UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE 222

Laboratoire de Catalyse Chimique

Institut de Science et d'Ingénierie Supramoléculaires (ISIS)

THÈSE de DOCTORAT

Présentée par :

Jing YI

Soutenue le : **30 Septembre 2022**

Pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie

Nonenzymatic analog of pyrimidine ribonucleotide biosynthesis

THÈSE dirigée par :

Prof. Joseph MORAN Directeur de thèse, Université de Strasbourg

RAPPORTEURS :

Dr. Yann TROLEZ Rapporteur, Université de Rennes

Prof. Erwan POUPON Rapporteur, Université de Paris-Saclay

AUTRES MEMBRES DU JURY :

Prof. Stéphane VUILLEUMIER Examinateur, Université de Strasbourg

Dedication

To my furry exuberant troublemaker Okay,

my sweetest selflessness intelligent and only sister 老姐,

my beloved parents 老爸老妈 and grandparents 外公外婆,

my soulmate, partner in crime 宝贝,

to people who lost their relatives or friends during COVID-19 and the Russo-Ukrainian War.

May the dead rest in peace and cherish the present.

愿逝者安息

愿世界和平

Table of Contents

List of Abbreviations

Ac	acetyle	**LC**	liquid chromatography
ALA	alanine	**LUCA**	Last Universal Common Ancestor
ASP	aspartic acid	**M**	molar
aTMP	amido triphosphate	**mg**	milligram
ATP	adenosine triphosphate	**MHz**	mega hertz
BC	before Christ	**mL**	milliliter
CAA	N-carbamoyl aspartic acid	**mM**	millimolar
CAP	carbamoyl aspartic acid	**mmol**	millimole
CO	carbon monoxide	**MS**	mass spectrometer
CO_2	carbon dioxide	**m/z**	mass to charge ratio
CoA	Coenzyme A	**NMR**	nuclear magnetic resonance
CTP	cytidine triphosphate	**°C**	degree celsius
Da	dalton	**OMP**	orotidine monophosphate
DAP	diamido phosphate	**ORO**	orotic acid
DCM	dichoromethane	**PEP**	phosphoenolpyruvate
DHO	dihyroorotic acid	**PNA**	protein nucleic acid
DMSO	dimethyl sulfoxide	**ppm**	part per million
DNA	deoxyribonucleic acid	**PRPP**	5-Phospho-D-ribose 1-diphosphate
equiv.	equivalent	**QTOF**	quadrupole time of flight
EtOAc	ethyl acetate	**R5P**	ribose 5-phosphate
GLY	glycine	**RNA**	ribonucleic acid
GLN	glutamine	**rTCA**	reductive tricarboxylic acid
GLU	glutamic acid	**r.t.**	room temperature
GTP	guanosine triphosphate	**s**	singlet
h	hour	**t**	time
H_2S	hydrogen sulfide	**temp.**	temperature
HAA	hydantoin 5-acetic acid	**TfOH**	triflic acid
HCl	hydrochloric acid	**TMP**	trimetaphosphate
HCN	hydrogen cyanide	**tR**	retention time
HFIP	1,1,1,3,3,3-hexafluoro-2-propanol	**URA**	uracil
HMP	sodium hexametaphosphate	**UMP**	uridine monophosphate
Hz	hertz	**UTP**	uridine triphosphate
IMP	inosine monophosphate	**UV**	ultra-violet
kJ	kilo joule	**V**	volt
KOCN	potassium cyanate	**w/w**	weight by weight
KOH	potassium hydroxide	**µL**	microliter

Acknowledgements

Firstly, I would like to thank my supervisor Prof. Joseph Moran for offering me the opportunity to work in this lab and for his efforts to help me with my fellowship application. His constant and patient guidance during my PhD helped me to develop my own way of thinking in science and in life. I've learned so much from him about presenting research the right way both in writing and in oral presentations. His encouragement allowed me to recover from failures and quickly find a new direction despite the challenges we all face in prebiotic chemistry. Through many discussions we have had, I got the chance to expand my horizons in academic life. He's always been a great role model of my life: through his hard work, creativity, self-discipline and optimism.

I would also like to thank Dr. Kamila Muchowska, for giving me advices on research and also for helping me correct my English writing from the beginning, and of course, for taking care of the whole lab with her great organization; Dr. David Leboeuf, for helping me review my scientific paper; and Dr. Pawel Dydio, who gave a lot of ideas to improve my project during my mid term defense.

Prebiotic chemistry is a field that faces many analytical constrains, but thanks to Cyril and Jean-Louis, I was able to analyze my reactions with NMR and MS. Additionally, I want to thank Wanhyalo, for supporting my MS analysis even if I bothered him late in the evening and on weekends. I also want to express my gratitude to Annia and Nathalie, our fantastic secretaries, who helped me complete many administrative procedures, gracias querida Annia, merci chère Nathalie.

My greatest gratitude goes to all the senior lab members I've worked with. Sreejith, Edward and Vuk, thank you for tutoring my master's interships in this lab. This has taught me a lot of useful lab skills and gave me a lot of advice that is still useful to this day. Ankita, whose joining the lab pulled me out of the black hole of getting stuck with analytical techniques, our work together on the *N*-glycosylation expanded my knowledge on sugar chemistry. Thank you, dear Ankita, for being a great colleague and best friend. Robert, for his help in NMR analysis and advice on writing scientific papers. Quentin, for sharing his synthesized active phosphate. Weiqiang, new to the lab, even though we've been working together only for few months on the glycosylation project, his rigorous attitude and creativity in science has influenced me a lot. Silvana, also new to the lab, has helped me with correcting my thesis. Marie and Nik, for giving me useful advice in chemistry and regarding career choices.

I also want to thank all other colleagues. Harpreet, for helping me with the nucleobase project and for the emotional support throughout these years that means so much to me; Emilie, for the correction of the French resume of this thesis and bringing laughs and happiness to the lab, without her, I wouldn't get any ear piercing; Sophia, for her help in the nucleobase project and her invitation to the Christmas dinner during the lockdown, we're both obsessed with the plants; Shaofei, for helping me with the fellowship application and being a good friend in this foreign country; Paul, for all the interesting conversations that we have made (travels, foods, cultures and economy); Cyprien, for sharing his chocolate cookies/cakes with me and I still can't forget his galette des rois with pistachio; Joris, for the fun time we had inside and outside the office and his good taste in music and in fashion; Maciek, for inviting me to his parties and his efforts in taking care of the coffee order and coffee machine; and also Abhijit, Jan, Elodie, Lucas, Florent, Robin, Valentyn, Edouard, Capucine, Corinna, Andrei for the contribution to the lovely lab atmosphere.

I want to thank specifically my master student, Louis, although the time we've worked together in the lab was short due to the pandemic, he helped me develop the oxidation step in the nucleobase project.

Last but not least, I'd like to acknowledge the China Scholarship Concil and Volkswagen grant for their financial support of this thesis.

At the end, please accept my greatest thanks, Prof. Erwan Poupon, Dr. Yann Trolez and Prof. Stéphane Vuilleumier, for taking the time to read and evaluate my manuscript and being part of my jury.

RESUME DE LA THESE DE DOCTORAT

Vers un analogue non-enzymatique de la biosynthèse des ribonucléotides pyrimidiques

1) Introduction

Deux visions s'opposent quant aux origines chimiques de la vie : l'hypothèse d'une origine génétique et celle d'une origine métabolique.La première approche "génétique" suppose que les polymères génétiques autoréplicatifs, probablement l'ARN, ont émergé directement d'une chimie prébiotique ressemblant peu aux voies biosynthétiques de la vie. De ce point de vue, la chimie prébiotique précoce est simplement un outil pour fournir des éléments de base pour un processus ultérieur auto-organisé. Un défi central pour cette vision de la chimie prébiotique est la synthèse directe, robuste et à haut rendement de ribonucléotides ou désoxyribonucléotides par tout moyen possible. Les travaux expérimentaux réalisés dans ce sens ont sans doute été l'axe principal du domaine au cours des 50 dernières années, avec des travaux expérimentaux récents de Sutherland, Carell, Powner, Benner, et Trapp.

