. 3.1b-c, Chapter 3). This observation is of interest, not only because of the industrial relevance of propene and the potential impact of biocatalysis routes enabling propene formation from renewables may have on global warming/energy

consumption, but also because the biological effects of propene are only poorly understood compared to those of ethylene. The potential effect of propene on ethylene receptor and other cellular receptors, hence the signalling effects in cells, are yet to be understood. Thus, for the first time it was shown that naturally occurring small-molecules can be efficiently transformed into propene via an enzyme catalysed process. Future studies could include the investigations of the effects of propylene on plant growth, besides being a competitor of ethylene action[62].

The results may serve as a starting point for further improving the catalytic activity of AtACCO, including by directed evolution. AtACCO variants with improved catalytic activity for norcoronamic acid turnover can be subjected to rounds of directed evolution to generate variants also catalysing the oxidation of coronamic acid. Having more active ACCO variants available, the effect of propene formation in bacteria can also be investigated, since bacteria naturally produce biosynthetic precursors of coronamic acid and norcoronamic acid, L-leucine and L-valine[57]. Enzymes involved in the biosynthesis of coronamic and norcoronamic acid and the cyanide detoxification enzyme can be encoded on plasmids and co-transfected to bacterial cells for continuous generation of higher olefins which can be monitored using e.g., gas chromatography-based assays.

The 1449B PsEFE:Mn(II):2OG/2OG derivative complex crystal structures obtained as part of this work reveal that the modes of 2OG binding to 2OG oxygenases may be

more diverse than currently perceived, including 2OG adopting a monodentate binding mode to the metal centre via the C5 carboxylate as opposed to a typical bidentate binding mode to the metal centre via the C1 carboxylate and C2 ketone. (Fig. 4.8-10) It will be of interest to investigate whether the unusual 2OG conformations in the active sites of 2OG oxygenases can also be crystallographically observed using Fe(II) as the cofactor rather than Mn(II). Although Martinez *et al.* reported that the presence of the PsEFE substrate L-arginine leads to significant conformational changes in the PsEFE active site which affects the 2OG binding mode/conformation[49], it is possible that the reaction outcome of PsEFE catalysis may reflect the conformation of 2OG in the active site.

In conclusion, the combined results provide new insight into the biosynthesis of ethylene, a remarkable biological signaling molecule. Although the currently attainable levels of biological ethylene production are likely far from being commercially viable, this situation may change in the future. The results also highlight the potential of engineering of ethylene form enzymes to produce alkenes and other commercially valuable chemicals.

Chapter 6. Materials and Methods

6.1 General

All chemicals were obtained from commercial sources (Sigma Aldrich, unless stated otherwise) and were used without further purification, unless stated otherwise. C-Methylated and C-ethylated ACC derivatives were synthesised by Dr. Xin Wen (University of Oxford). All DNA oligo primers were obtained from Merck (primer sequences are shown in Chapter 7.2). A pET-28a (+)-TEV plasmid encoding for *Amborella trichopoda* aminocyclopropanecarboxylate oxidase (AtACCO; EC: 1.14.17.4; UNIPROT ID: W1NXW4), codon optimised for expression in *E. coli* was obtained from Genscript.

6.2 Protein Expression and Purification

6.2.1 Bacterial Growth Media Preparation

The bacterial growth medium (2xYT) was prepared from commercially available 2xYT broth mixture (Melford). The media (31g) was dissolved per litre of sterile-filtered water and autoclaved at 121℃ for 15 min. 2xYT agar plates were prepared by

dissolving the 2xYT broth mixture (31 g/L) together with agar (10 g/L), followed by sterilising using an autoclave at 121℃ for 15 min.

6.2.2 Bacterial Growth

Protein encoding plasmids were transformed into T7 Express LysY competent cells (New England Biolabs). Cells were grown in 2xYT media containing kanamycin (30 μg/mL) at 37℃ with agitation (180 rpm), until the OD_{600} of the cell culture reached ~0.8. Recombinant protein production was then induced by the addition of isopropyl-β-thiogalactopyranoside (IPTG, final concentration: 0.5 mM). Cells were incubated at 18 ℃ for 16 h before being harvested by centrifugation (11325 x g, 10 min, 4℃. The resultant cell pellet was stored at -80℃.

6.2.3 Protein Purifications

The following general procedure was used. A frozen cell pellet was resuspended in cold (4℃) binding buffer (20 mM imidazole, 500 mM NaCl, 20 mM Tris, pH 7.8) and lysed at 4℃ using a high-pressure cell disrupter (Continuous Flow Cell Disruptor CF1; Constant Systems Ltd) at 20 kpsi. The resultant lysates were centrifuged (23000 rpm, 15 min, 4 ℃) and the supernatant was purified at 4℃ using a Ni(II)-affinity chromatography (Histrap HP column, cytiva) and an ÄKTA Pure chromatography system (GE Healthcare). An elution gradient of increasing imidazole concentrations (from: 20 mM imidazole, 500 mM NaCl, 20 mM Tris, pH 7.8; to: 500 mM imidazole, 500 mM NaCl, 20 mM Tris, pH 7.8) was used. Fractions containing the desired ACCO protein (identified by SDS-PAGE analysis) were combined and concentrated to 5 mL using an Amicon Ultra centrifugal filter (30 kDa MWCO; 4255 g, 4℃). The concentrated protein solution was subjected to buffer exchange to lower the imidazole concentration by addition of 5 mL of Tobacco etch virus (TEV) cleavage buffer solution (20 mM Tris, 150 mM NaCl, pH 7.8, $10\%_{v/v}$ glycerol), mixing and concentrating down to 5 mL in the same centrifugal filter. This process was repeated

twice. The N-terminal His_6-tag was cleaved overnight using recombinant TEV protease (sequence given in the in Supporting Information Supplementary Table S6) at 4℃. The resultant mixture was purified using Ni(II)-affinity chromatography using Histrap binding buffer (20 mM imidazole, 500 mM NaCl, 20 mM Tris, pH 7.8). The flow-through was collected, concentrated, and purified using size-exclusion chromatography (300 mL HiLoad 26/60 Superdex 75 pg column or 120 mL HiLoad 16/60 Superdex 75 pg column; ÄKTA Pure chromatography system; buffer: 25 mM Tris, pH 7.5; flow rate: 1.0 mL/min). Eluted fractions containing >95% pure target protein (estimated from Invitrogen SDS-PAGE gel and MS analyses) were combined, concentrated and aliquoted. The aliquots were flash frozen in liquid nitrogen and stored at -80℃. The protein concentration was determined with a NanoDrop One microvolume UV-Vis spectrometer (Thermo Scientific), assuming all protein molecules are active. Fresh aliquots of protein were used for biochemical experiments.

6.3 ^{1}H NMR assays

^{1}H NMR assays were performed using either a Bruker AVIII 600 MHz spectrometer or a Bruker AVIII 700 MHz NMR spectrometer with a 5-mm inverse triple-resonance-inverse cryoprobe.

6.3.1 Assay Preparation

Aqueous stocks of $NaHCO_3$ (500 mM), L-ascorbate (500 mM), ammonium iron (II) sulphate (100 mM) and substrate (ACC or ACC-derivatives, 100 mM) were freshly prepared on the day assays were performed. The stock solutions of $NaHCO_3$, L-ascorbate, ACC/ACC derivatives, along with a stock (80 mM) of (trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP-d_4) solution, which was used as an internal standard (0 ppm), were diluted in buffer (50 mM KH_2PO_4/K_2HPO_4, 10%$_{v/v}$ D_2O, pH 8.0) to give $NaHCO_3$ (40 mM), L-ascorbate (32 mM), ACC/ACC-

derivatives (4 mM) and TMSP-d_4 (8 mM) final assay concentrations. The ammonium iron (II) sulphate stock solution was diluted in water to 0.5 mM. All stock solutions were incubated on ice.

6.3.2 Sample Preparation

^{1}H NMR assays were performed using a similar setup as reported for PsEFE assay[48]. In brief, a final reaction volume of 180μL was used with the following concentrations: 400 µM substrate (ACC or ACC-derivatives), 10 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe(II), and 800 μM TMSP-d_4. The indicated stock solutions were added into a 1.5 mL Eppendorf tube in the following order: buffer solution (50 mM KH_2PO_4/K_2HPO_4, 10%$_{v/v}D_2O$, pH 8.0; variable volume), then $NaHCO_3$ (40 mM; 18 μL); then L-ascorbate (32 mM; 18 μL); then ammonium iron(II) sulphate (0.5 mM;18 μL); then ACC/ACC-derivatives (4 mM; 18μL); then TMSP-d_4 (8 mM; 18μ L); then ACCO/ACCO variants (variable volume). The assay mixture was mixed by pipetting, and transferred to a 3 mm Bruker NMR tube, then hand centrifuged (Eisco Lab Crank Hand Centrifuge).

6.3.3 NMR Instrument Operation

The NMR tubes were transferred as rapidly as possible to the spectrometer and inserted into the probe (approximately 1 min). The probe was locked to 10%$_{v/v}$ D_2O, followed by auto-tuning (command “atma” in Topspin 4.3.1, Bruker), auto-shimming (command “topshim”), pulse calibration (command “puc”), getprosol (command “getprosol”), receiver gain adjustment (command “rga”), and consecutive acquisitions were started. The variable delay was set to zero and the number of scans per acquisition was set to 64 to ensure a sufficient signal-to-noise ratio. The relaxation

delay was minimised to reduce the time-gap between acquisitions. Water suppression was achieved using perfect echo-modified WATERGATE experiments[63]. Digital resolution of the free induction delay (FID) was adjusted to 0.3 Hz for accuracy in integration.

6.3.4 Data Processing

Peak integrals were normalised via comparisons with that of TMSP-d_4. Spectra were subjected to automatic phase and baseline correction. Integral precision was attenuated by zero filling using MestReNova x86_64 (Mestrelab Research) and integrals calibrated and scaled to the internal standard. The enzyme turnover was measured in terms of substrate depletion since the product is gaseous, assuming all depleted substrate was converted to the product. The peak corresponding to one of the protons from the substrate (either ACC or ACCO derivatives) with the lowest level of protein noise interference was integrated. The substrate concentration for each timepoint in a single time course was calculated by comparing the integral number to that for the internal standard. The substrate depletion curve was then plotted after calculating the substrate concentration at each timepoint in a single time course.

6.4 Mutagenesis, Plasmid Preparation and Sequencing

6.4.1 PCR Protocol

All AtACCO mutants were obtained from wt AtACCO using the Quikchange site-directed mutagenesis[64]. The primer pair contained an overlapping sequence, causing the formation of a homologous sequence at both terminals of the PCR products. The PCR product was directly transformed into 5-alpha *E. coli* competent cells (New England Biolabs) which ligase the PCR product into a plasmid via homologous

recombination.

The primer pair and template used for the cloning of each mutant are shown in Section 7. All PCR reactions were performed with Q5 High-Fidelity DNA Polymerase (New England Biolab) with the following protocol. For a 20 µL reaction mixture:

a) PCR Setup

Component	20 µL reaction	Final Concentration
5X Q5 Reaction Buffer	4 µL	1X
10 mM dNTPs	0.4 µL	200 µM
10 µM Forward Primer	1 µL	0.5 µM
10 µM Reverse Primer	1 µL	0.5 µM
2 ng/µL Template Plasmid	1 µL	0.1 ng/µL
Q5 DNA Polymerase	0.2 µL	0.02 U/ µL
Nuclease-Free Water	12.4 µL	-

b) Thermocycling conditions:

Step	Temperature	Duration
Initial Denaturation	98℃	30s
35 Cycles	98℃ 60℃ 72℃	10s 30s 3 min 30 s
Final Extension	72℃	5 min
Hold	12℃	2 min

The PCR products were subjected to template plasmid clean up using DpnI restriction enzyme (New England Biolab). 0.5 µL of DpnI enzyme was added to 10 µL PCR

product. The reaction mixture was incubated at 37℃ for 15 min, then at 80 ℃for 15 min to deactivate the enzyme. The processed PCR product was directly transformed into 5-alpha competent *E. coli* cells (New England Biolabs).

6.4.2 Plasmid Preparation and Sequencing

Cells were grown in 5 mL LB Broth (Thermo Fisher Scientific) overnight and harvested using centrifugation (4000 rpm, 10 min). The resultant cell pellets were purified using Plasmid Miniprep Kit (Zymopure, Cat No. D4210). The resulted plasmids were stored at -20℃. A 5 µL sample of 100 ng/µL was sent for Sanger Sequencing (Source Bioscience) using T7 promoter forward primer and T7 promoter forward primer.

6.5 Crystallography

6.5.1 Broad Screen for ACCO Crystallography

Crystallisation trials were conducted with the *His_6*-tag-cleaved ACCO mutants in the presence of ACC & $CoCl_2$. Mixtures were incubated on ice containing AtACCO (18 mg/mL), $CoCl_2$•$4H_2O$ (3 mM), and ACC (20 mM). The solutions were then centrifuged (4000 g, 5 min, 4 °C) to remove any precipitate.

High-throughput crystallisation trials were performed in 96-well, 3-subwell Swisssci plates (Molecular Dimensions & Hampton Research, Table 1). The crystallisation supernatant was dispensed as 150 nL sitting drops into the sub-wells (2:1, 1:1, and 1:2 sample to well ratio) using a Mosquito crystal robot (Sptlabtech). Crystals were grown using the vapour diffusion method at room temperature. Putative protein crystals were

harvested from the PegION screen and directly cryo-cooled in liquid nitrogen for data collection. Crystals were mounted in a cryo-stream 100K and data were collected at the Diamond Synchrotron Light Source (Didcot UK) on beamlines I04-1 and I02.

1. JCSG-*plus*™ HT-96 Eco Screen	13. Morpheus HT-96
2. PACT *premier* HT-96 Eco Screen	14. Wizard Classic 1 and 2 Block
3. SG1 Screen HT-96	15. PEG/Ion HT™
4. Stura FootPrint Combination HT-96/FX-96	16.The LMB Crystallization Screen
5. NeXtal Anion Suite	17. Crystal Screen HT™
6. The LFS ECO HT-96 MD1-122-ECO	18. Index™
7. NeXtal Cations Suite	19. ProPlex™ HT-96/FX-96
8. The BCS Eco Screen HT-96	20. MembFac HT™
9. PEGRx HT™	21. Natrix HT™
10. SaltRx HT™	22. Morpheus II HT-96
11. Morpheus Fusion HT-96	23. Morpheus III HT-96 MD1-116
12. The MPD Suite	24. MIDAS*plus*™ HT-96 MD1-107

Table 1. List of screening kits used for crystallisation broad screening

6.5.2 Broad Screen for PsEFE Crystallography

Crystallisation trials were conducted with the tag-cleaved 1449B PsEFE and different co-substrates, including 2OG, 3-methyl-2OG, 4-methyl-2OG, and 3-ethyl-2OG. Mixtures were incubated on ice containing 1449B PsEFE (11 mg/mL), $MnCl_2$•$4H_2O$ (3 mM), and the 2OG/2OG derivative (10 mM). The solutions were then centrifuged (4000 x g, 5 min, 4 °C) to clear any precipitate.

High-throughput trials were performed in 96-well, 3-subwell Swisssci (Molecular Dimensions & Hampton Research, Table 1). The crystallisation supernatant was

dispensed as 150 nL sitting drops into the sub-wells (2:1, 1:1, and 1:2 sample to well ratio) using a Mosquito crystal robot (Sptlabtech). Crystals were grown using the vapour diffusion method at 17 °C. Putative protein crystals were harvested from the PEGION screen and directly cryo-cooled in liquid nitrogen for data collection. Crystals were mounted in a cryo-stream 100K and data were collected at the Diamond Synchrotron Light Source (Didcot UK) on beamlines I04-1 and I02.