Cependant, sans négliger l'implication indéniable des ribonucléotides et de l'ARN dans l'origine de la vie, à partir de la fin des années 1980, des difficultés conceptuelles de l'approche d'une "origine génétique" a emergé la proposition d'une "origine métabolique". Cette approche considère que l'origine de la vie a impliqué l'apparition spontanée d'un réseau de réaction auto-organisé, déclenché par un besoin thermodynamique de relâcher les gradients redox géochimiques. Les catalyseurs initiaux permettant la réactivité auto-organisée auraient été les minéraux, les argiles et les métaux, mais certaines matières organiques produites par le réseau pourraient en outre rétroagir comme catalyseurs pour renforcer les réactions existantes et en permettre de nouvelles. Un tel réseau réactionnel aurait été historiquement conforme au métabolisme biologique. De plus, parce que les systèmes complexes, tels que le métabolisme, sont difficiles à perturber et à réécrire, certaines similitudes énergétiques et chimiques devraient exister avec le métabolisme aujourd'hui. Dans cette perspective alternative, les ribonucléotides et leurs oligomères auraient été produits par le réseau de réaction prébiotique, peut-être initialement en petites quantités, mais y auraient pris de l'importance parce qu'ils offraient une fonction catalytique ou régulatrice critique au sein de la chimie auto-organisée existante. Selon ce principe, un défi central de la chimie prébiotique est de comprendre quelles auraient été les

conditions d'une telle chimie auto-organisée. Une façon d'atteindre cet objectif consiste à rechercher des conditions qui récapitulent les variantes non enzymatiques des réactions et processus métaboliques essentiels, y compris ceux de la synthèse de molécules génétiques. Plusieurs conditions, éventuellement très différentes entre elles, pourraient être trouvées pour accéder à des réactions particulières ou voies non-enzymatiques, et certaines de ces conditions peuvent n'avoir aucun rapport avec l'origine de la vie. Cependant, cette stratégie devrait éventuellement identifier les conditions permettant à un système plus large de pouvoir s'auto-organiser. Les efforts expérimentaux dans cette direction se sont jusqu'à présent concentrés sur les analogues catalysés par les métaux des voies de fixation du carbone, la biosynthèse des acides aminés et des cofacteurs et le métabolisme des sucres. Les recherches sur les conditions purement chimiques qui déclenchent des réactions de biosynthèse des ribonucléotides dans un contexte prébiotique ont été limitées, comme nous le verrons ci-dessous.

2) Résultats et discussions

a. Transformations séquentielles

Il existe deux voies générales de biosynthèse des ribonucléotides: la voie de sauvetage qui régénère les ribonucléotides après leur dégradation et la voie *de novo* qui les construit à partir des métabolites essentiels qui sont les acides aminés, les composés C1 et le phosphoribosylpyrophosphate (**PRPP**), forme activée de ribose-5-phosphate. Les trois premières étapes de la biosynthèse *de novo* des pyrimidines sont illustrées par les flèches noires du schéma 1. Premièrement, l'aspartate (**ASP**) réagit avec le carbamoyl phosphate (**CAP**) pour donner du carbamoyl aspartate (**CAA**), cette première étape est catalysée par la carbamoyl phosphate synthétase dans le système biologique. [34] Le **CAA** subit ensuite une cyclisation déshydratante pour donner du dihydroorotate (**DHO**), la réaction est catalysée par la dihydroorotase. Le **DHO** est oxydé en orotate (**ORO**) par divers cofacteurs impliqués dans le transfert d'électrons, cette oxydation est catalysée par la dihydroorotate déshydrogénase. Plus tard dans la voie, qui n'est pas illustrée dans le schéma 1, **ORO** se couple avec le **PRPP** pour donner l'orotidine 5'-monophosphate et, dans les étapes suivantes, les ribonucléotides canoniques de pyrimidine, qui sont l'uridine-5'-monophosphate et la cytidine-5'-monophosphate. La conversion de **ASP** en **ORO** est donc l'étape déterminante dans la formation des ribonucléotides pyrimidiques.

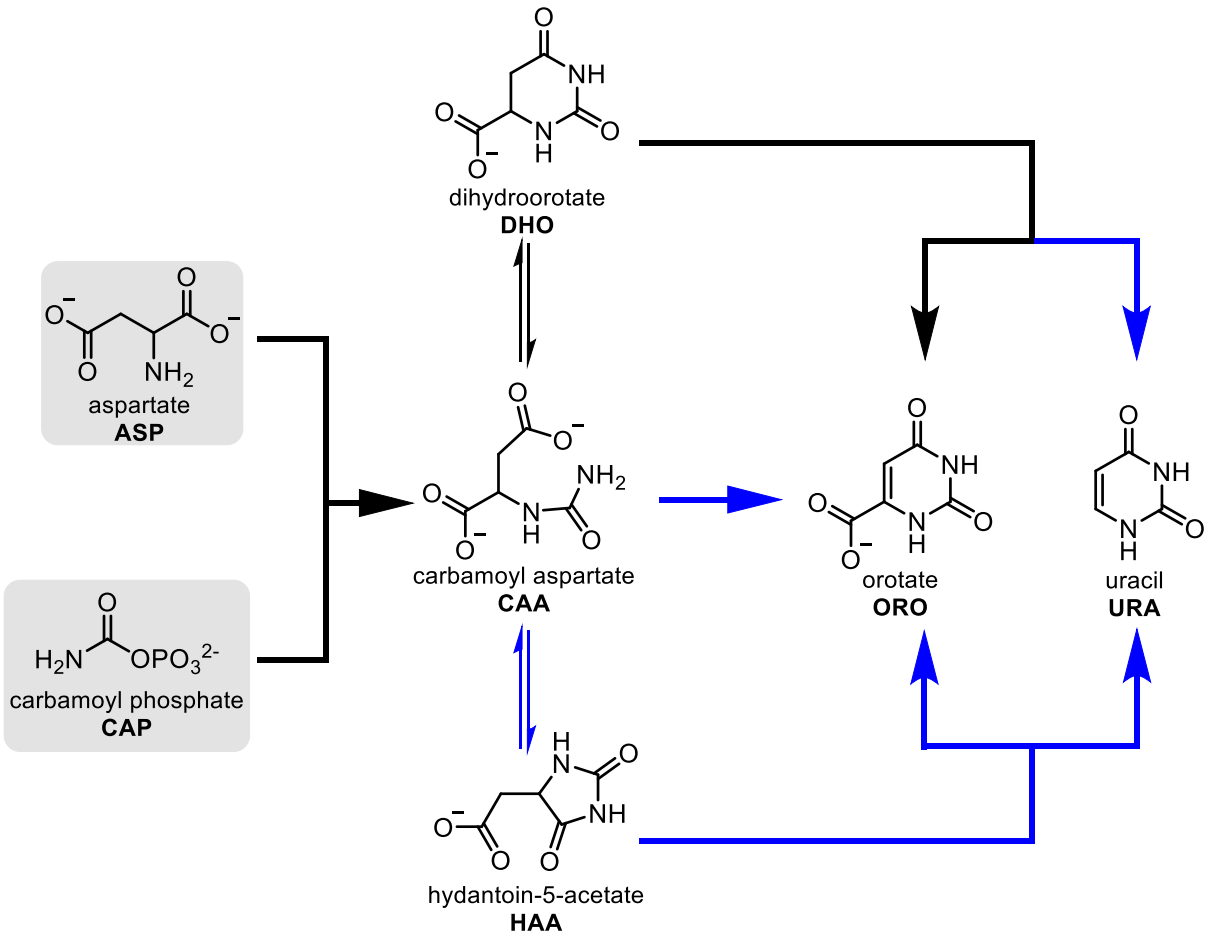


Schéma 1. Les trois premières étapes de la biosynthèse de novo des ribonucléotides pyrimidiques convertissent l'aspartate (ASP) en orotate (ORO), comme le montrent les flèches noires. Les réactions supplémentaires observées dans ce rapport sont décrites par des flèches bleues.

Au laboratoire, il a été démontré que la synthèse non-enzymatique de l'**ORO** était dépendante de la photochimie. Ici, nous examinons tout d'abord ces trois premières étapes clés individuellement: le transformation de l'**ASP** en **CAA** en utilisant le **CAP** comme donneur du groupement carbamoyl, cette étape est faite en présence de $NaHCO_3$ à 60 ºC pour 16 h, le rendement peut atteindre 63% par la mesure de ^{1}H RMN quantitative. Ensuite, la cyclisation déshydratante du **CAA** peut générer deux produits **DHO** et **HAA**, la formation du produit penta-cyclique **HAA** est favorable en général sous condition thermal, mais cette sélection peut être réversée en ajoutant du $CuSO_4$ en milieu légèrement basique, néonmoins seulement 1% de **DHO** est produit. La dernière étape avant la formation de la nucléobase concerne l'oxydation du **DHO** en **ORO**, la condition optimale est l'utilisation de MnO_2 comme oxidant, 40 mol% de $CuSO_4$ serves comme catalyst, 57% de l'**ORO** est obtenu après le chauffage pendant une nuit,

une petite quantité de l'**URA** (2%) est aussi observée avec l'**ORO**. Cette condition est aussi optimale pour convertir le produit majoritaire **HAA** en **ORO**, avec la basification de la solution en utilisant NaOH après 16 h. Donc la conversion accidentelle du **CAA** en **HAA** n'est pas nécessairement une impasse, car il peut être orienté vers le même produit final, l'**ORO**.

Pour conclure, nous constatons que ces réactions se produisent toutes de manière non enzymatique dans de l'eau à 60 ºC sans avoir besoin de lumière UV. (Figure 1) Différents

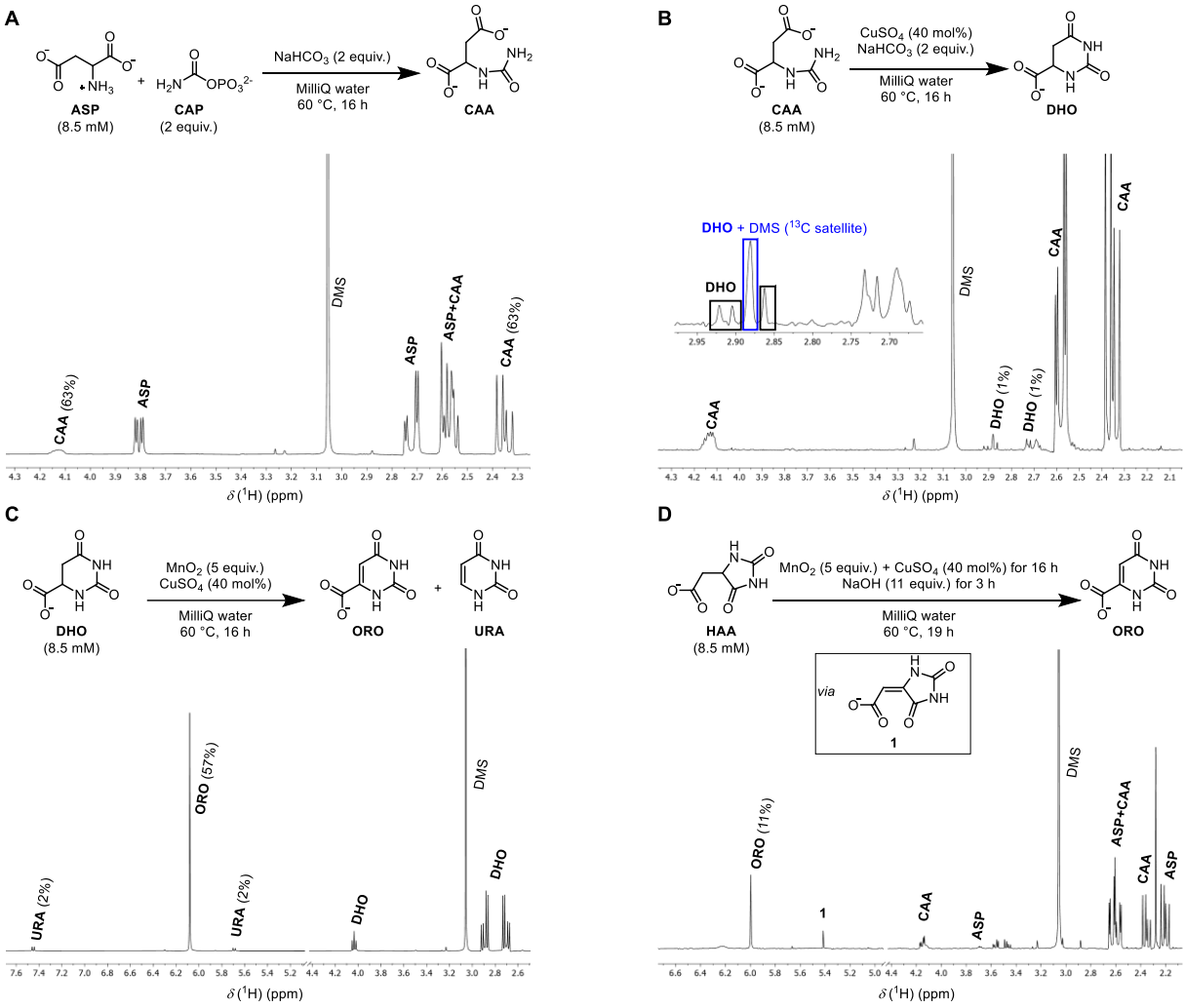


Figure 1. Etude d'étapes réactionnelles individuelles. (**A**) Spectre ^{1}H RMN des produits de réaction de l'aspartate **ASP** et du carbamoyl phosphate **CAP**. δ, déplacement chimique. (**B**) ^{1}H RMN des produits de la réaction de cyclisation de l'aspartate de carbamoyle **CAA**. (**C**) ^{1}H RMN des produits de réaction oxydative du dihydroorotate **DHO**. (**D**) ^{1}H RMN des produits de la réaction oxydative de l'hydantoïne-5-acétate (**HAA**). Les rendements ont été déterminés par ^{1}H RMN quantitative avec des techniques de suppression d'eau, la diméthylsulfone comme étalon interne.

paramètres tels que la température de la réaction, le pH, la concentration des substrats et le criblage de différents métaux ou oxidants ont été testés pour trouver les conditions optimales pour chaque étape. Les résultats présentés dans la Figure 1 ne sont pas les plus optimales, mais ils sont les adaptés pour trouver une condition universelle dans le monde primitif.

b. Transformation monotope

Notre intérêt s'est ensuite porté sur l'évaluation de la faisabilité de rassembler ces trois étapes de transformation en un seul pot. Nous avons découvert un analogue de la biosynthèse des pyrimidines dans lequel certaines étapes sont favorisées par des métaux. La version en un seul pot de cette transformation a été réalisée par l'addition successive des réactifs et par des changements de pH à des temps donnés. Malgré la stabilité chimique des différents intermédiaires réctionnels, cette intervention par l'humain est nécessaire. Par exemple, le **CAP** se décompose rapidement en présence du sel Cu^{2+}, le **DHO** n'est pas stable et s'hydrolyse en **CAA** en milieu alkalin sous NaOH. Ces réactions indésirables ne sont pas évitables dans la solution car leurs cinétiques sont plus favorables que les transformations désirables. Au final, les conditions réactionnelles de cette transformation en un seul pot ont été choisies parmi les conditions optimisées dans l'étude séquentielle précédente et ont été présentées dans dans le schéma. (Figure 2) Ces résultats sont analysés également par la mesure de ^{1}H RMN quantitative mais aussi par la LC-QTOF/MS avec une courbe de calibration et l'acid maleic comme étalon interne.