6.5.3 Optimisation for PsEFE Crystallisation

Crystallization conditions were optimised for 1449B PsEFE based on the conditions of PEG/Ion H4 (0.03 M Citric acid, 0.07 M BIS-TRIS propane/ pH 7.6, 20% w/v Polyethylene glycol (PEG) 3350) with varied polyethylene glycol concentrations.

0.03M Citric Acid/0.07M BIS-TRIS propane 18% w/v PEG 3350	0.03M Citric Acid/0.07M BIS-TRIS propane 19% w/v PEG 3350	0.03M Citric Acid/0.07M BIS-TRIS propane 20% w/v PEG 3350	0.03M Citric Acid/0.07M BIS-TRIS propane 21% w/v PEG 3350	0.03M Citric Acid/0.07M BIS-TRIS propane 22% w/v PEG 3350	0.03M Citric Acid/0.07M BIS-TRIS propane 23% w/v PEG 3350
30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 135 µL 40% PEG3350 135 µL H_2O	30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 142.5 µL 40% PEG3350 127.5 µL H_2O	30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 150 µL 40% PEG3350 120 µL H_2O	30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 157.5 µL 40% PEG3350 112.5 µL H_2O	30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 165 µL 40% PEG3350 105 µL H_2O	30 µL 0.3M Citric Acid/ 0.7M BIS-TRIS propane 172.5 µL 40% PEG3350 97.5 µL H_2O

Table 2. Optimised crystallography conditions for 1449B PsEFE crystallography.

The final volume of crystallisation reagents prepared into each well was 300 uL (Table 2). The crystallisation supernatant was dispensed as sitting drops into the wells (crystallisation supernatant (µL): reagent (µL) = 2:1, 1:1, and 1:2 sample to well ratio)

manually. Crystals were grown using the vapour diffusion method at 17 °C. Putative protein crystals were harvested and directly cryo-cooled in liquid nitrogen for data collection. Crystals were mounted in a cryo-stream 100K and data were collected at the Diamond Synchrotron Light Source (Didcot UK) on beamlines I04-1 and I02.

References

1 K. M. Sundaram, M. M. Shreehan and E. F. Olszewski, in *Kirk-Othmer Encyclopedia of Chemical Technology*, ed. Kirk-Othmer, Wiley, 1st edn., 2010, pp. 1–39.

2 Z. Li, M. Åhman, L. J. Nilsson and F. Bauer, *Renewable and Sustainable Energy Reviews*, 2024, **199**, 114540.

3 P. Del Campo, M. T. Navarro, S. K. Shaikh, M. D. Khokhar, F. Aljumah, C. Martínez and A. Corma, *ACS Catal.*, 2020, **10**, 11878–11891.

4 W. E. Luttrell and N. P. Giles, *J. Chem. Health Saf.*, 2016, **23**, 34–36.

5 R. N. Webb, T. D. Shaffer and A. H. Tsou, in *Kirk-Othmer Encyclopedia of Chemical Technology*, ed. Kirk-Othmer, Wiley, 1st edn., 2003.

6 Y. Lin and S. Tanaka, *Appl Microbiol Biotechnol*, 2006, **69**, 627–642.

7 M. Zhang and Y. Yu, *Ind. Eng. Chem. Res.*, 2013, **52**, 9505–9514.

8 M. Dubois, L. Van Den Broeck and D. Inzé, *Trends in Plant Science*, 2018, **23**, 311–323.

9 K. Shin, S. Lee, W.-Y. Song, R.-A. Lee, I. Lee, K. Ha, J.-C. Koo, S.-K. Park, H.-G. Nam, Y. Lee and M.-S. Soh, *Plant and Cell Physiology*, 2015, **56**, 572–582.

10 B. Van De Poel, D. Smet and D. Van Der Straeten, *Plant Physiol.*, 2015, **169**, 61–72.

11 W. Mou, Y.-T. Kao, E. Michard, A. A. Simon, D. Li, M. M. Wudick, M. A. Lizzio, J. A. Feijó and C. Chang, *Nat Commun*, 2020, **11**, 4082.

12 F. Xu, Y. Peng, Z.-Q. He and L.-L. Yu, *BMC Plant Biol*, 2023, **23**, 163.

13 R. F. Lacey and B. M. Binder, *Journal of Inorganic Biochemistry*, 2014, **133**, 58–62.

14 C. Merchante, J. Brumos, J. Yun, Q. Hu, K. R. Spencer, P. Enríquez, B. M. Binder, S. Heber, A. N. Stepanova and J. M. Alonso, *Cell*, 2015, **163**, 684–697.

15 H. Guo and J. R. Ecker, *Cell*, 2003, **115**, 667–677.

16 T. Potuschak, E. Lechner, Y. Parmentier, S. Yanagisawa, S. Grava, C. Koncz and P. Genschik, *Cell*, 2003, **115**, 679–689.

17 T. Nakano, K. Suzuki, N. Ohtsuki, Y. Tsujimoto, T. Fujimura and H. Shinshi, *J Plant Res*, 2006, **119**, 407–413.

18 S.-D. Yoo and J. Sheen, *Plant Signaling & Behavior*, 2008, **3**, 848–849.

19 X. Meng, J. Xu, Y. He, K.-Y. Yang, B. Mordorski, Y. Liu and S. Zhang, *The Plant Cell*, 2013, **25**, 1126–1142.

20 D. M. Penrose, B. A. Moffatt and B. R. Glick, *Can. J. Microbiol.*, 2001, **47**, 77–80.
21 B. R. Glick, Z. Cheng, J. Czarny and J. Duan, *Eur J Plant Pathol*, 2007, **119**, 329–339.
22 Y. Li, M. Qi, Q. Zhang, Z. Xu, Y. Zhang, Y. Gao, Y. Qi, L. Qiu and M. Wang, *JoF*, 2022, **9**, 55.
23 M. Steinke and G. O. Kirst, *Journal of Experimental Marine Biology and Ecology*, 1996, **201**, 73–85.
24 R. F. Lacey and B. M. Binder, *Plant Physiol.*, 2016, **171**, 2798–2809.
25 F. I. Rodríguez, J. J. Esch, A. E. Hall, B. M. Binder, G. E. Schaller and A. B. Bleecker, *Science*, 1999, **283**, 996–998.
26 N. Papon and B. M. Binder, *Trends in Microbiology*, 2019, **27**, 193–196.
27 W. Wang, J. J. Esch, S.-H. Shiu, H. Agula, B. M. Binder, C. Chang, S. E. Patterson and A. B. Bleecker, *The Plant Cell*, 2007, **18**, 3429–3442.
28 A. Hérivaux, T. Dugé De Bernonville, C. Roux, M. Clastre, V. Courdavault, A. Gastebois, J.-P. Bouchara, T. Y. James, J.-P. Latgé, F. Martin and N. Papon, *mBio*, 2017, **8**, e01739-16.
29 C. Eckert, W. Xu, W. Xiong, S. Lynch, J. Ungerer, L. Tao, R. Gill, P.-C. Maness and J. Yu, *Biotechnol Biofuels*, 2014, **7**, 33.
30 S. Mansouri and A. W. Bunch, *Microbiology*, 1989, **135**, 2819–2827.
31 B. N. Shipston and A. W. Bunch, *Microbiology*, 1989, **135**, 1489–1497.
32 J. A. North, A. R. Miller, J. A. Wildenthal, S. J. Young and F. R. Tabita, *Proc. Natl. Acad. Sci. U.S.A.*, DOI:10.1073/pnas.1711625114.
33 J. A. North, A. B. Narrowe, W. Xiong, K. M. Byerly, G. Zhao, S. J. Young, S. Murali, J. A. Wildenthal, W. R. Cannon, K. C. Wrighton, R. L. Hettich and F. R. Tabita, *Science*, 2020, **369**, 1094–1098.
34 Md. S. Islam, T. M. Leissing, R. Chowdhury, R. J. Hopkinson and C. J. Schofield, *Annu. Rev. Biochem.*, 2018, **87**, 585–620.
35 L. J. Walport, R. J. Hopkinson and C. J. Schofield, *Current Opinion in Chemical Biology*, 2012, **16**, 525–534.
36 W. S. Aik, R. Chowdhury, I. J. Clifton, R. J. Hopkinson, T. Leissing, M. A. McDonough, R. Nowak, C. J. Schofield and L. J. Walport, in *2-Oxoglutarate-Dependent Oxygenases*, eds. C. Schofield and R. Hausinger, The Royal Society of Chemistry, 2015, pp. 59–94.
37 I. J. Clifton, M. A. McDonough, D. Ehrismann, N. J. Kershaw, N. Granatino and C. J. Schofield, *Journal of Inorganic Biochemistry*, 2006, **100**, 644–669.
38 S. Markolovic, T. M. Leissing, R. Chowdhury, S. E. Wilkins, X. Lu and C. J. Schofield, *Current Opinion in Structural Biology*, 2016, **41**, 62–72.
39 M. A. McDonough, C. Loenarz, R. Chowdhury, I. J. Clifton and C. J. Schofield, *Current Opinion in Structural Biology*, 2010, **20**, 659–672.
40 S. Khan, Z. Sehar, Nidhi, M. Albaqami and N. A. Khan, *Environmental and Experimental Botany*, 2023, **212**, 105379.
41 L. Brisson, N. El Bakkali-Taheri, M. Giorgi, A. Fadel, J. Kaizer, M. Réglier, T. Tron, E. H. Ajandouz and A. J. Simaan, *J Biol Inorg Chem*, 2012, **17**, 939–949.

42 Z. Zhang, J.-S. Ren, I. J. Clifton and C. J. Schofield, *Chemistry & Biology*, 2004, **11**, 1383–1394.
43 S. Zhao, L. Wu, Y. Xu and Y. Nie, *Nat. Prod. Rep.*, 2025, **42**, 67–92.
44 S. Markolovic, Q. Zhuang, S. E. Wilkins, C. D. Eaton, M. I. Abboud, M. J. Katz, H. E. McNeil, R. K. Leśniak, C. Hall, W. B. Struwe, R. Konietzny, S. Davis, M. Yang, W. Ge, J. L. P. Benesch, B. M. Kessler, P. J. Ratcliffe, M. E. Cockman, R. Fischer, P. Wappner, R. Chowdhury, M. L. Coleman and C. J. Schofield, *Nat Chem Biol*, 2018, **14**, 688–695.
45 D. R. Dilley, Z. Wang, D. K. Kadirjan-Kalbach, F. Ververidis, R. Beaudry and K. Padmanabhan, *AoB PLANTS*, DOI:10.1093/aobpla/plt031.
46 R. P. Hausinger, S. B. J. S. Rifayee, M. G. Thomas, S. Chatterjee, J. Hu and C. Z. Christov, *RSC Chem. Biol.*, 2023, **4**, 635–646.
47 H. Weingart, H. Ullrich, K. Geider and B. Völksch, *Phytopathology®*, 2001, **91**, 511–518.
48 S. Dhingra, Z. Zhang, C. T. Lohans, L. Brewitz and C. J. Schofield, *Journal of Biological Chemistry*, 2024, **300**, 107546.
49 S. Martinez, M. Fellner, C. Q. Herr, A. Ritchie, J. Hu and R. P. Hausinger, *J. Am. Chem. Soc.*, 2017, **139**, 11980–11988.
50 N. M. W. Brunhuber, J. L. Mort, R. E. Christoffersen and N. O. Reich, *Biochemistry*, 2000, **39**, 10730–10738.
51 H. P. J. De Wild, *Journal of Experimental Botany*, 2003, **54**, 1537–1544.
52 T. F. Cummings, *Occupational Medicine*, 2004, **54**, 82–85.
53 P. John, *Physiologia Plantarum*, 1997, **100**, 583–592.
54 R. E. Mitchell, M. C. Pirrung and G. M. McGeehan, *Phytochemistry*, 1987, **26**, 2695–2697.
55 N. E. Hoffman, S. F. Yang, A. Ichihara and S. Sakamura, *Plant Physiol.*, 1982, **70**, 195–199.
56 M. C. Pirrung, J. Cao and J. Chen, *Chemistry & Biology*, 1998, **5**, 49–57.
57 R. E. Mitchell, S. A. Young and C. L. Bender, *Phytochemistry*, 1994, **35**, 343–348.
58 M. Melotto, W. Underwood and S. Y. He, *Annu. Rev. Phytopathol.*, 2008, **46**, 101–122.
59 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, *Nature*, 2024, **630**, 493–500.
60 J. Baltrusaitis, *ACS Sustainable Chem. Eng.*, 2017, **5**, 9527–9527.
61 I. Rafiqul, C. Weber, B. Lehmann and A. Voss, *Energy*, 2005, **30**, 2487–2504.
62 H. H. A. Dollwet and R. E. Seeman, *Plant Physiol.*, 1975, **56**, 552–554.
63 R. W. Adams, C. M. Holroyd, J. A. Aguilar, M. Nilsson and G. A. Morris, *Chem.*

Commun., 2013, **49**, 358–360.
64 H. Liu and J. H. Naismith, *BMC Biotechnol*, 2008, **8**, 91.

Supplementary Information

Figure S1. Deconvoluted electrospray ionisation (ESI) mass spectrum of L196A ACCO. calculated mass: 35739.87 Da; Observed mass: 35740.26 Da.

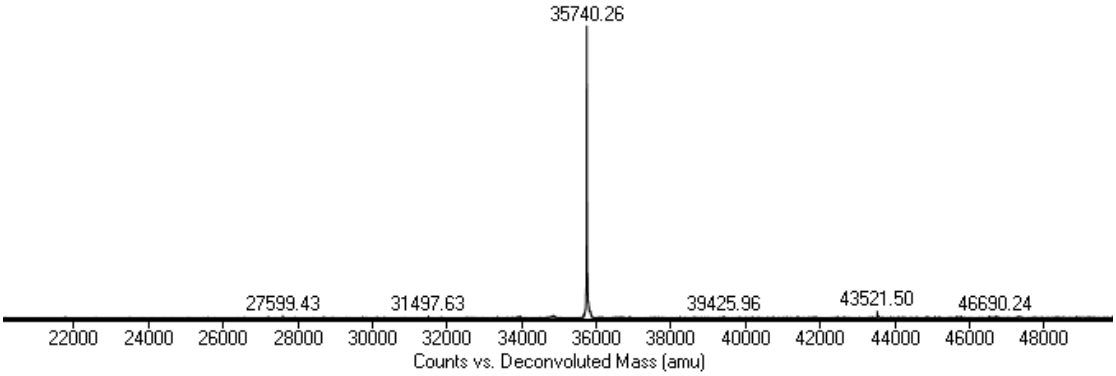


Figure S2. Deconvoluted electrospray ionisation (ESI) mass spectrum of K159A ACCO. Calculated mass: 35724.85 Da; Observed mass: 35724.93 Da.

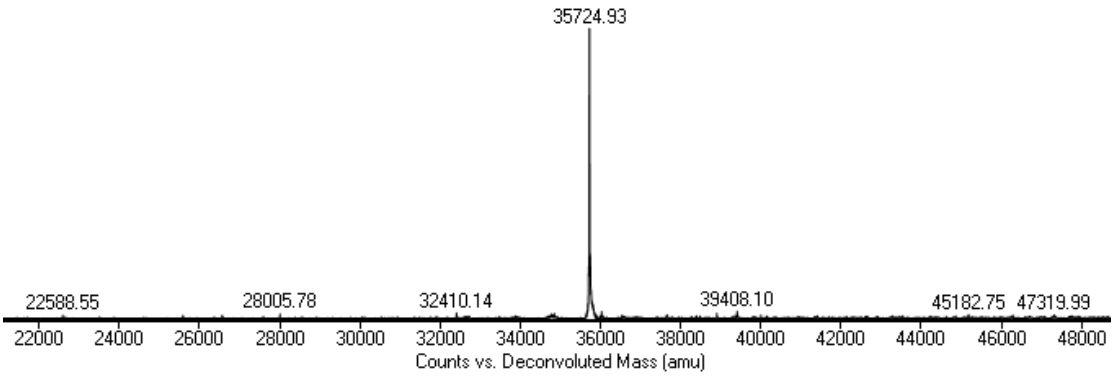


Figure S3. Deconvoluted electrospray ionisation (ESI) mass spectrum of K159E ACCO. Calculated mass: 35782.89 Da; Observed mass: 35782.51 Da.