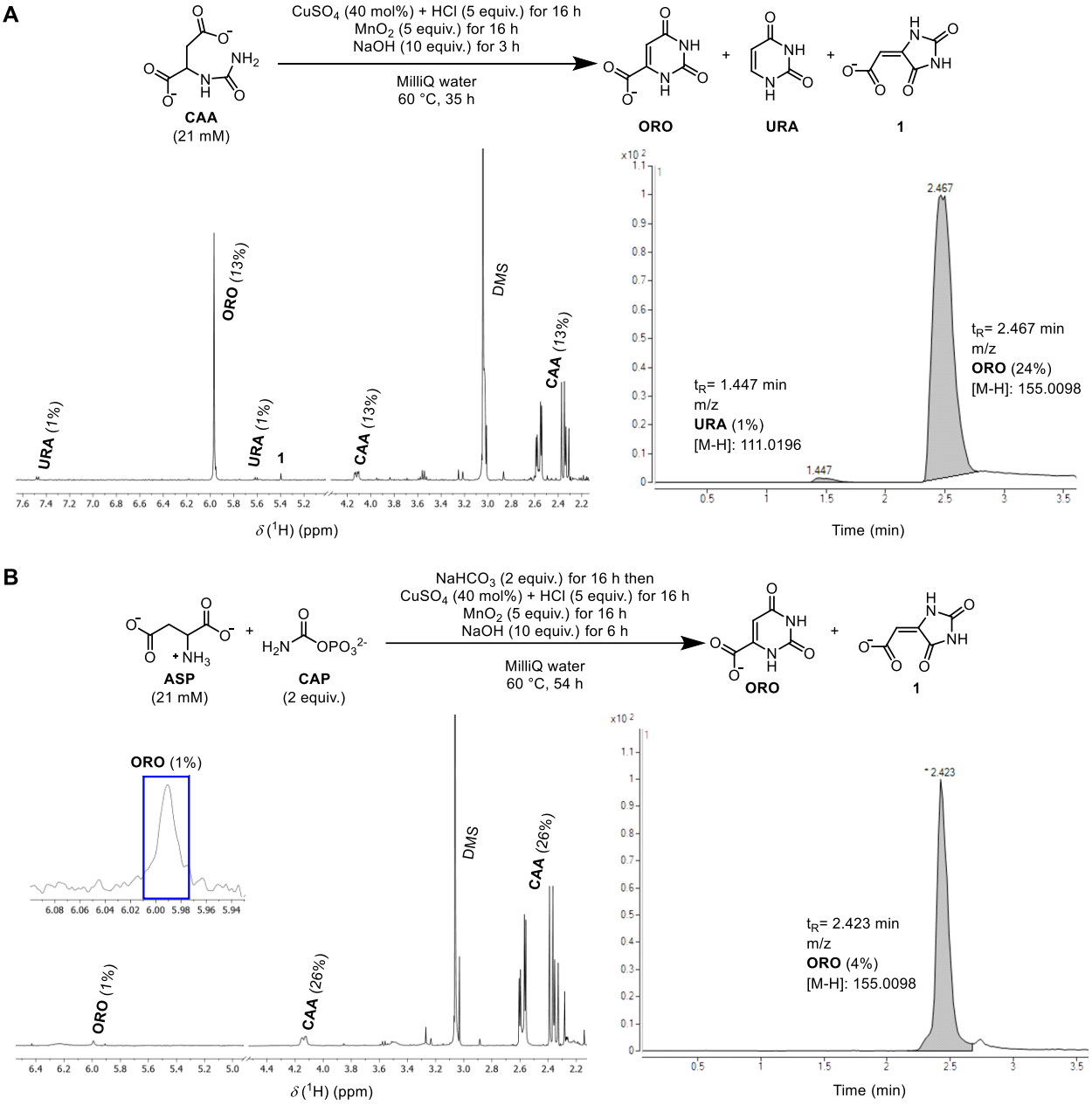


Figure 2. Formation en un seul pot de bases de pyrimidine. (**A**) ^{1}H RMN et LC-QTOF-MS des produits de la réaction multi-étapes de l'aspartate de carbamoyle (**CAA**). (**B**) ^{1}H RMN et LC-QTOF-MS des produits de la réaction en multi-étapes de l'aspartate (**ASP**) et du phosphate de carbamoyle (**CAP**). [M-H] à m/z = 155,0098 a été choisi comme fragment caractéristique pour l'**ORO**, [M-H] à m/z = 111,0196 a été choisi comme fragment caractéristique pour l'**URA**.

3) Conclusion générale

En plus des réactions spécifiques de la voie de biosynthèse présentées ici, des réactions non enzymatiques parallèles qui ne sont pas trouvées en biologie se produisent également mais convergent *in fine* vers un même point biologique : l'orotate. Cette voie parallèle, considérée dans le cadre d'une origine métabolique, appuie l'hypothèse selon laquelle les réactions non enzymatiques ont modelé l'évolution de la voie biologique. Le raffinement évolutif du réseau de réactions non enzymatiques ramifié et redondant illustré dans le schéma 1 (flèches noires et bleues) aboutirait logiquement à une voie linéaire simplifiée comme celle que l'on trouve en biologie aujourd'hui (flèches noires uniquement). Conformément à cette idée, certaines enzymes catalysent toujours un équilibre entre **CAA**, **HAA** et **DHO**, même s'il ne semble pas y avoir d'utilisation biologique pour le **HAA**. Les résultats présents peuvent être considérés comme une première étape expérimentale vers l'incorporation de molécules génétiques dans un cadre métabolique auto-cohérent pour l'origine de la vie. Les perspectives de recherche sur ce projet devraient porter sur l'optimisation de la version non enzymatique en un seul pot de cette voie où les réactions pourraient se produire sans modifications contrôlées des conditions, ainsi que sur l'étude de l'assemblage non enzymatique des ribonucléotides de pyrimidine parallèlement à leur biosynthèse.

PART I

1 Research background

1.1 Abiogenesis

Abiogenesis is defined as the natural development where life has been built up from non-living matter. The transition from inorganics to an organic world involves an evolutionary process with increasing complexity, rather than a single event. This process would have involved different stages: the prebiotic synthesis of simple organic metabolites, molecular self-replication, self-assembly and autocatalysis.[1] Thus, the study of abiogenesis is often focused on the determination of chemical reactions that could give rise to life without the presence of enzymes, under the environmental conditions of the early Earth.

1.1.1 Timeline of multidisciplinary achivements to understand the origins of life

1.1.1.1 Background

Around 300 BC, an ancient Greek natural philosopher, Aristotle, began to pursue the problem of life's origins through detailed and systematic observations. In his theory of spontaneous generation, he insisted that "*living creatures could arise from nonliving matter and without descent from similar organisms*". Centuries later, in the early 1800's, the Swedish chemist, Berzelius, one of the founders of modern chemistry, stated that the synthesis of organic compounds necessitates living organisms, a hypothesis called vitalism. However, both theories were considered disproven after Friedrich Wöhler's urea synthesis in 1828. Wöhler obtained urea by heating inorganic ammonium cyanate, and his result was commenmorated as a milestone in modern chemistry because it showed that biological molecules were not dependent on life but could also be generated abiotically.[2] This idea was supported by Louis Pasteur's 1861 fermentation experiments, which proved that bacteria and fungi did not appear in sterile and nutrient rich medium but arose only from other bacteria and fungi.[3] Following this compelling evidence, Charles Darwin assumed that the origin of life may have begun in a "warm little pond". A primordial soup containing all sorts of inorganic precursors would have given rise to simple biological molecules –formed by purely chemical means and ready to undergo even more complex processes. Darwin's theory of evolution proposed that all living species have descended from a common ancestor through natural selection, which is now widely accepted as a fundamental concept in science.[4]

In the early 1920s, Oparin and Haldane hypothesized that the plausible primitive conditions on Earth were favorable for the evolution of matter, transforming inorganics into organic molecules, which lead to the formation of higher living organisms.[5] Their hypothesis was corroborated 30 years later, once Miller and Urey obtained 20 amino acids from 4 initial inputs including water, methane, ammonia, and hydrogen, using continuous electrical sparks.[6] These incredible result became inspiration for Fox's "proteinoid world". In the late 1950s, Fox discovered that amino acid condensed into polypeptides under simple, high-temperature conditions, and suggested that life-like polymers would self-replicate, leading to the development of higher living organisms. Although the proteinoid world idea was soon dropped by the scientific world, these early chemical experiments provided first insights into understanding how biological reactions could have possibly occured in the early Earth's environment. They soon became the cornerstone of the modern research area investigating the origins of life.