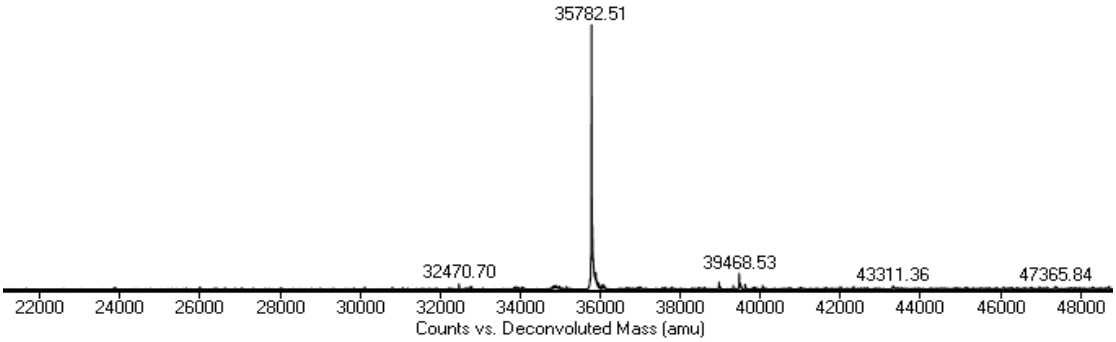


Figure S4. Deconvoluted electrospray ionisation (ESI) mass spectrum of K159R ACCO. Calculated mass: 35724.85 Da; Observed mass: 35724.93 Da.

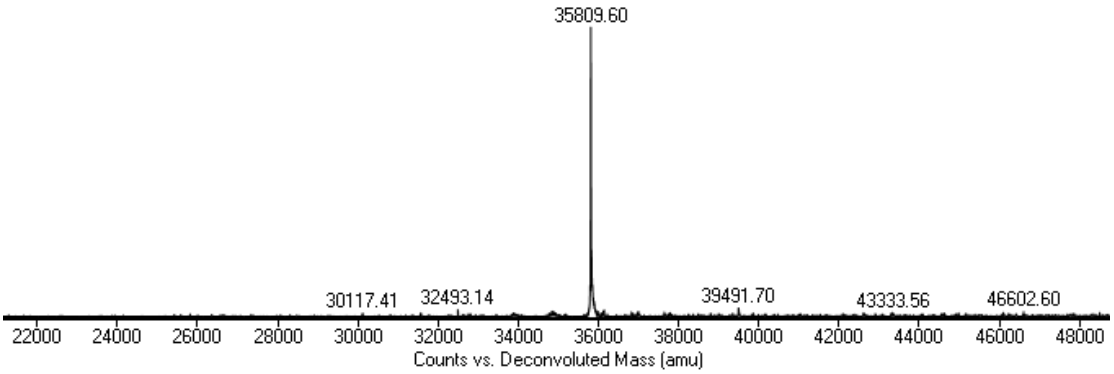


Figure S5. Deconvoluted electrospray ionisation (ESI) mass spectrum of K290E ACCO. Calculated mass: 35782.89 Da; Observed mass: 35782.61 Da.

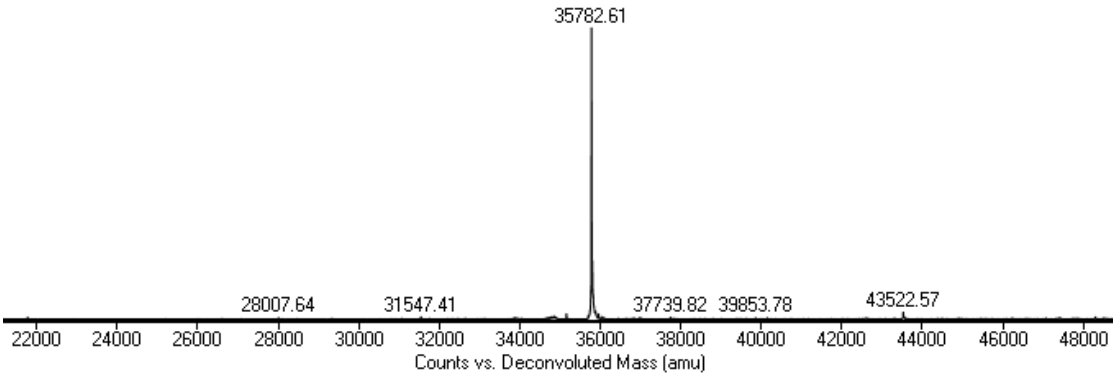


Figure S6. Deconvoluted electrospray ionisation (ESI) mass spectrum of K290R ACCO. Calculated mass: 35809.96 Da; Observed mass: 35809.62 Da.

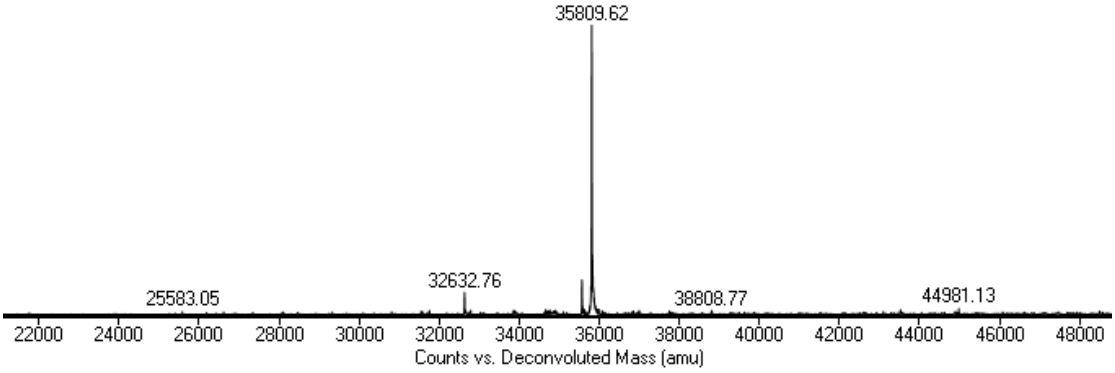


Figure S7. Deconvoluted electrospray ionisation (ESI) mass spectrum of ACCO Y91F. Calculated mass: 35765.95 Da; Observed mass: 35765.63 Da.

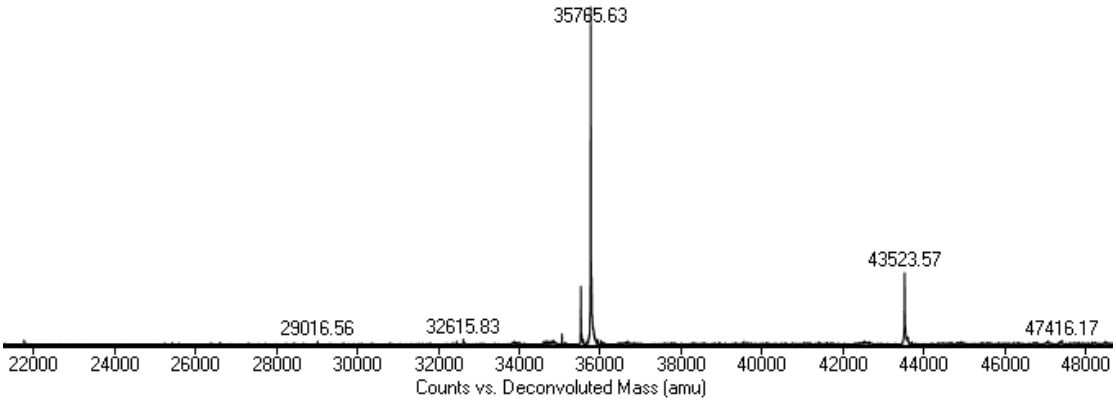


Figure S8. Deconvoluted electrospray ionisation (ESI) mass spectrum of Y91H ACCO. Calculated mass: 35755.91 Da; Observed mass: 35755.50 Da.

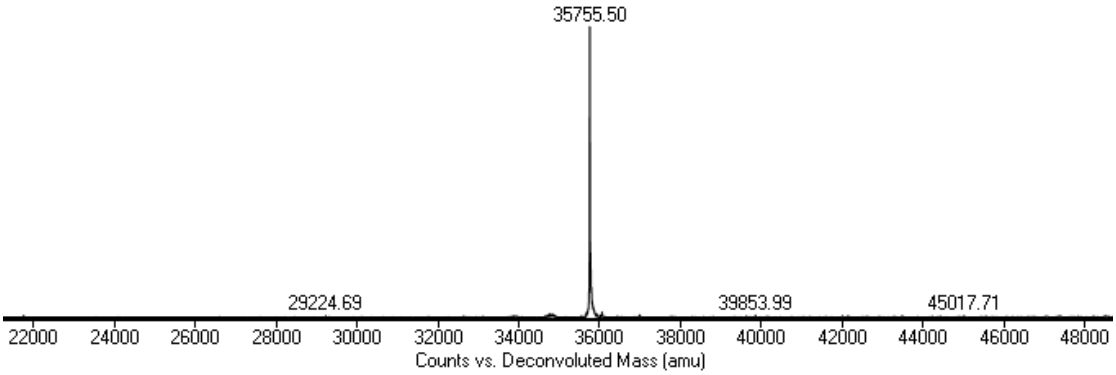


Figure S9. Deconvoluted electrospray ionisation (ESI) mass spectrum of Y163F ACCO. Calculated mass: 35765.95 Da; Observed mass: 35765.65 Da.

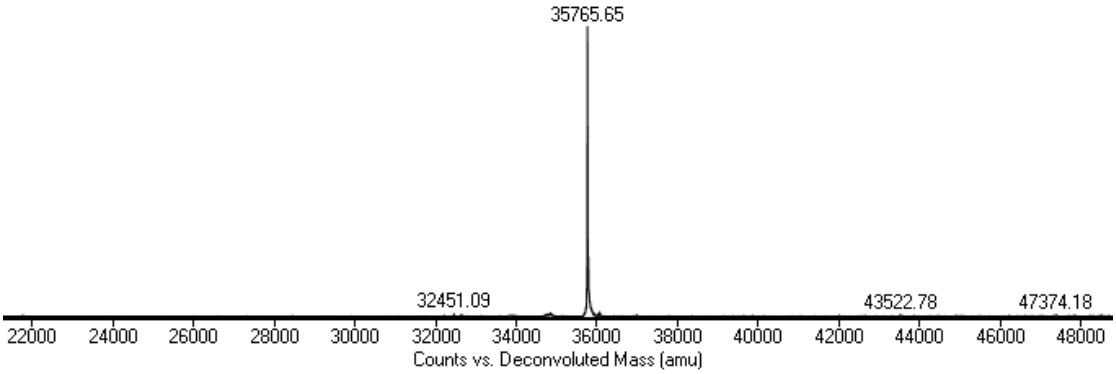


Figure S10. Deconvoluted electrospray ionisation (ESI) mass spectrum of Y163H ACCO. Calculated mass: 35765.91 Da; Observed mass: 35755.62 Da.

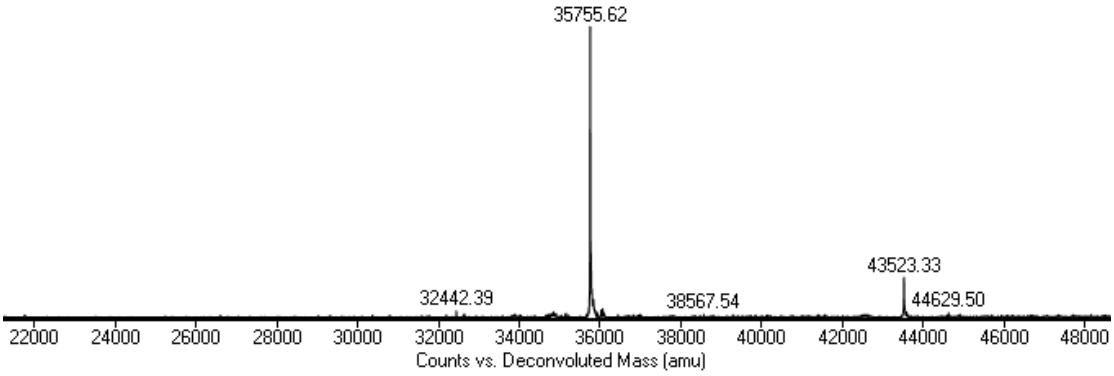


Figure S11. Deconvoluted electrospray ionisation (ESI) mass spectrum of R245Q ACCO. Calculated mass: 35753.89 Da; Observed mass: 35753.63 Da.

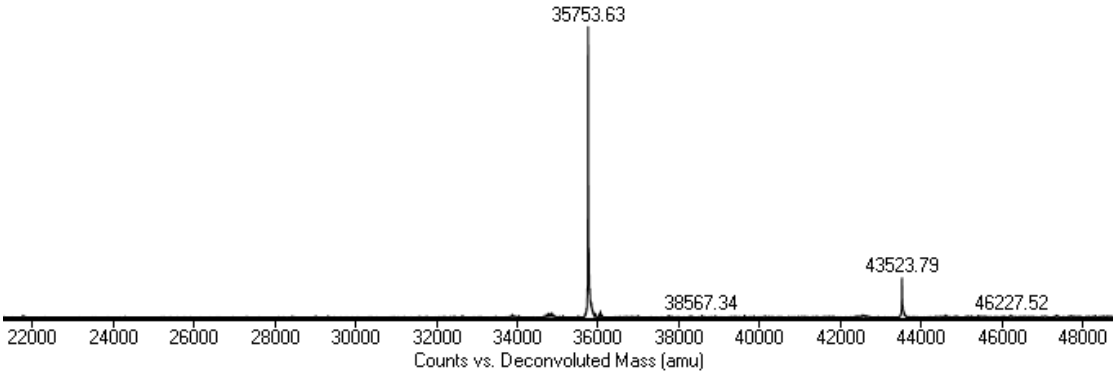


Figure S12. Deconvoluted electrospray ionisation (ESI) mass spectrum of S247A ACCO. Calculated mass: 35765.95 Da; Observed mass: 35765.66 Da.

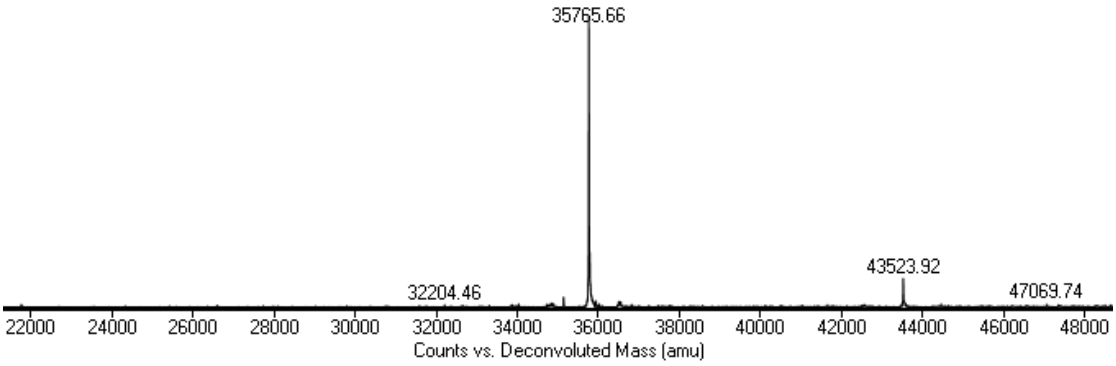


Figure S13. Deconvoluted electrospray ionisation (ESI) mass spectrum of K290A ACCO. Calculated mass: 35724.85 Da; Observed mass: 35724.71 Da.

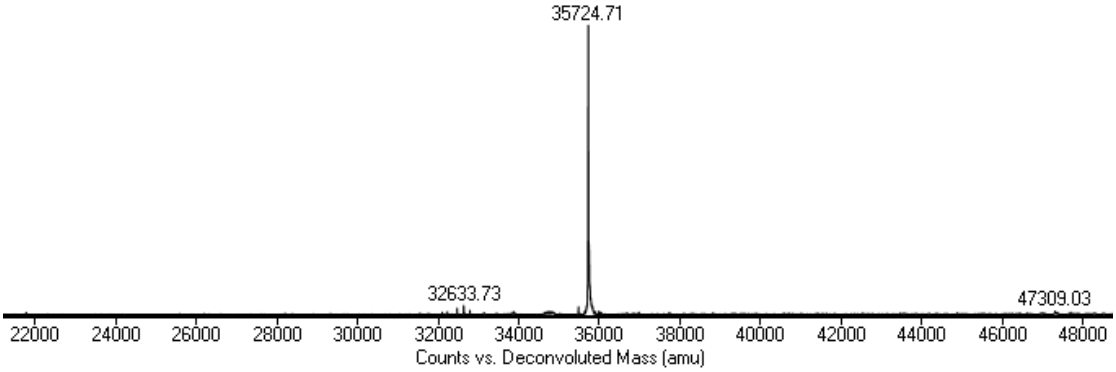


Figure S14. Deconvoluted electrospray ionisation (ESI) mass spectrum of N217Q ACCO. Calculated mass: 35795.98 Da; Observed mass: 35794.72 Da.

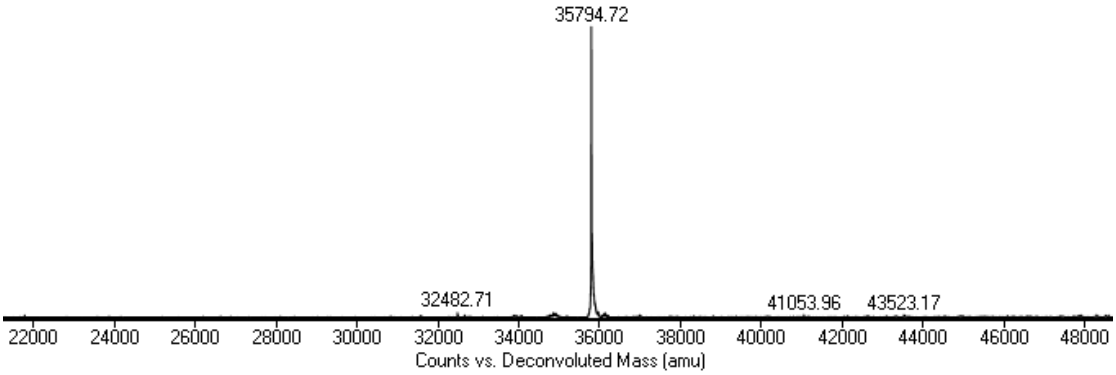


Figure S15. Deconvoluted electrospray ionisation (ESI) mass spectrum of Y286F ACCO. Calculated mass: 35795.98 Da; Observed mass: 35794.72 Da.

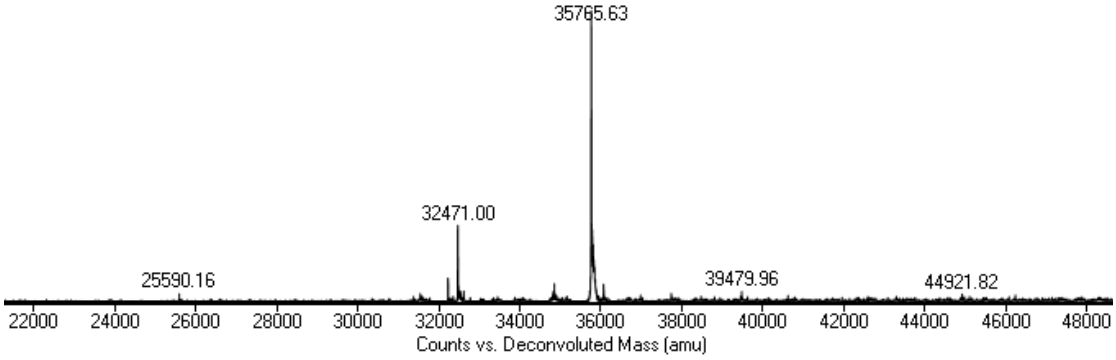


Figure S16. Deconvoluted electrospray ionisation (ESI) mass spectrum of N253A ACCO. Calculated mass: 35738.92 Da; Observed mass: 35738.70 Da.

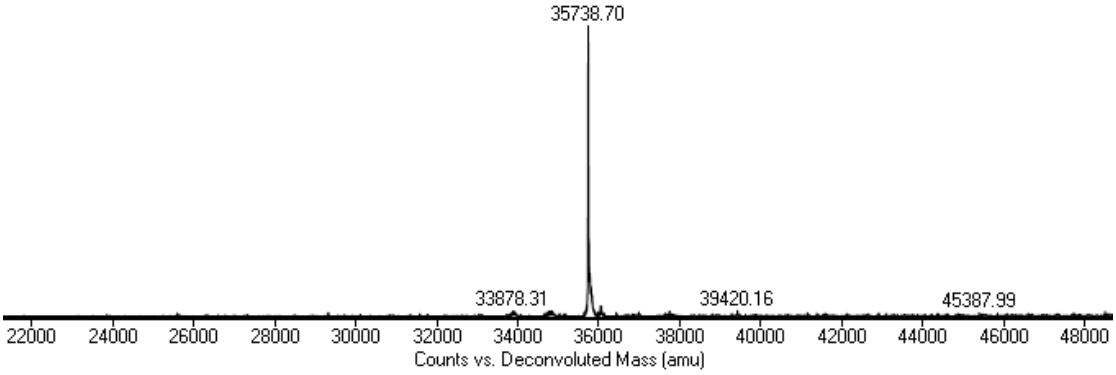


Figure S17. Deconvoluted electrospray ionisation (ESI) mass spectrum of L187V ACCO. Calculated mass: 35767.92 Da; Observed mass: 35767.52 Da.

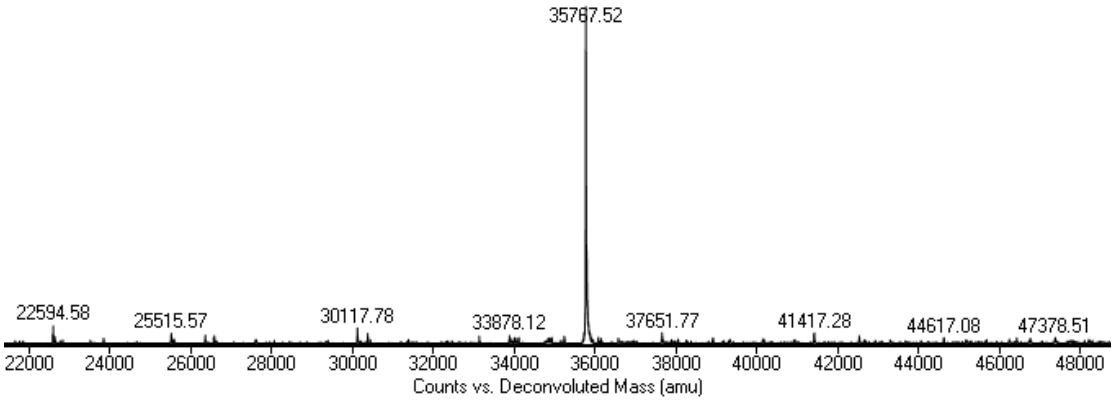


Figure S18. Deconvoluted electrospray ionisation (ESI) mass spectrum of L187A ACCO. Calculated mass: 37401.78 Da; Observed mass: 37400.48 Da.

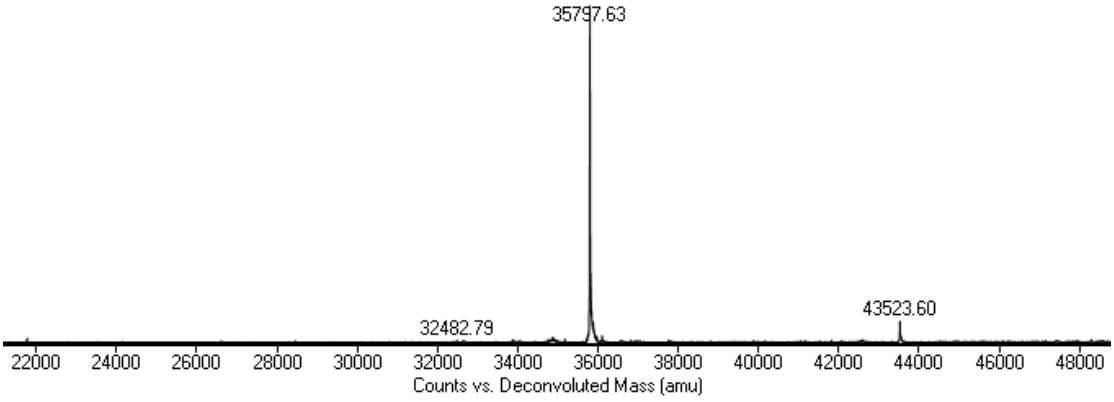


Figure S19. Deconvoluted electrospray ionisation (ESI) mass spectrum of S247C ACCO. Calculated mass: 35798.01 Da; Observed mass: 35797.63 Da.

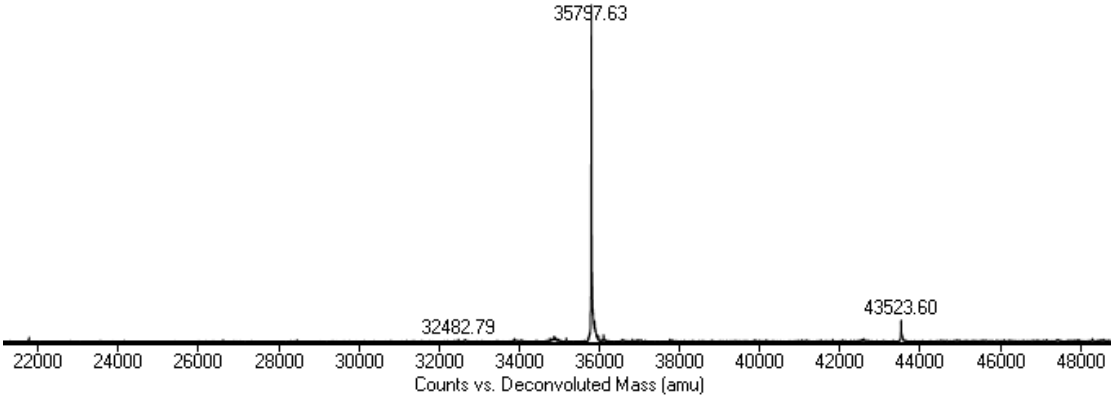


Figure S20. Deconvoluted electrospray ionisation (ESI) mass spectrum of R245K ACCO. Calculated mass: 35753.94 Da; Observed mass: 35753.78 Da.

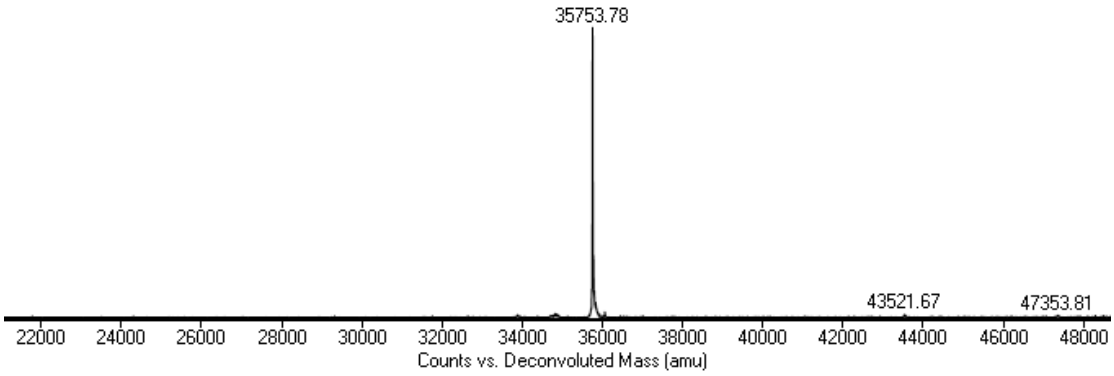


Figure S21. Deconvoluted electrospray ionisation (ESI) mass spectrum of N253Q ACCO. Calculated mass: 35795.98 Da; Observed mass: 35794.72 Da.

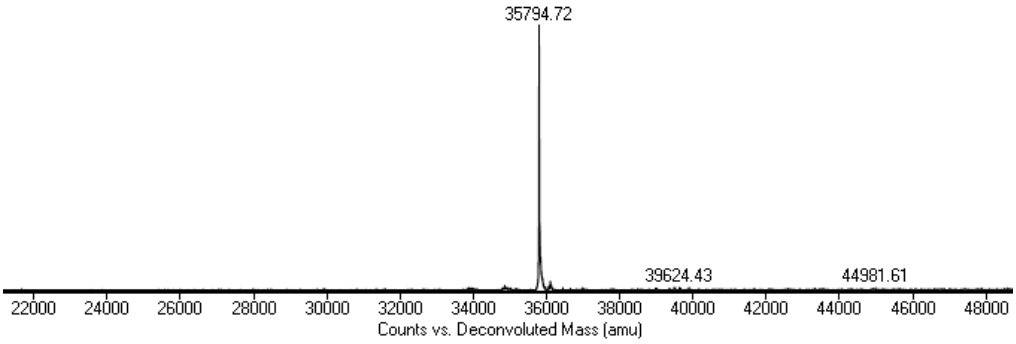


Figure S22. Deconvoluted electrospray ionisation (ESI) mass spectrum of V185G ACCO. Calculated mass: 37445.83 Da; Observed mass: 37444.69 Da.

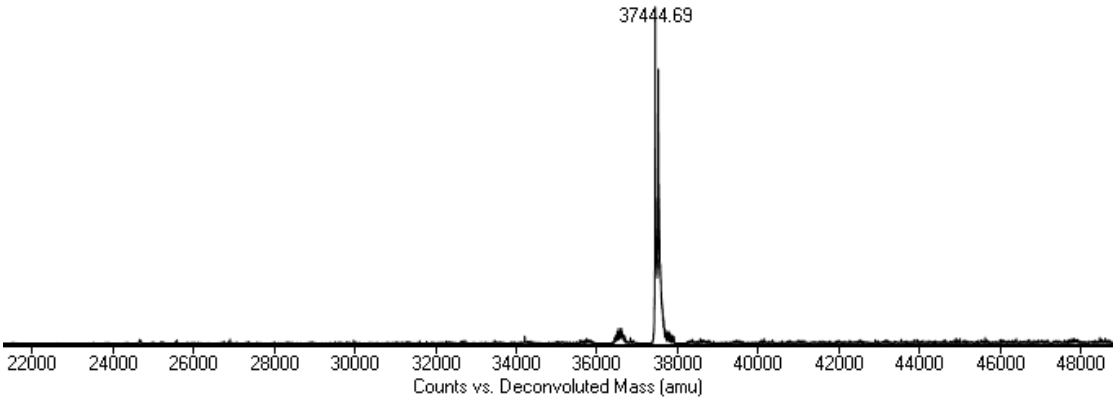


Figure S23. Deconvoluted electrospray ionisation (ESI) mass spectrum of N217G ACCO. Calculated mass: 37430.86 Da; Observed mass: 37428.81 Da.