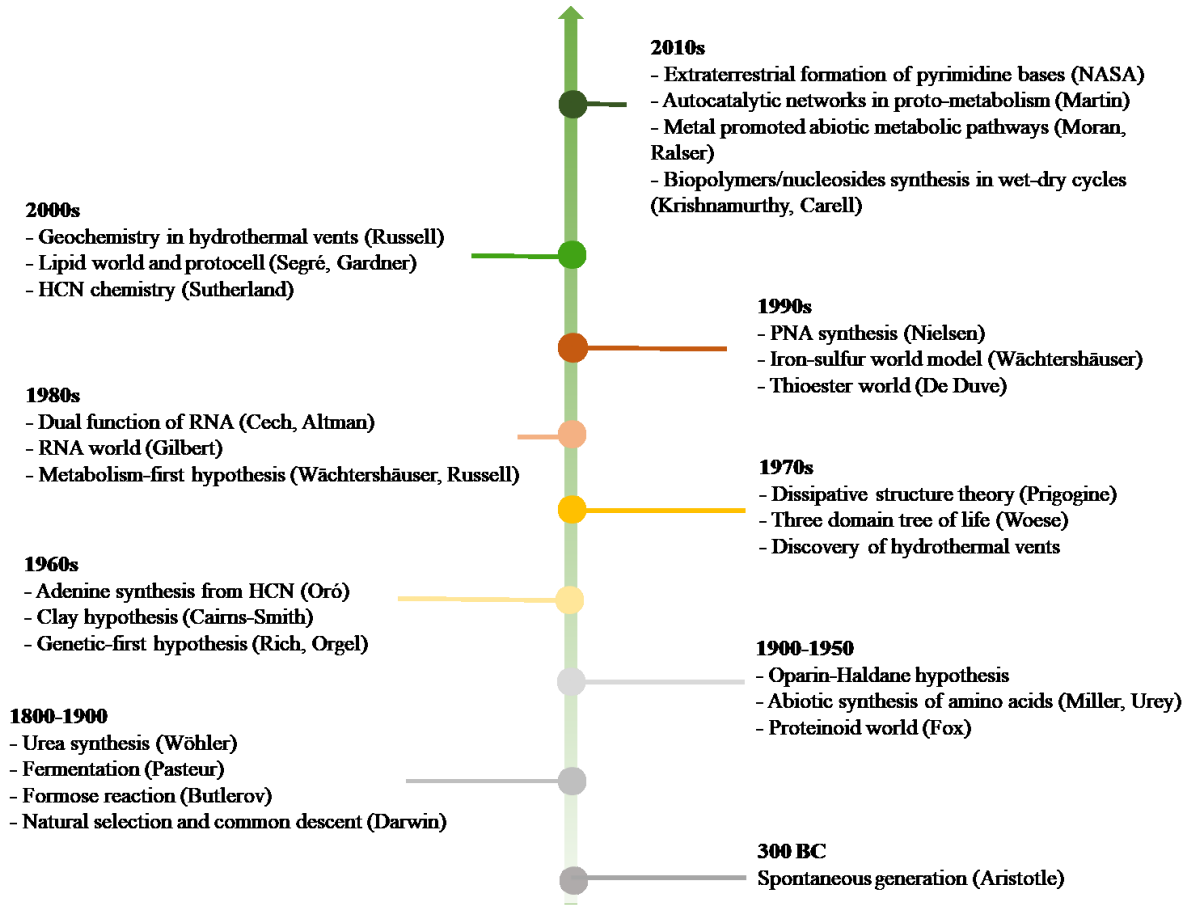


Figure 1: Timeline of multidisciplinary achievements in Origin of Life research

1.1.1.2 RNA world versus Metabolism-First

The very first attempt to investigate a plausible way of synthesizing the building blocks of genetics was by Oró in 1960s, who obtained adenine, one of the nucleobases, by heating HCN in aqueous ammonia for several days at moderate temperature.[7] Later, Cairns-Smith put forward the clay hypothesis and suggested that self-replication of clay minerals could serve as a "vehicle" for the "genetic takeover" while the synthesis of complex proto-organic molecules could be catalyzed by the surface of clay crystals.[8] In 1982, the Nobel Prize winners Cech and Altman disclosed the dual function of RNA: it can serve both as a genetic information carrier and as catalyst in biology.[9] These results supported the "genetics-first" proposal by Rich and Orgel from the 1960s, positing that the emergence of life was mostly reliant on the non-enzymatic synthesis of RNA-based genetic material and ribozymes.[10] Nobel laureate Gilbert later named this concept "RNA world" in 1986 and hypothesised that the self-replicating RNA or pre-RNA molecules existed before the evolution of proteins and DNA.[11] The discovery of clay-directed RNA polymerization by Ferris in 1988 and the prebiotic nucleoside synthesis by Eschenmoser in 1991 gave evidence to support the RNA world hypothesis.[12,13] The peptide nucleic acid synthesis achieved by Nielsen in 1991, as a genetic material which presents high robustness character under extreme conditions, plausibly answer to the critical question which was imposed by the instability of RNA.[14]

In contrast to the previously discussed replicator-first scenarios, some scientists have different opinions on how prebiotic evolution was initiated. In the 1970s, Prigogine in his pioneering work on the dissipative structure theory (Figure 2a) set out to understand how such a thermodynamic open far-from-equilibrium primordial Earth system could have led to the emergence of the first living cell. His mathematical model described how the constant dissipation of energy in chemical systems containing interacting matter, such as CHNOPS sources and hetero- or homogeneous catalysts, could have lead to the emergence of new structures due to internal self-assembly under non-equilibrium conditions.[15] At the same time, Woese brought attention to the genetic code and biological taxonomy. He discovered the third domain: Archaea, which was new and distinct from Bacteria and Eucarya that had been considered as two major divisions of organisms in the structure of the tree of life. His three-domain system resulted not only in the concept of the last universal common ancestor (LUCA) (Figure 2b), but also in the idea that Archaea might have strong evolutionary connections to the

first living organisms on Earth as many of them were found to be highly resistant to extreme environments.[16, 17, 18]

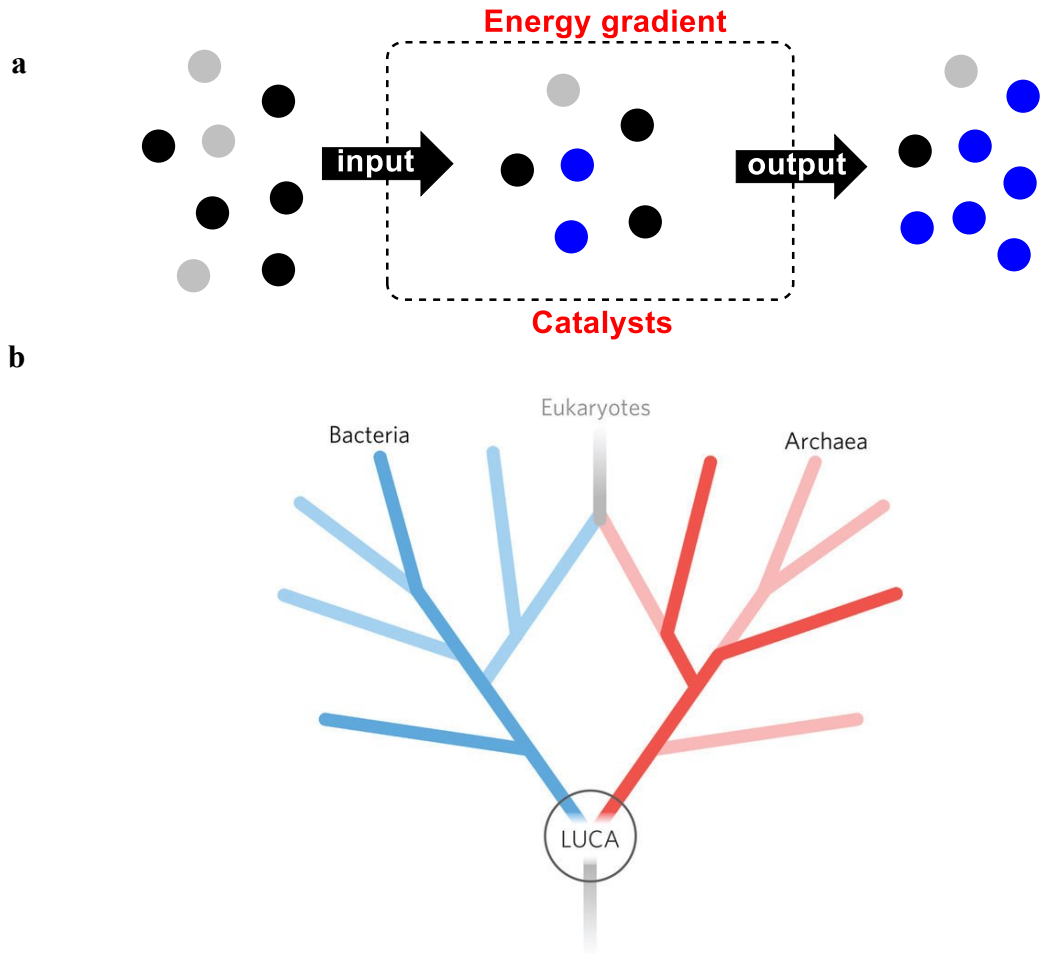


Figure 2. a. Simplified dissipative structures of thermodynamic non-equilibrium system, gray and black circles are initial inputs, blue circles are newly formed products, the dashed square is the environment containing the spontaneous flow of energy and catalysts. *b*. Phylogeny for the Last Universal Common Ancestor's (LUCA) genes (figure reproduced from *Weiss, 2016* [19]).

In the 1970s, the first deep water survey uncovered a hydrothermal vent, namely the "dark smoker", which was reported by scientists from the Scripps Institution of Oceanography with evidence for submarine hydrothermal vents along the East Pacific Rise. These hot ocean vents, which are also acidic and rich in sulfites or sulfur-bearing minerals, became discussed in the origin of life field and was hypothesized to provide highly suitable thermodynamic environment where prebiotic reactions might be allowed on the early Earth (Figure 3).[20] Russell and co-workers later predicted in his water-world theory, that alkaline type hydrothermal vents would also have existed in the early deep-sea environment and would have been the cradle for

life. Such alkaline hydrothermal vents could possibly produce energy for prebiotic reactions by creating proton gradients, due to the proton exchange at the interface between different vents fluid and the electron gradient, due to the metal-rich environment. These two phenomenons would operate as abiotic version of mitochondria that life uses to drive its metabolic reactions.[21, 22] However, this alkaline hydrothermal vent was not discovered until 2001 by Kelley and co-workers.[23] In 1995, several years after receiving his Nobel Prize in Physiology or Medicine, De Duve focused on the thioester world proposal and declared his intention to contest the "RNA world" hypothesis. He insisted that the primordial world must resemble the present-day metabolism by the formation of the same type of activated molecules: thioesters, which act as a simpler version of energy supplier in all living organisms before the phosphate-based energy transduction appeared, must be mandatory to the emergence of other metabolites.[24] Around the same time, Wächtershäuser proposed the iron-sulfur world hypothesis and suggested that early metabolism must have predated genetics. In his hypothesis, organic molecules were produced in reactions occurring in the "dark smoker" area of the hydrothermal vents and were turned into ligands for transition metal catalysts which gave further feedback and feed-forward effects to the growth of complexity on the early Earth. This autotrophic scenario furnished a brand new perspective to biochemical retrodiction and also highlighted the importance of self-organization in biochemical networks.[25] Based on this idea, a lot of work has been done in both the geochemical and chemical fields, showcasing reactions that were able to mimic the metabolic reactions of making essential building blocks for life under alkaline hydrothermal-like conditions.[26] These results corroborated the metabolism-first approach and put it in an extremely challenging position against the traditional replicator-first hypotheses.

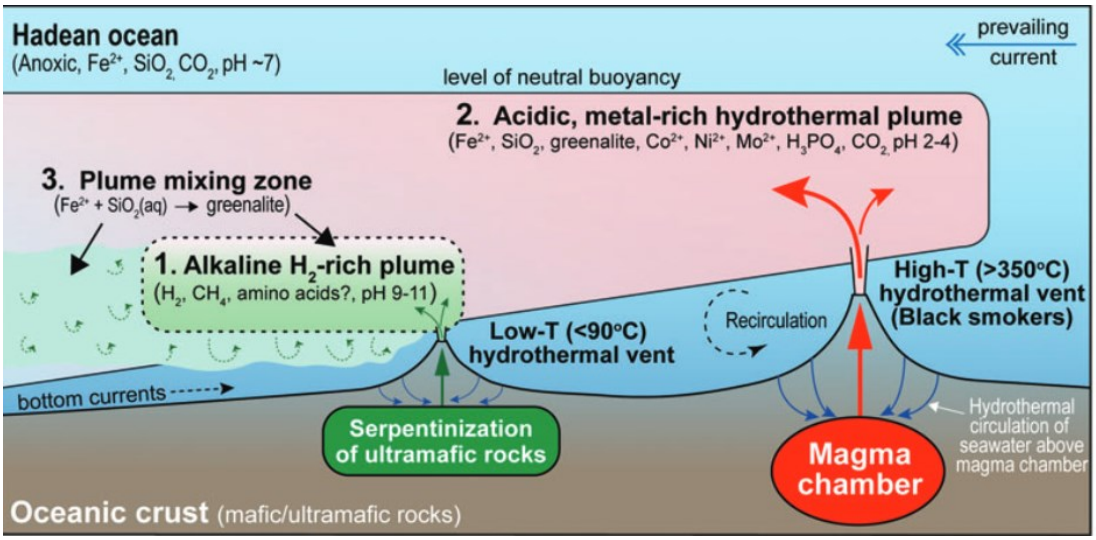


Figure 3. Representative scheme of acidic (black smoker) and alkaline deap sea hydrothermal vent in the Hadean ocean. (figure reproduced from *Rasmussen 2021* [27])

The contradictory “chicken or the egg” problem (whether life emerged from genetics or from metabolism) is still debated, as multiple prebiotic scenarios have been reviewed and examined both theoretically and experimentally. The main critiques between these two principal hypotheses could be outlined as the following questions: (1) If the evolution is relevant to the contemporary biochemistry? (2) Is there any heredity? (3) What was the energy source? (4) What were the initial chemical components? (5) Were the systems selective towards the evolutionary path? (6) Were these pre-biochemical reactions robust? (7) How is the origin of homochirality addressed? (Table 1)

Table 1. Strengths and weaknesses of genetics-first and metabolism-first hypotheses

Feature	Genetics-First	Metabolism-First
Continuity in chemistry	Lacking	Yes
Heredity	Yes	Lacking
Energy source	Solar	Hydrothermal vents, Solar, Lightning
Initial chemical substrate	HCN or related cyanides	Simple geo-chemicals (CHNOPS and minerals)
Selective chemistry	Yes	Lacking
Robust chemistry	Lacking	Questionable
Chiral questiom	Yes	Lacking

The major reason why our lab supports the metabolism-first paradigm is because of the high degree of similarity it expresses between proto-metabolism and modern biochemistry.[28] This principle of continuity[29], leaves a significant mark in origins of life studies and aims to decipher the historical pre-Darwinian evolution. It offers a strong heuristic tool to track down the ancient paths or types of reactions that early Earth employed to build up such rigid and highly organized reaction networks based on the richly detailed knowledge of current science. Instead of enzymatic reactions in biochemistry, proto-metabolic networks could be established with a

simpler version of geologically available catalyst, such as minerals, and could anticipate the formation of essential building blocks that were rational for the metabolic evolution. In contrast to this prebiotic model, the "RNA world" studies are sometimes very different from the biochemical organization of a comtemporary living system. For example, the Strecker amino acid synthesis and the cyanide chemistry-based nucleotide synthesis do not resemble the known biological metabolic networks.[30, 31]

Ubiquity is another reason to refute the "RNA world" hypothesis.[32] Unlike the proto-metabolic reactions which were able to take place within widespread environments, cyanide chemistry was found to be incapable of being initiated under ancient geological conditions. Several laboratory results have demonstrated that the multi-step reactions towards nucleosides[31, 33, 34] or polypeptides[35] require very specific reaction conditions for each step and even sometimes purifications which was, under prebiotic environment, unachievable without compartimentalization or human intervention. Moreover, the prerequisite for a constant supply of HCN or cyanoacetylene, which were not abundant in the primitive terrestrial atmosphere, remains controversial in this hypothesis.[36]

Even though RNA polymers are considered to be genetic carriers as well as catalysts to their own replication in contemporary biology, it is never being reported experimentally in a prebiotic system that they can do both jobs at the same time.[37, 38] Therefore, the question of how does the first self-replicating RNA emerge without template-directed information replication became a major problem.[39] The formation of such complex molecules or their precursors are generally low yielding and present high kinetic barriers in the absence of activating agents or catalysts.[40] On the other hand, in the metabolism-first hypothesis such reactions are promoted by Earth abundant minerals which consequently enhance the reaction kinetics.[26, 41] Even though preliminary metabolic reactions did not provide the intermediates which could play the role as RNA replicators, predictions from computational modelling combined with experiments indicate that the metabolic networks display at least one obligatory autocatalytic core. A major example is the reverse Krebs cycle, which is conceptually parallel to the enzyme-catalysed biological reverse Krebs cycle.[42] Thus, the major argument in the origin of life came up with the ability of running such autocatalytic network independently without enzymes, but with continuous input of one of the metabolites of the network. Muchowska *et al.* recently provided promising experimental evidence supporting this concept, but also demonstrated that parasitic reactions occurred in the metal-catalyzed reactions. In fact, chemical selectivity poses a big

problem in metal-promoted prebiotic reaction networks, which is not the case in enzymatic catalysis.[43]

Since the origin of hereditary replicators is still debatable, it is advocated to dissect the large polymers (mass > 5000 Da) in order to track back the essential building blocks that may have existed in early Earth environments.[29,44] The need of combining this top-down (analytic) approach and bottom-up (synthetic) approach, which involves the *de novo* synthesis and assembly of these building blocks that attempt to construct the structural and functional units of such polymers, is urged for the progress of origin of life field. Following this purpose, genetic-first approach cannot be regarded as satisfactory due of the imperative requirement for structurally complex large polymers (RNAs with mass > 5000 Da). On the other hand, metabolism-first hypothesis, just like the modern metabolic networks, bypasses this obligation and can presumably uptake small geological molecules (mass < 500 Da) which are achiral and widely distributed on the early Earth. This last prebiotic scenario was likely to be adopted by nature and shows advantageous robustness that accelerates the accumulation of organic molecules.