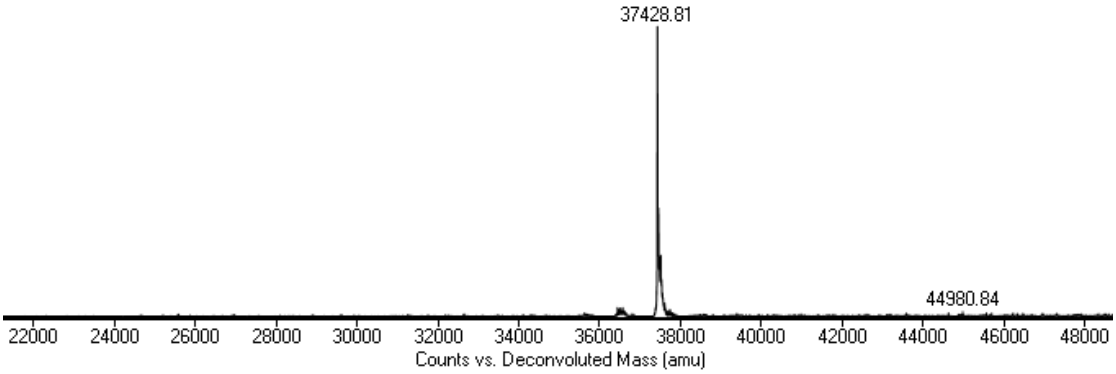


Figure S24. Deconvoluted electrospray ionisation (ESI) mass spectrum of the L187A/S247A AtACCO. Calculated Mass: 37429.83; Observed mass: 37430.67 Da.

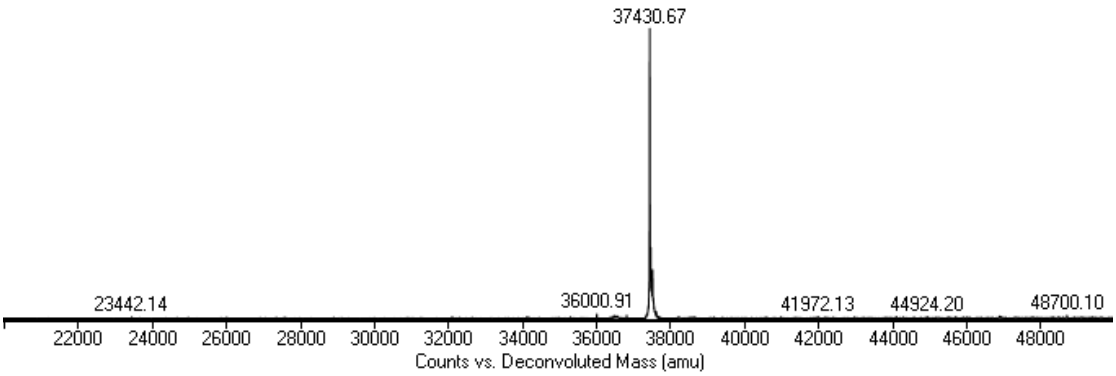


Figure S25. Deconvoluted electrospray ionisation (ESI) mass spectrum of L187G/ S247G ACCO. Calculated mass: 37401.78 Da; Observed mass: 37400.48 Da.

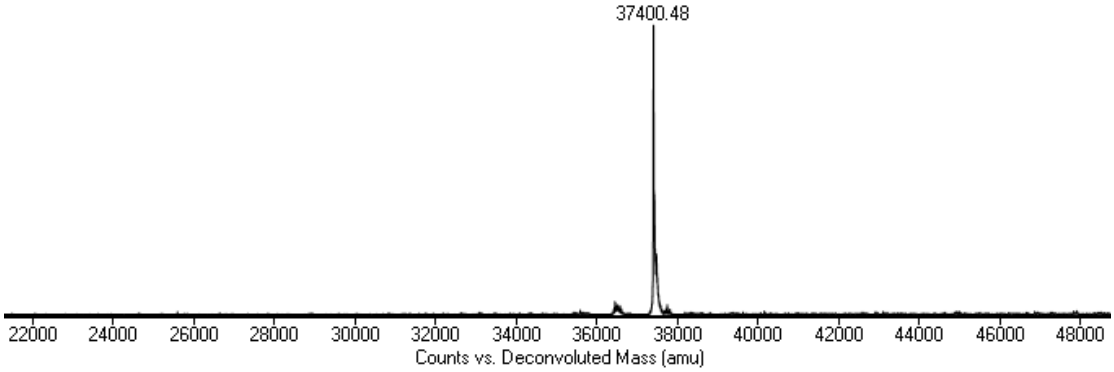


Figure S26. Analysis of ACCO catalysis co-factor requirements. 700 MHz ^{1}H NMR spectra monitoring the reaction of: a) $NaHCO_3$, L-ascorbate, Fe(II), ACC, ACCO; b) $NaHCO_3$, L-ascorbate, Fe(II), ACC; c) $NaHCO_3$, L-ascorbate, Fe(II); d) $NaHCO_3$, L-ascorbate, ACC; e) $NaHCO_3$, Fe(II), ACC; f) L-ascorbate, Fe(II), ACC; g) ACC only, h) Fe(II) only, i) L-asc only and j) $NaHCO_3$. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

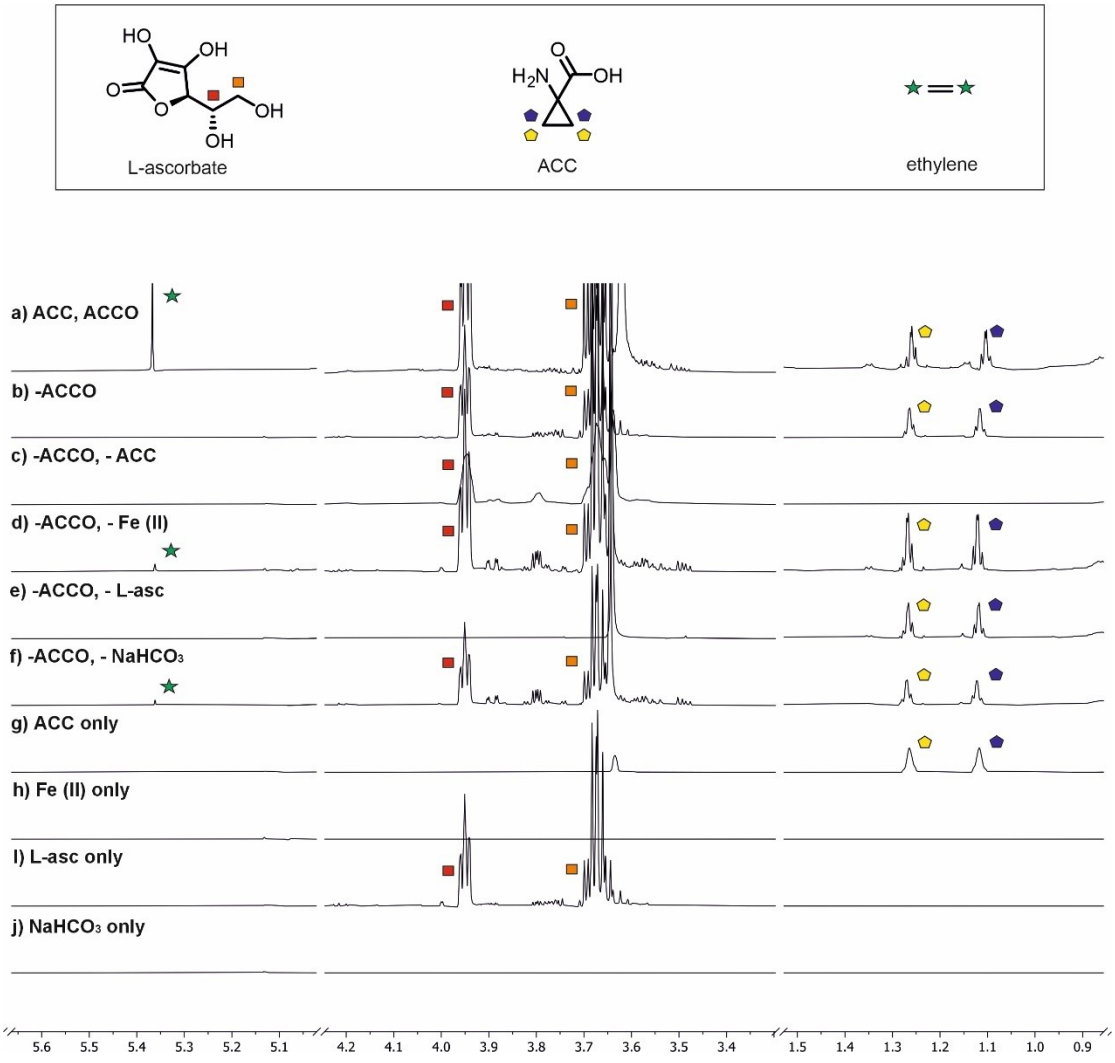


Figure S27. Analysis of ACCO catalysis with (1*S*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with (1*S*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid after: **a)** 2 h, **b)** 1 h and **c)** 0 h. **d)** ^{1}H NMR specturm of (1*S*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

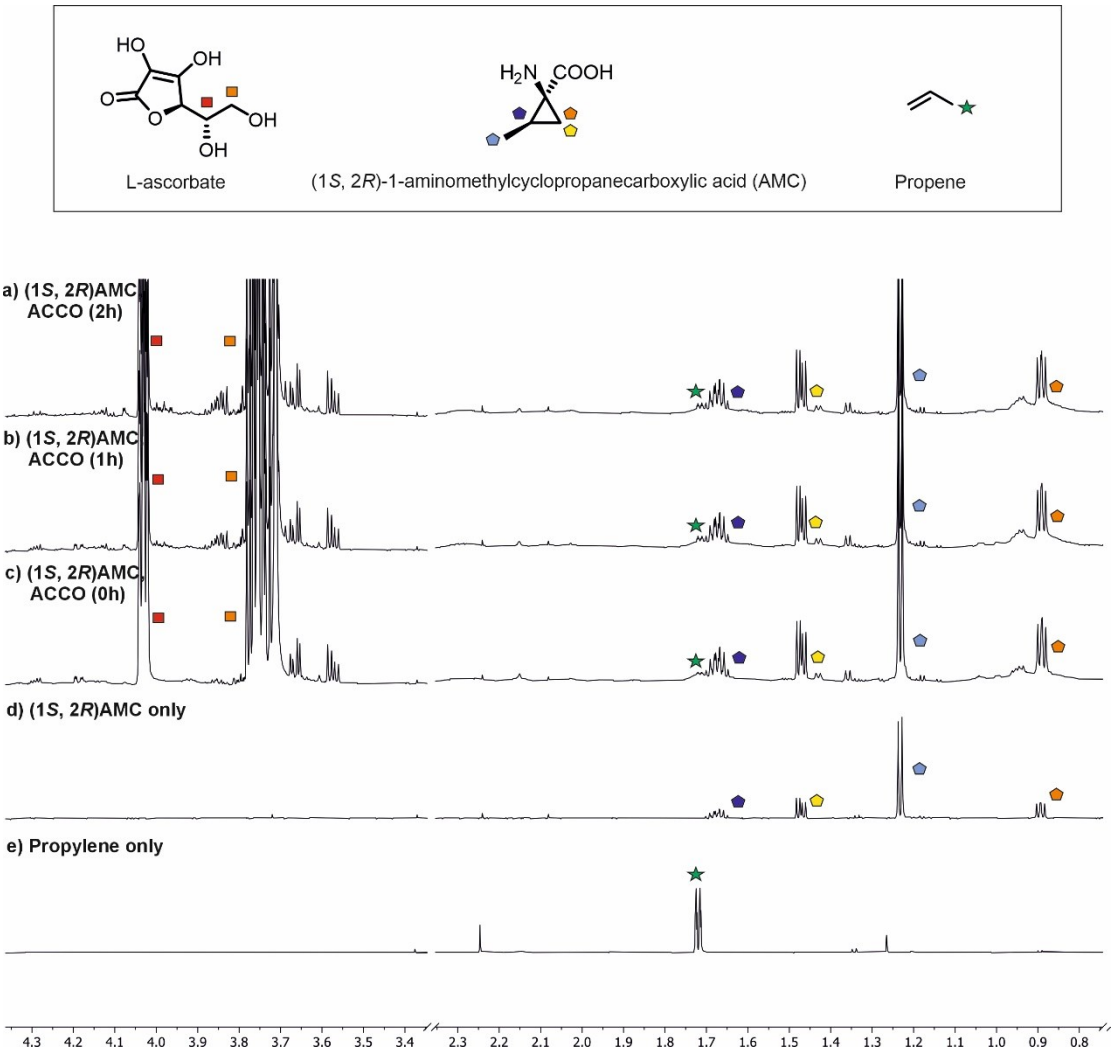


Figure S28. Analysis of ACCO catalysis with (1*S*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with (1*S*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of (1*S*, 2*S*)-1-aminomethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

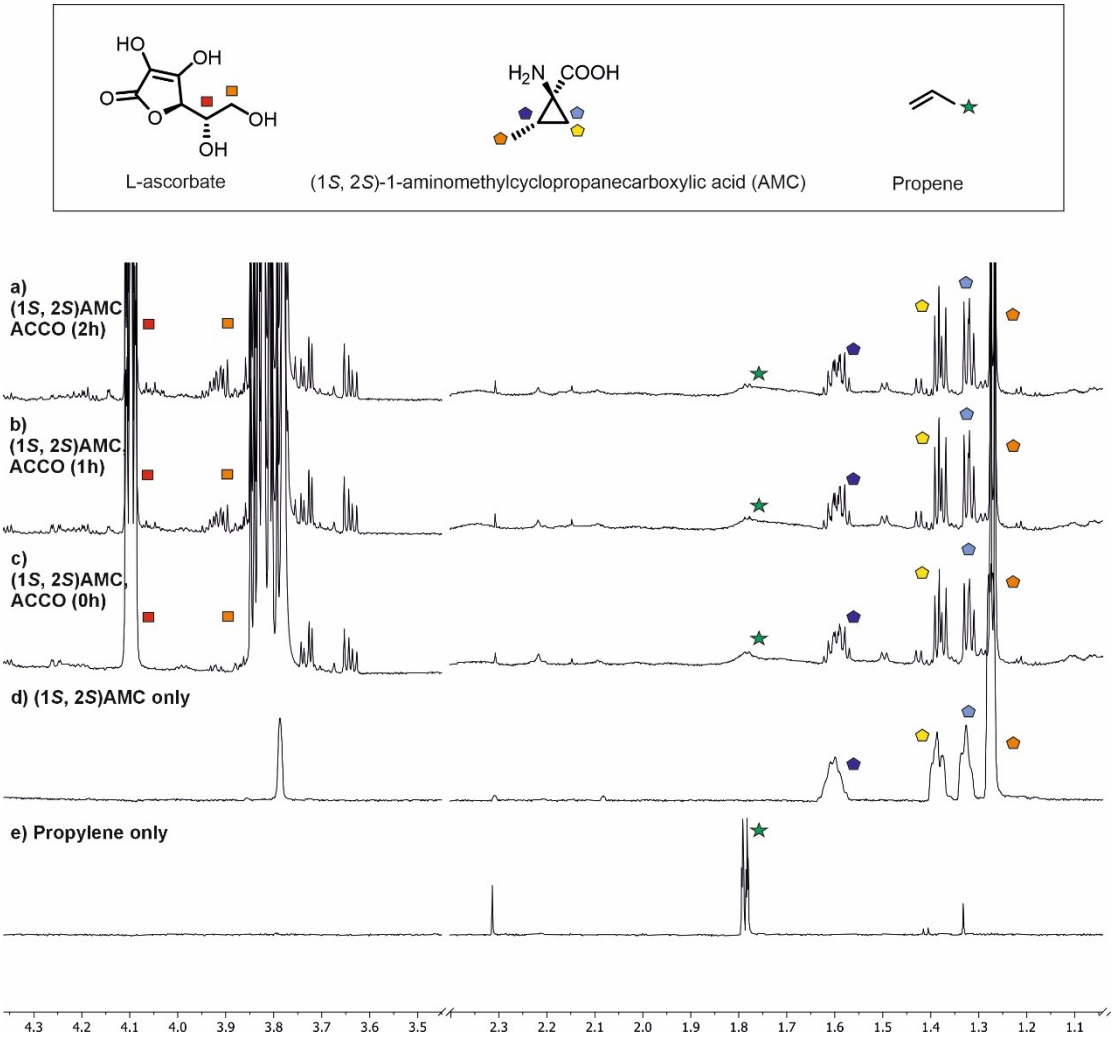


Figure S29. Analysis of ACCO catalysis with (1*R*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR of (1*R*, 2*R*)-1-aminomethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