Following the above discussion, the metabolism-first approach appears as more convincing to me. In order to illustrate this specific hypothesis, I will describe in more detail the key concepts of the metabolism-first hypothesis, giving examples of different prebiotic scenarios.

1.2 Metabolic origins of life

1.2.1 Metabolic pathway evolution

Several models have been proposed to explain the possible mechanism of the emergence of metabolic pathways. The initial definition of metabolic networks was given by Horowitz in 1945: it is a sequence of consecutive reactions where each of them is catalyzed by an individual enzyme.[45] He firstly addressed the retrograde pathway evolution model and assumed that the hereditary organisms capable of Darwinian evolution predated the emergence of metabolic pathways. In his hypothesis, the depletion of functional metabolite stimulates selectively and positively the abiotic formation of an enzyme that catalyzes its formation from any available precursor already existing in the primitive environment. The subsequent depletion of this functional metabolite, together with its precursor, would prompt the emergence of new pathways in the reverse direction to its synthesis (Figure 4a). However, the reason why abiotic reactions led to the formation of these precursors, metabolites or enzymes and the way how it

happened remains unclear in this theory. Therefore, we speculate this retrograde model would have occurred at a more complex level of biological evolution, much later than the prebiotic stage.

An alternative model of forward pathway evolution was carried out by Granick to compensate the previously shown contradictions.[46] He mentioned that "*Each step had to be an end product of the biosynthetic chain and serve a useful function. Later, an additional step would be elaborated to produce a new end product that could carry out the same function more efficiently, but the previous step would still be essential, for it would serve as an intermediate*". The major point of his hypothesis is that the evolution of metabolism generates not only new enzymes but also new metabolites through enzymatic reactions (Figure 4b). Nonetheless, the absence of nonenzymatic reactions made this model inapplicable in the early prebiotic context.

The observation in divergent metabolic pathways that ancestral enzymes possessing various functionalities could yield different end products from identical essential substrates, inspired Ycas and Jensen to propose the generalists-to-specialists model (Figure 4c).[47, 48] However, these consecutive reactions not only lead to divergence, but would also drive to convergence as similar reactions could have emerged independently due to chemical necessity. Thus, phylogenetic evidence is indispensable to support the divergent paradigm. A complementary scenario by Lazcano and Miller called patch assembly, formalized that the simultaneous recruitment of multiple enzymes from different pathways could give rise to brand new ones (Figure 4d).[49] The current topology of the metabolic networks is most likely adopted from the patchwork model because the greater part of metabolic pathways uses enzymes of individual evolutionary origins. Nevertheless, this patchwork model does not explain how pathways may have initially emerged under primitive conditions.

To account for the shortcomings of the models reviewed above, Tawfik proposed an integrated metabolite-enzyme coevolution model that unifies different ideas from the past and could also be applied in the prebiotic context.[50] This model gave prominence to the nonenzymatic transformations and claimed their important role to help new enzymatic pathways emerge at the origin of life. In this hypothesis, side products from either enzymatic reactions or nonenzymatic reactions, namely underground metabolism, served as precursors or stimulators to the origins of specialized enzymes that catalyzed the formation of these products (Figure 4e).[51] Consequently,the spontaneous emergence of new enzymes is no longer necessary as metabolic reactions could have been enzyme-free, in contrast to the previously described models. Inspired by this hypothesis, Moran suggested that "*in the context of the earliest steps*

in the origin of life, we need to consider proto-metabolic pathways that predated enzymatic biochemistry, and which would, by definition, have been entirely nonenzymatic" (Figure 4f).[26] This protometabolic scenario implies that small molecule catalysts such as metal ions, minerals or organocatalysts could assist and accelerate nonenzymatic reactions.

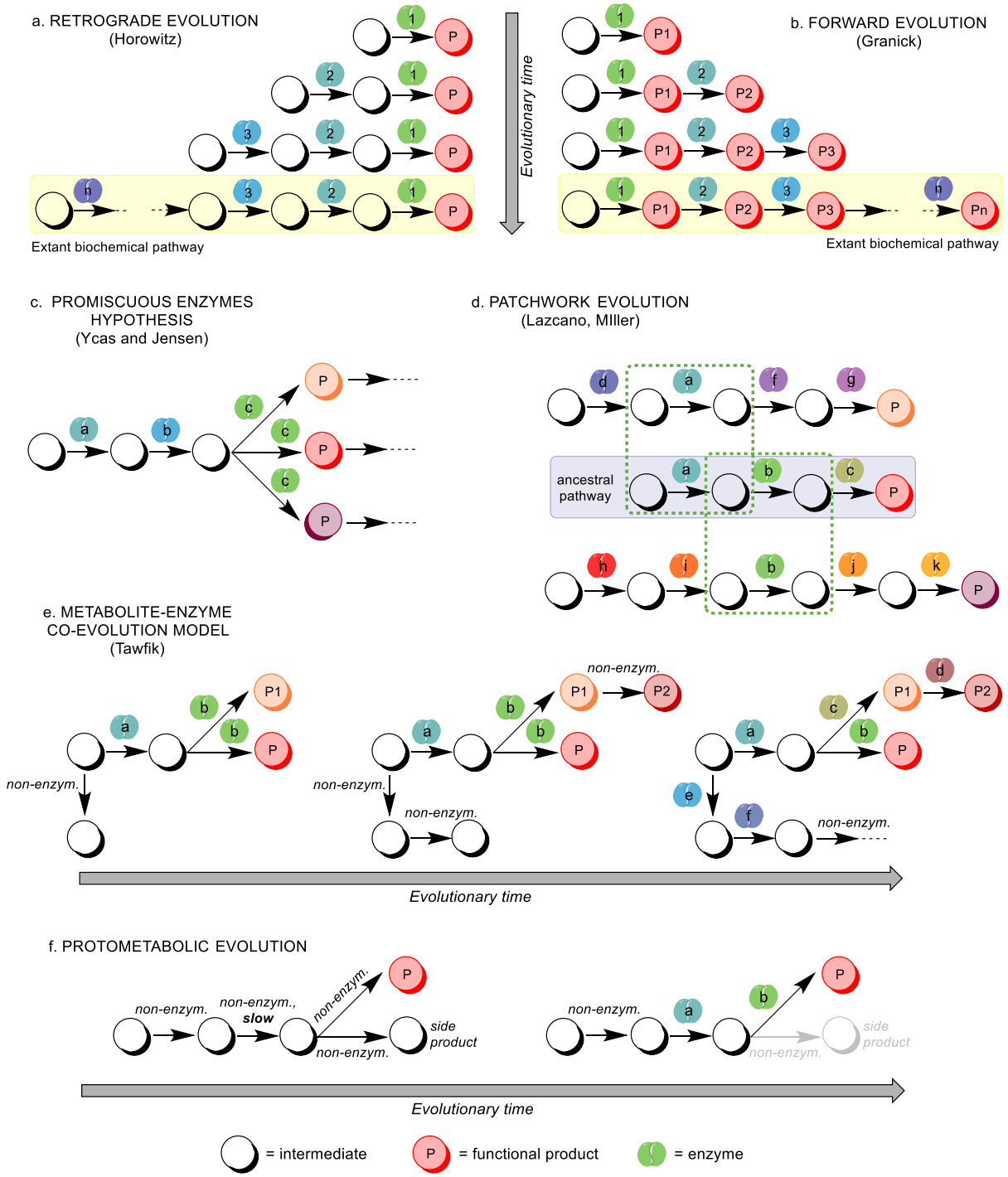


Figure 4. An overview of different models for metabolic pathway evolution. Different colors correspond to different enzymes or end products (figure reproduced from *Muchowska, 2019* [26]).