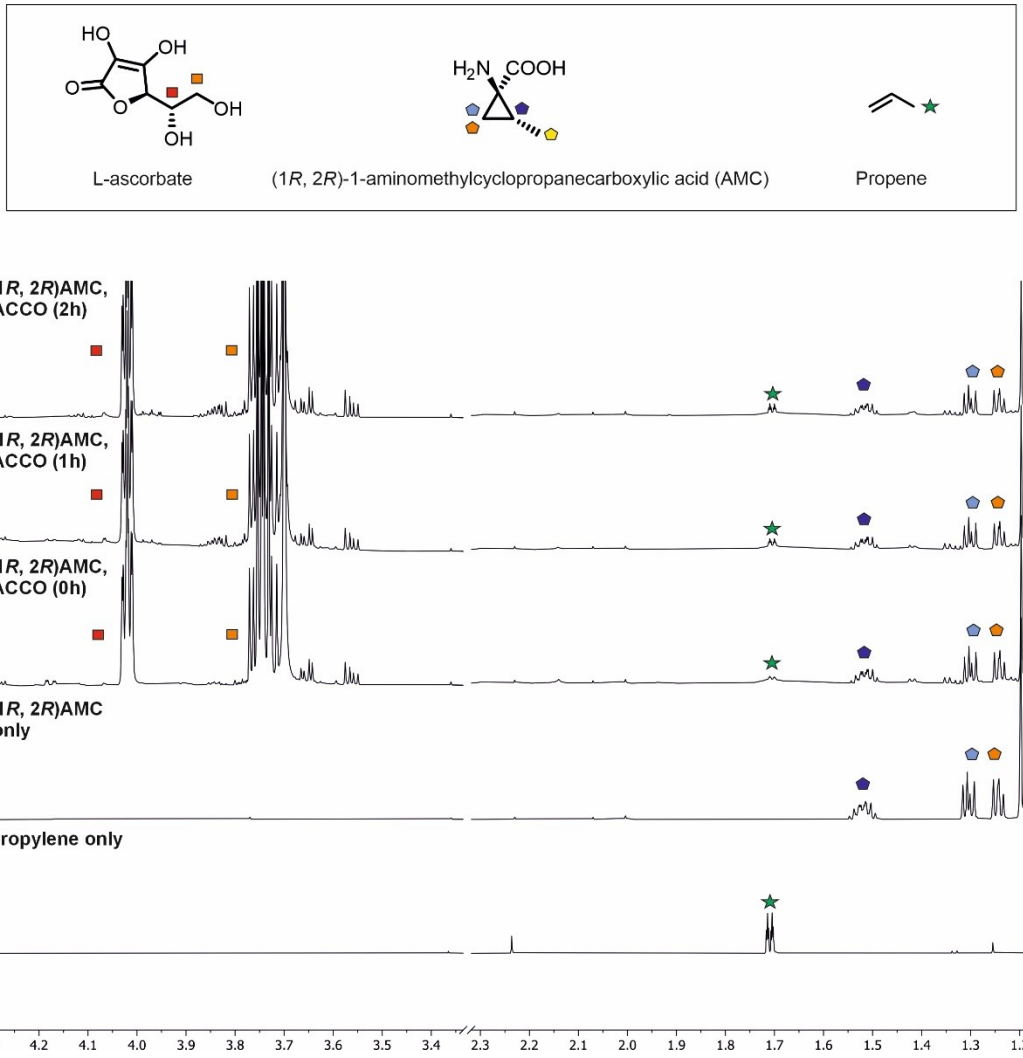
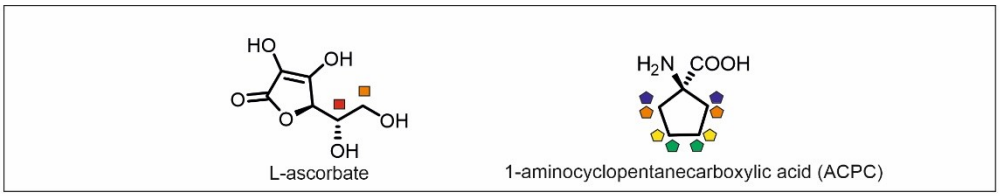


Figure S30. Analysis of ACCO catalysis with 1-aminocyclopentanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of 1-aminocyclopentanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.



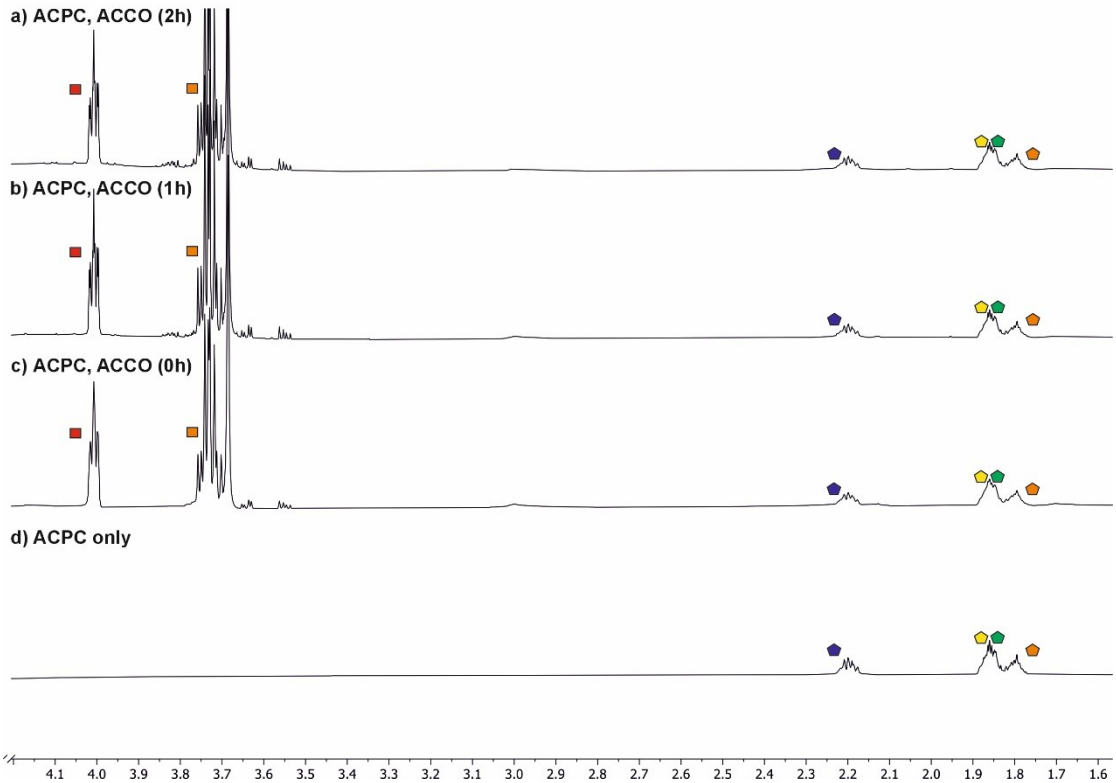


Figure S31. Analysis of ACCO catalysis with L-azetidine-2-carboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of L-azetidine-2-carboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.

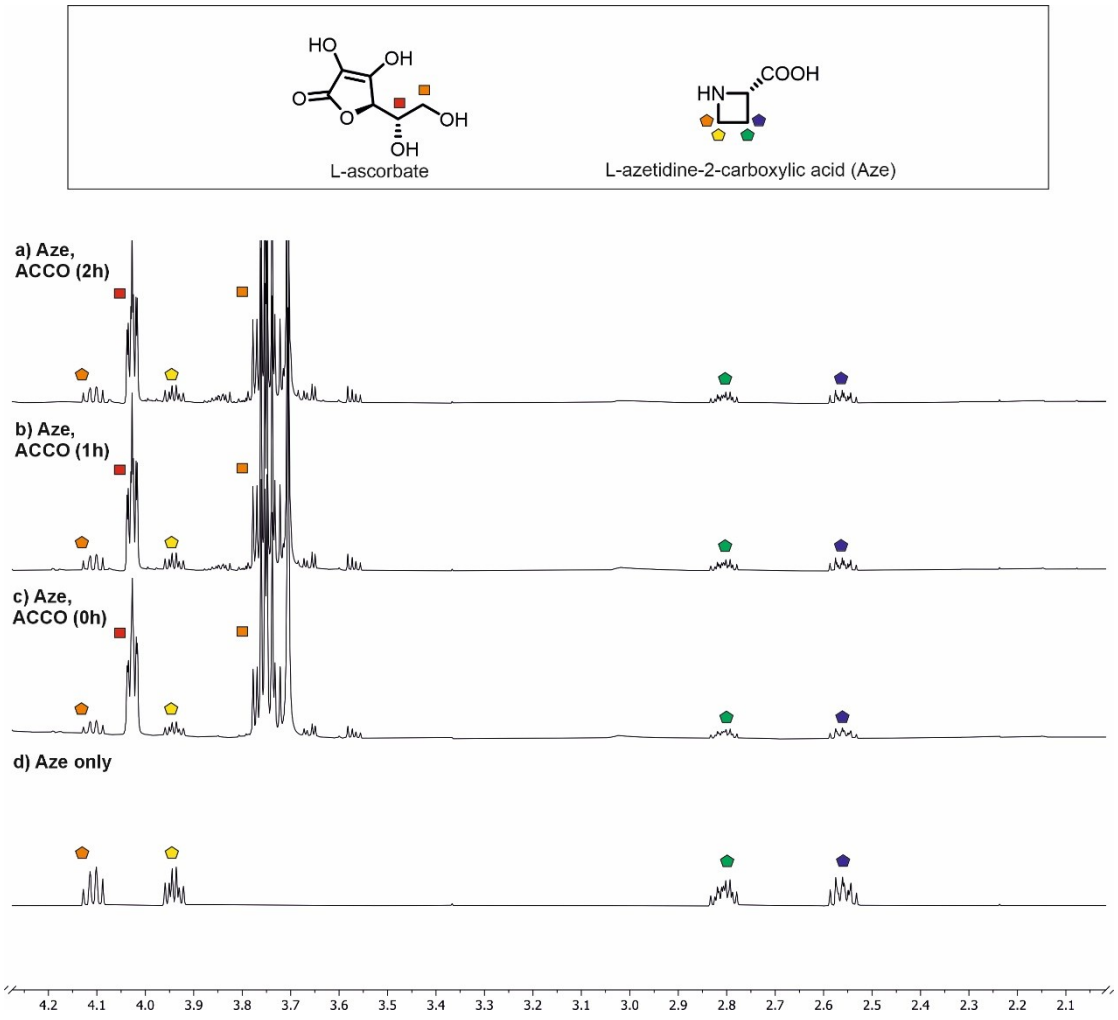


Figure S32. Analysis of ACCO catalysis with 2-amino-2-ethylbutanoic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of 2-amino-2-ethylbutanoic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.

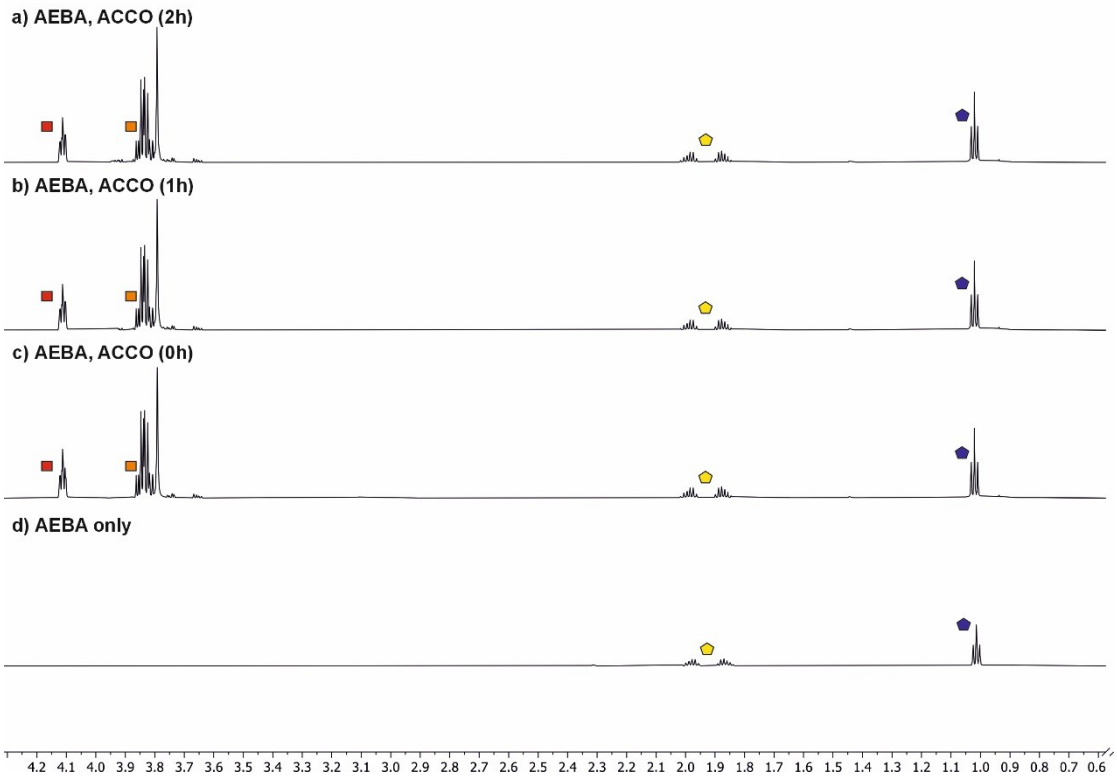
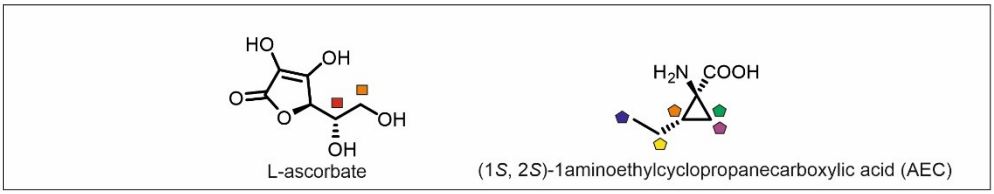


Figure S33. Analysis of ACCO catalysis with (1*S*, 2*S*)-1-aminoethylcyclopropanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of (1*S*, 2*S*)-1-aminoethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.



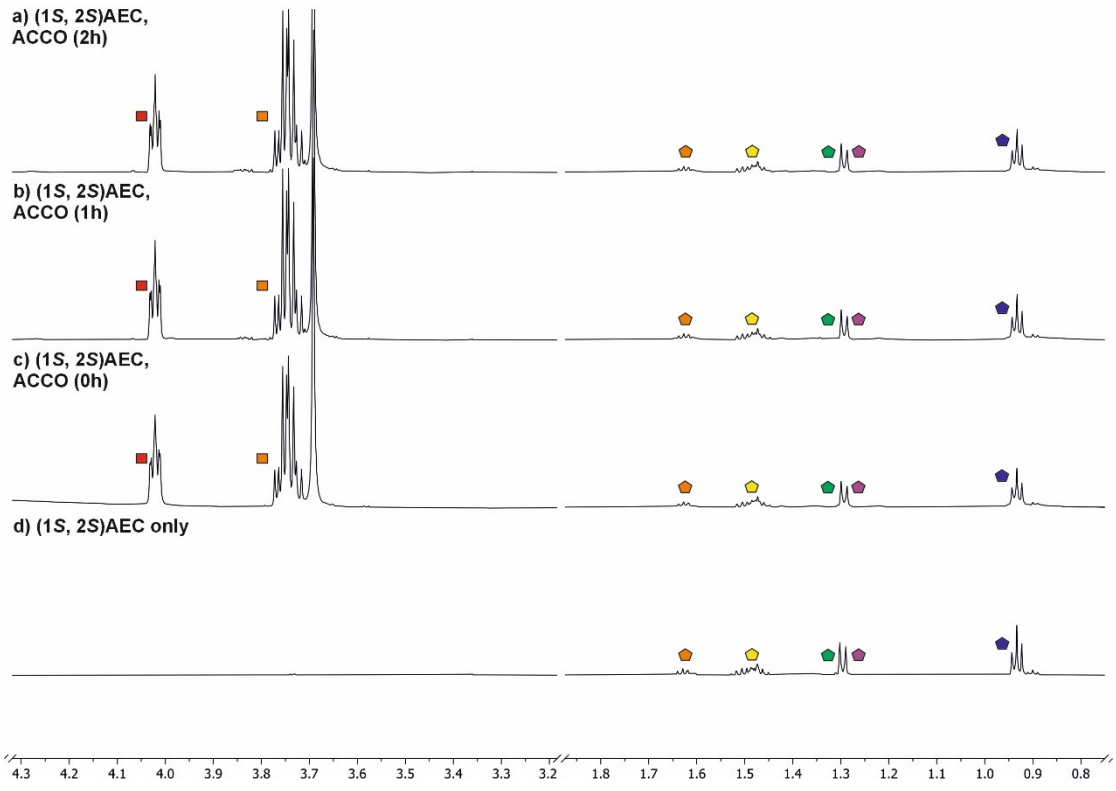


Figure S34. Analysis of ACCO catalysis with (1*R*, 2*R*)-1-aminoethylcyclopropanecarboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of (1*R*, 2*R*)-1-aminoethylcyclopropanecarboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.

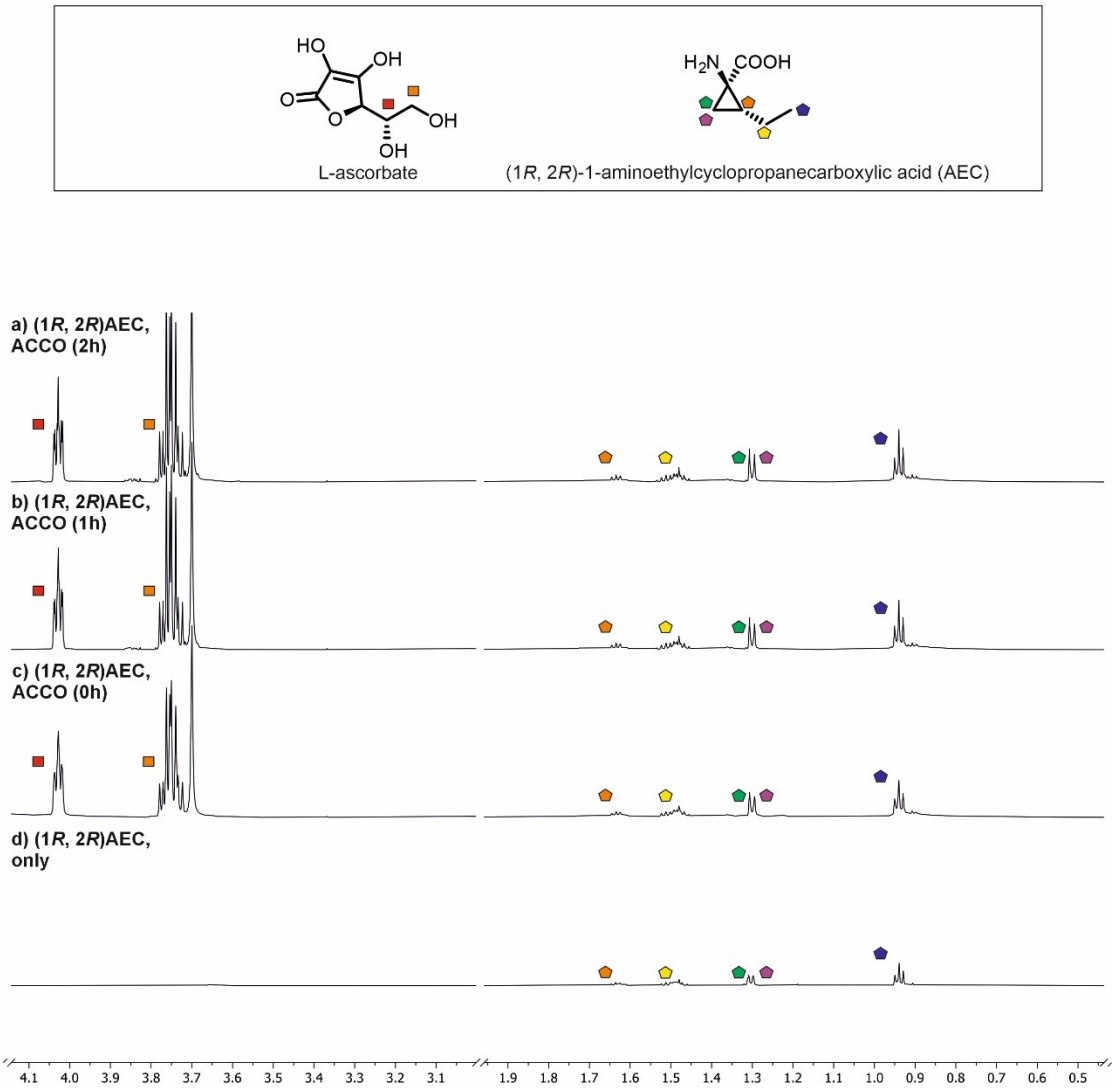


Figure S35. Analysis of ACCO catalysis with (1*R*, 2*R*, 3*S*)-1-amino-2,3-dimethylcyclopropane-1-carboxylic acid as a substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated AtACCO with ACC after **a)** 2h, **b)** 1h and **c)** 0h. **d)** ^{1}H NMR spectrum of (1*R*, 2*R*, 3*S*)-1-amino-2,3-dimethylcyclopropane-1-carboxylic acid only in buffer. Conditions: 400 µM substrate, 20 µM ACCO, 4 mM $NaHCO_3$, 3.2 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O). No observable catalytic reaction observed.

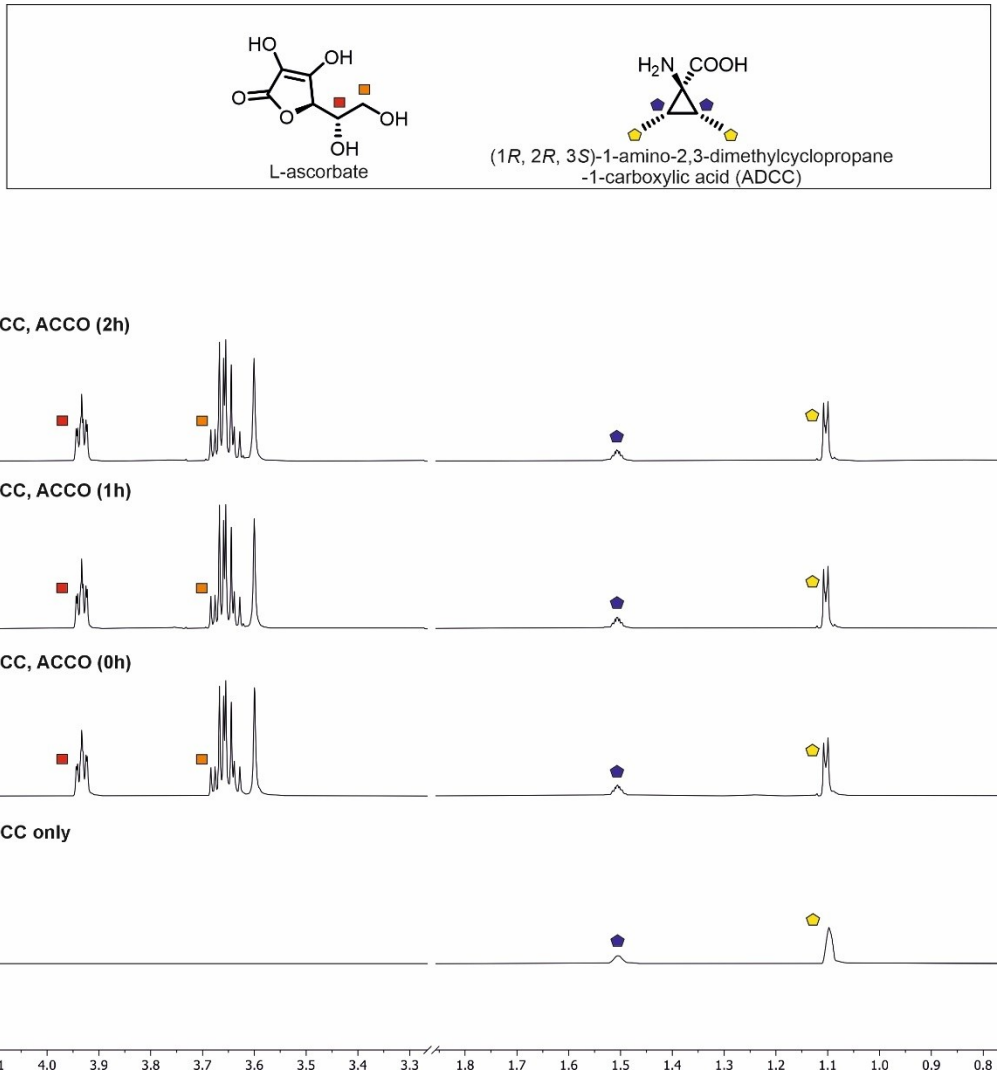


Figure S36. Analysis of Glycinea PsEFE catalysis with L-arginine as a substrate and 2OG as a co-substrate. 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated glycinea PsEFE with 2OG after **a)** 2h, **b)** 1h and **c)** 0h. ^{1}H NMR spectrum of **d)** L-arginine and **e)** 2OG only in buffer. Conditions: 400 µM 2OG, 10 µM ACCO, 0.5 mM L-Argininr, 0.5 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).

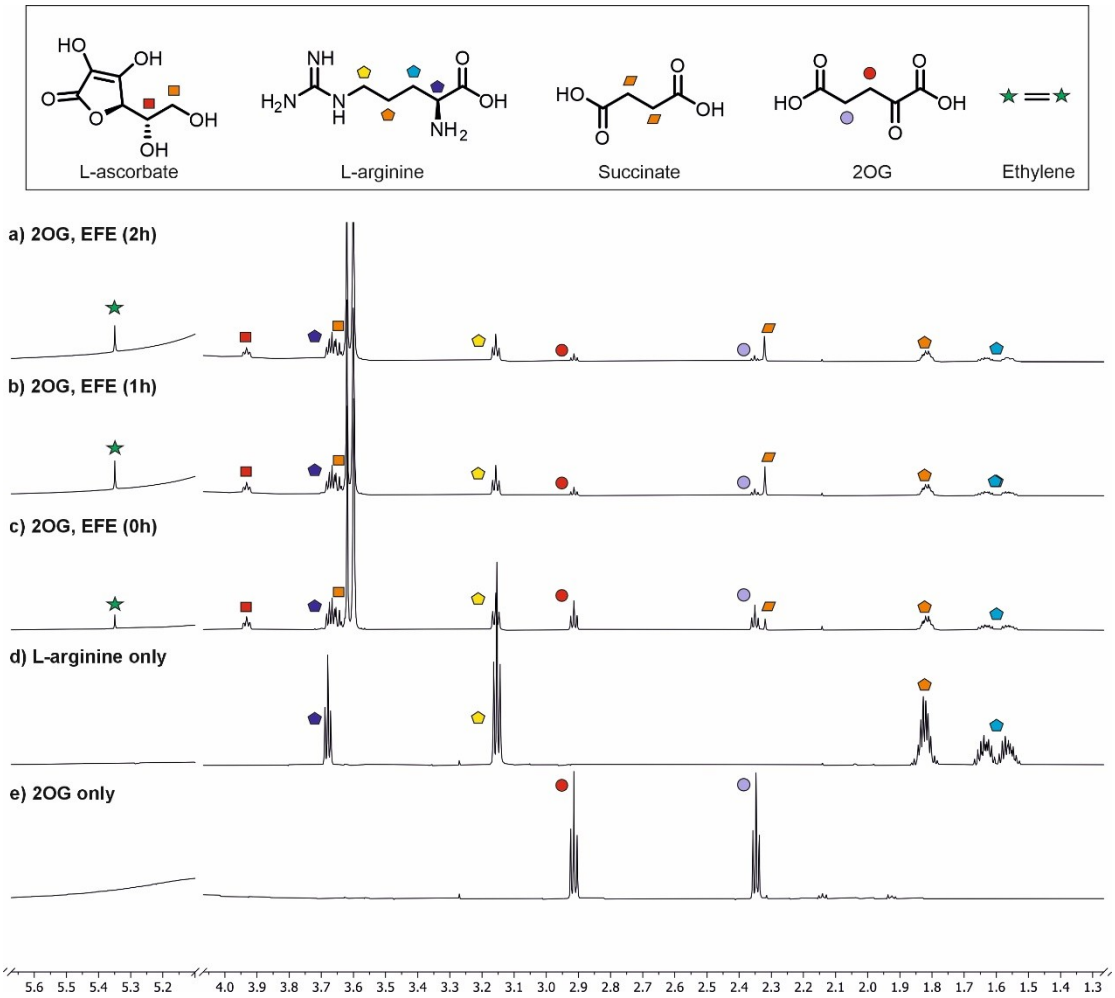
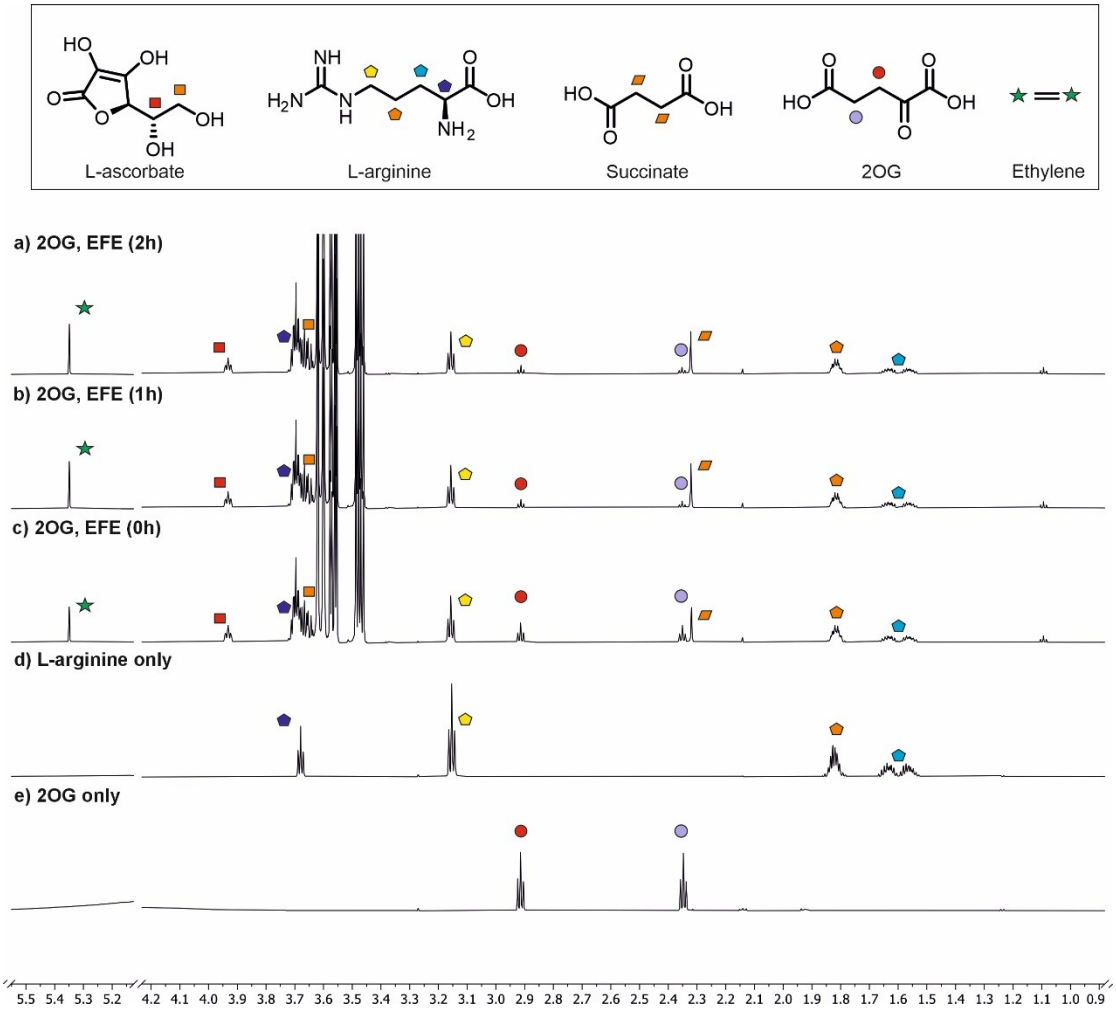


Figure S37. Analysis of 1449B PsEFE catalysis with L-arginine as a substrate and 2OG as a co-substrate. Representative 700 MHz ^{1}H NMR spectra monitoring the reaction of isolated 1449B PsEFE with 2OG after **a)** 2h, **b)** 1h and **c)** 0h. ^{1}H NMR spectrum of **d)** L-arginine and **e)** 2OG only in buffer. Conditions: 400 µM 2OG, 10 µM ACCO, 0.5 mM L-Argininr, 0.5 mM L-ascorbate, 50 µM Fe (II), 800 µM TMSP-d_4 in 50 mM potassium phosphate buffer at pH 8 (10% $_{v/v}$ D_2O).



Supplementary Table S-1. Primer pairs designed for site-directed mutagenesis of AtACCO encoded on a pET28a-SUMO vector. Codon-optimised substitutions for *E. coli* expression are highlighted in red.

Mutation	Primer pairs (5'-3')	Length

s			(nt)
K159A	fwd	agttcatatccaccgtgtcc	20
	rev	ggtggatatgaactaactgcggtgccgaaggtcgg	35
K159E	fwd	agttcatatccaccgtgtcc	20
	rev	ggtggatatgaactaacttcggtgccgaaggtcgg	35
K159R	fwd	agttcatatccaccgtgtcc	20
	rev	ggtggatatgaactaaccctggtgccgaaggtcgg	35
K290A	fwd	tccaggcaaaggagcc	16
	rev	cctttgcctggaatgcctgaccaacataaagcttcatatag	41
K290E	fwd	tccaggcaaaggagcc	16
	rev	cctttgcctggaattcctgaccaacataaagcttcatatag	41
K290R	fwd	tccaggcaaaggagcc	16
	rev	cctttgcctggaatctctgaccaacataaagcttcatatag	41
Y91F	fwd	aaggtgctttcccaatcc	18
	rev	ggaaagcacctttttcctgcgccacctg	28
Y91H	fwd	aaggtgctttcccaatcc	18
	rev	ggaaagcacctttcacctgcgccacctg	28
Y163F	fwd	tgaactaactttggtgccg	19
	rev	ccaaagttagttcatttccaccgtgtccgc	30
R245Q	fwd	attaccattggtctgcgc	18
	rev	agaccaatggtaatcagatgagcatcgcgtcctt	34
S247A	fwd	catgcgattaccattggtct	20
	rev	tggtaatcgcatggccatcgcgtccttctacaatc	35
N217Q	fwd	ataacaatgctgtgtctcagag	22
	rev	acacagcattgttattcaaatcggcgaccaactg	34
L196A	fwd	aactgctgaaggacgg	16
	rev	tccttcagcagttgagcgccagacacctgcg	31
Y286F	fwd	gttggtcagaaattccaggc	20
	rev	gaatttctgaccaacaaaaagcttcatatagtcctcaaacac	42
N253A	fwd	tagaaggacgcgatgctc	18
	rev	cgcgtccttctacgcgccgggtagcgatgct	31
N253Q	fwd	tagaaggacgcgatgctc	18
	rev	cgcgtccttctaccagccgggtagcgatgct	31

Supplementary Table S-2. Primer pairs designed for site-directed mutagenesis of AtACCO encoded on a pET28a-Tev vector. Codon-optimised substitutions *E. coli* expression are highlighted in red.

Mutations	Primer pairs (5’-3’)	Length

			(nt)
V185G	fwd	acccgcatcggtgt	14
	rev	ccgatgcgggtggcctgggactgctgttccaggaccc	37
N217G	fwd	aatcacgatgctgtgacg	18
	rev	acagcatcgtgattggaatcggtgatcagctggaa	35
S247G	fwd	attgcgagcttctataaccc	20
	rev	tagaagctcgcaatgcccatacggttaccgttggttt	37
L187A/ S247A	fwd	atcggtgtgcgcac	14
	rev	cgcacaccgatgcgggtggtctggtgctggcgtttcaggacccgcaggttt	51

Supplementary Table S-3. DNA and protein sequences for AtACCO as encoded on a pET28-SUMO vector for *E. coli* expression.

ACCO	GFSFPVVDLQELEGGERKSAMELINDACENWGFFEVVNHGLS

(SUMO) Protein Sequence	QEFMDQVESLTKEHYRKYMEKRFKDEVAERVLKKEEEVKDLD WESTFYLRHLPSSNISEIPDLDHEYRRVMKEFAGVIEKLAEKLL DVLCENLGLEKGYLKKAFQGKNGYPTFGTKVSSYPPCPRPELV KGLRAHTDAGGLVLLFQDPQVSGLQLLKDGEWVDVPPLRHSI VINIGDQLEVITNGRYKSVMHRVVAQTNGNRMSIASFYNPGSD AVIFPAPTLLKKETAEYPKFVFEDYMKLYVGQKFQAKEPRFET MKAMETVSLGPIATA
ACCO (SUMO) DNA Sequence	GGATTCTCATTTCCCGTAGTTGATCTACAAGAATTGGAAGGT GGCGAACGTAAGAGCGCAATGGAACTGATCAACGACGCCTG CGAGAACTGGGGTTTTTTCGAGGTCGTGAATCATGGTTTAAG CCAGGAGTTCATGGATCAGGTAGAATCGCTGACCAAAGAGC ACTACCGCAAATACATGGAGAAACGTTTCAAAGACGAAGTT GCTGAGCGCGTGCTGAAGAAAGAGGAAGAGGTGAAGGACC TGGATTGGGAAAGCACCTTTTACCTGCGCCACCTGCCGTCCT CTAACATTAGCGAAATTCCGGATCTGGACCATGAATATCGTCG TGTCATGAAGGAGTTCGCCGGGGTGATCGAAAAGCTGGCGG AGAAGCTGTTGGACGTGCTGTGCGAAAACCTGGGTTTGGAG AAGGGCTACCTGAAAAAGGCGTTTCAAGGTAAAAACGGCTA CCCGACCTTCGGCACCAAAGTTAGTTCATATCCACCGTGTCC GCGTCCGGAACTGGTTAAAGGCTTGCGTGCGCACACCGATG CGGGTGGTCTCGTCTTGTTGTTTCAAGATCCGCAGGTGTCTG GCCTGCAACTGCTGAAGGACGGCGAATGGGTTGATGTTCCG CCTCTGAGACACAGCATTGTTATTAACATCGGCGACCAACTG GAGGTGATCACCAACGGCCGTTATAAAAGCGTTATGCATCGT GTTGTTGCGCAGACCAATGGTAATCGCATGAGCATCGCGTCC TTCTACAATCCGGGTAGCGATGCTGTGATCTTTCCGGCACCG ACGCTGCTCAAGAAGGAGACGGCGGAATACCCGAAATTCGT GTTTGAGGACTATATGAAGCTTTATGTTGGTCAGAAATTCCA GGCAAAGGAGCCGCGTTTTGAGACGATGAAGGCTATGGAAA CCGTTTCCTTAGGTCCGATTGCGACTGCC

Supplementary Table S-4. DNA and protein sequences for AtACCO as encoded on a pET28-TEV vector for *E. coli* expression.

ACCO (TEV) Protein Sequence	MGFSFPVVDLQELEGGERKSAMELINDACENWGFFEVVNHGL SQEFMDQVESLTKEHYRKYMEKRFKDEVAERVLKKEEEVKDL DWESTFYLRHLPSSNISEIPDLDHEYRRVMKEFAGVIEKLAEKL LDVLCENLGLEKGYLKKAFQGKNGYPTFGTKVSSYPPCPRPEL VKGLRAHTDAGGLVLLFQDPQVSGLQLLKDGEWVDVPPLRHS IVINIGDQLEVITNGRYKSVMHRVVAQTNGNRMSIASFYNPGSD AVIFPAPTLLKKETAEYPKFVFEDYMKLYVGQKFQAKEPRFET MKAMETVSLGPIATA
ACCO (TEV) DNA Sequence	ATGGGTTTCAGCTTTCCGGTGGTTGACCTGCAGGAACTGGA GGGTGGCGAACGCAAGAGCGCGATGGAGCTGATTAACGATG CGTGCGAAAACTGGGGTTTCTTTGAGGTGGTTAACCACGGC CTGAGCCAGGAGTTCATGGACCAAGTTGAAAGCCTGACCAA GGAGCACTACCGTAAGTACATGGAAAAGCGTTTCAAGGATG AAGTGGCGGAGCGTGTTCTGAAGAAAGAGGAAGAGGTGAA AGACCTGGATTGGGAGAGCACCTTCTACCTGCGTCACCTGCC GAGCAGCAACATTAGCGAAATCCCGGACCTGGATCACGAGT ATCGTCGTGTGATGAAGGAATTTGCGGGCGTTATCGAAAAGC TGGCGGAGAAACTGCTGGACGTTCTGTGCGAAAACCTGGGT CTGGAGAAAGGCTATCTGAAGAAAGCGTTCCAGGGCAAGAA CGGCTACCCGACCTTTGGTACCAAAGTTAGCAGCTATCCGCC GTGCCCGCGTCCGGAACTGGTGAAAGGTCTGCGTGCGCACA CCGATGCGGGTGGCCTGGTGCTGCTGTTCCAGGACCCGCAG GTTTCTGGTCTGCAACTGCTGAAAGATGGTGAATGGGTGGAT GTTCCGCCGCTGCGTCACAGCATCGTGATTAACATCGGTGAT CAGCTGGAAGTTATTACCAACGGCCGTTACAAGAGCGTGATG CACCGTGTGGTTGCGCAAACCAACGGTAACCGTATGAGCATT GCGAGCTTCTATAACCCGGGCAGCGACGCGGTTATCTTTCCG GCGCCGACCCTGCTGAAGAAAGAAACCGCGGAGTACCCGA AGTTCGTTTTTGAGGATTACATGAAACTGTATGTGGGTCAGA AGTTCCAAGCGAAAGAACCGCGTTTTGAGACCATGAAAGCG ATGGAAACCGTGAGCCTGGGTCCGATTGCGACCGCG

Supplementary Table S-5. DNA and Protein sequences for 1449B PsEFE as encoded on a pET28-TEV vector for *E. coli* expression.

1449B PsEFE Protein Sequence	MTNLQTFELPTEVIGSAADISLGRALIQAWQKDGILQIKTDSEQ NRKTQEAMAASKQFCKEPLTFKSSCVSDLTYSGYVASGEEVTA GKPDFPEIFTVCKDLPVSDQRVKAGWPCHGPVPWPNNTYQKS MKAFMGELGLAGERLLKLTALGFELPINTFTDLTRNGWHHMR VLRFPPQTSTMSSGIGAHTDYGLLVIAAQDDVGGLYIRPPVEGE KRNRNWLPGESSAGMFEHDDPWTYVTPVQNVWTVFPGDILQF MTCGQLLSTPHKVRLNTRERFACAYFHEPNFEACAYQVFEPSG NERIHYGEHFTSMFMRCYPDRITTKRIHKDNRLAHFK
1449B PsEFE Protein Sequence	ATGACAAATTTACAAACTTTTGAACTACCCACCGAAGTAATC GGCAGCGCTGCCGATATCAGCCTGGGCCGTGCACTTATCCAG GCATGGCAGAAGGACGGCATCCTTCAGATTAAAACCGACAG CGAACAAAACCGTAAAACCCAAGAAGCAATGGCGGCGAGC AAACAGTTCTGCAAGGAGCCGCTGACCTTCAAGAGCAGTTG CGTCAGCGATTTAACTTACAGCGGCTATGTTGCATCCGGGGA GGAAGTGACGGCTGGCAAACCGGACTTCCCGGAAATTTTTA CGGTGTGCAAGGACTTGCCAGTCAGCGACCAGCGTGTGAAA GCCGGCTGGCCGTGTCATGGCCCGGTTCCGTGGCCCAACAA CACGTACCAAAAGTCCATGAAGGCGTTTATGGGTGAACTGG GTCTGGCGGGCGAGCGCCTGCTGAAATTGACCGCTCTGGGT TTTGAGCTGCCGATTAACACCTTTACTGACCTGACGCGTAAC GGTTGGCATCATATGCGCGTGCTGCGTTTTCCGCCACAGACC TCGACCATGTCCTCTGGCATTGGTGCGCACACCGACTATGGT CTGTTGGTTATTGCGGCACAGGATGATGTTGGTGGTCTCTACA TCCGCCCTCCGGTTGAAGGTGAGAAACGTAACCGCAACTGG CTGCCGGGTGAGTCTTCCGCGGGTATGTTCGAGCACGATGAC CCGTGGACGTACGTGACCCCGGTTCAAAATGTTTGGACCGTC TTCCCGGGTGATATTCTGCAATTTATGACCTGTGGTCAGCTGT TGTCGACCCCGCATAAGGTGCGCTTGAATACCCGTGAACGTT TCGCCTGCGCGTACTTCCACGAACCGAACTTTGAGGCGTGC GCCTACCAAGTGTTCGAGCCGAGCGGCAATGAACGTATCCA CTATGGCGAGCACTTTACGAGCATGTTCATGCGTTGTTATCCG GATCGTATCACCACCAAACGCATCCACAAGGATAATAGACTG GCGCATTTCAAGTAA

Supplementary Table S-6. Protein sequence of proteases used during recombinant protein production.

SenPII Protease	MEKEISNALGHGPQDEILSSAFKLRITRGDIQTLKNYHWLNDEVINFYMNLLV ERNKKQGYPALHVFSTFFYPKLKSGGYQAVKRWTKGVNLFEQEIILVPIHRKV HWSLVVIDLRKKCLKYLDSMGQKGHRICEILLQYLQDESKTKRNSDLNLLEW THHSMKPHEIPQQLNGSDCGMFTCKYADYISRDKPITFTQHQMPLFRKKMVW EILHQQLL
TEV Protease	GHIVWPDYANILKEVFGGARMACVTSAHMAGANGSILKKAEETSRATMHKP VIFGEDYVTEADLPYTPLHLEVNAEMERMYYLGRRALTHGKRRKVSVNNKR NRRRKVAKTYVGRDSIVEKIVVPHTERKVDTTTAVKDTCNEVSTQLVHNSMP KRKKQKNFLPATSLSNVYAQTWSIVRKRHMQVEIISKKSVRAKVKRFEGSVQL FASVRHMYGERKRVDLRIDNWQQKTLLDLAKRFKNERVDQSKLTFGSSGLVL RQGSYAPAHWYRHGMFIVRGRSDGMLVDARAKVTFAVCYSMTHY