1.2.2 Ecological and energetic perspectives on metabolism

Taking a deep look at modern biochemistry, living organisms are usually devided into two individual variants: autotrophic organisms that are self-sustainable by producing their own metabolites from inorganic sources (CHNOPS elements) and energy, such as light (photoautotrophs) and chemical energy (chemoautotrophs), in the environment; whereas heterotrophic organisms are dependant on other organisms due to the inability of completing this self-sustained process by themselves (Figure 5).[52]

Another key factor to distinguish the organisms is to analyze the type of metabolic reactions used to provide energy supply for the system to survive. Mammals employ catabolic metabolism to generate energy from food consumption. The mechanism goes through oxidative processes that break down C-C bonds of metabolites and release CO_2. Thus, electron acceptors such as O_2, SO_4^{2-}, NO_3^- etc are required in these organisms. The second type of organisms, using anabolic metabolism, reductively builds up the carbon skelton of organic molecules by uptaking CO_2 as a carbon source, and uses H_2, H_2S, or other sources as electron donors. This type of mechanism usually occurs in many of the bacteria and archaea.[53]

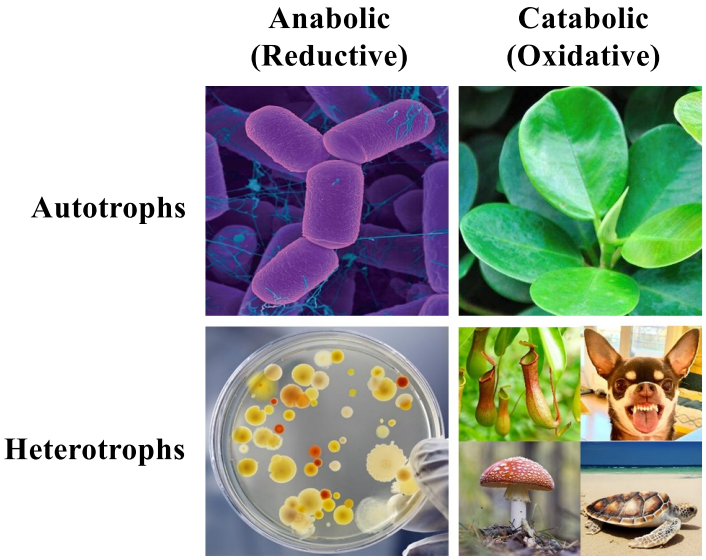


Figure 5. The classification of metabolisms and organisms.

On an ecosystem level, all these organisms that use two types of metabolism and two types of food source contribute to the global biological carbon cycle, where large biomolecules are made from CO_2 and then broken down back to CO_2 on the global scale. Given the biological and geological evidence[54], it is often considered that the prebiotic scenario was based on anabolic chemoautotrophic processes[55], instead of photosynthesis, which was thought to be a later

development, and that life recruited inorganic electron donors to fix CO_2 and to promote chemosynthesis.[56]

1.2.3 Five universal metabolic precursors in biology

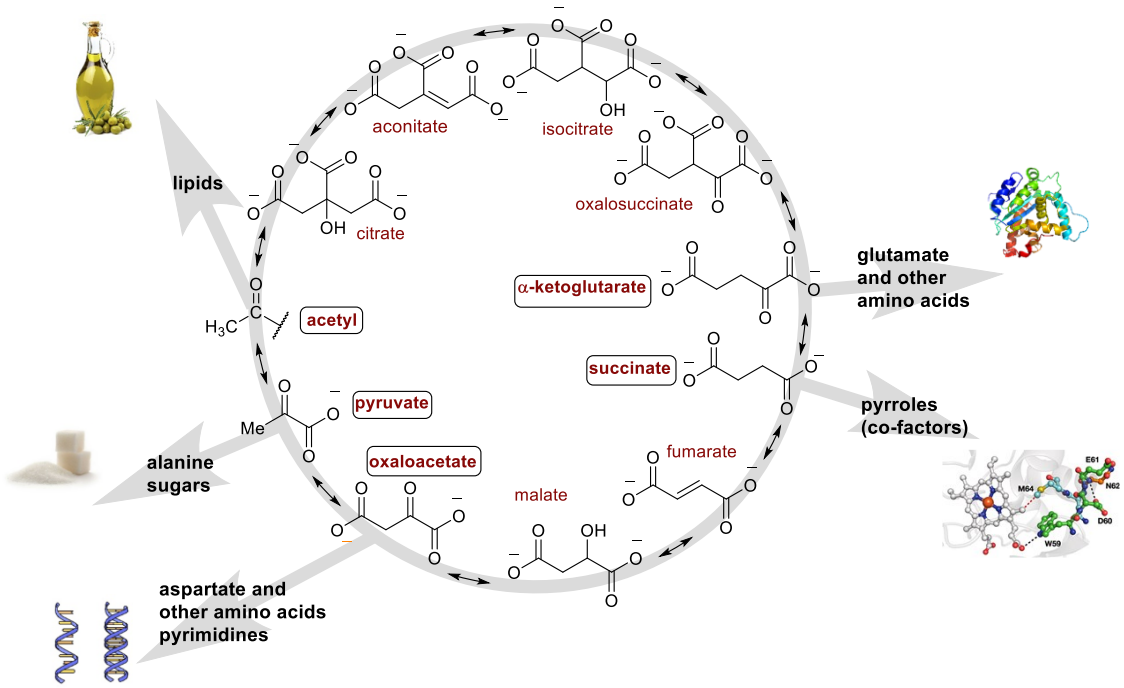


Figure 6. The five universal metabolic precursors and their location in the (r)TCA cycle.

From the above discussion, we may notice that modern metabolic networks possibly evolved from prebiotic pathways that fixed CO_2.[57] Smith and Morowitz identified five universal metabolic precursors in current biochemistry that could be regarded as the first set of metabolites that were delievered from nonenzymatic reactions of CO_2 under primitive conditions.[52] These universal metabolic precursors are: (1) acetyl, the biosynthetic precursor to fatty acids, (2) pyruvate, the precursor to certain amino acids and sugars, (3) oxaloacetate, the precursor to various amino acids and pyrimidines, (4) succinate, the precursor to cofactors, and (5) α-ketoglutarate, the precursor to other amino acids. Recent phylometabolic analyses of biological systems indicate the location of these compounds in known metabolic reactions: predominantly in the tricarboxylic acid cycle (also called the TCA cycle, Krebs cycle or citric acid cycle).[52] The TCA cycle includes 11 small carboxylic acids that are interconvertible. It also involves biochemical oxidations that break down these carboxylic acids into CO_2 and release energy in the form of ATP and thioesters (Figure 6). Unlike in most of the organisms

nowadays, this cycle is also found to operate in the reverse direction (reductive TCA or rTCA cycle). Some organisms like archaea and bacteria, whose genes show little homology across all domains of life that are thought to have derived directly from LUCA, are very well known autotrophic systems making their metabolites from CO_2 using the rTCA cylcle.[58]

1.2.4 Reverse-TCA cycle and its metabolic primer

The rTCA cycle plays a key role in biochemistry as its sequential reactions continuously produce the five universal metabolic precursors. The topology of this core metabolic cycle is autocatalytic. Each full turn of the cycle allows for an exponential self-amplification of the whole network, meaning the quantity of intermediates will double every time.[52] Nevertheless, running this cycle as a perpetual motion machine becomes problematic, especially when enzymatic reactions had not yet emerged on the early Earth. A classical autocatalytic network (Figure 7, left) would collapse if the overall yield became <50% and parasitic reactions would become predominant.[59] A stabilized autocatalytic system was thus proposed. This model introduced a primary feeder pathway that supplies constantly the fuel molecule (Figure 7, right). In the context of the rTCA cycle this molecular fuel refers to acetate, as the which could replenish the low-yielding network to survive even in the absence of enzymatic catalysis.[57] This type of network organisation, if present in prebiotic chemistry, could have led to the emergence of more complex biological polymers.

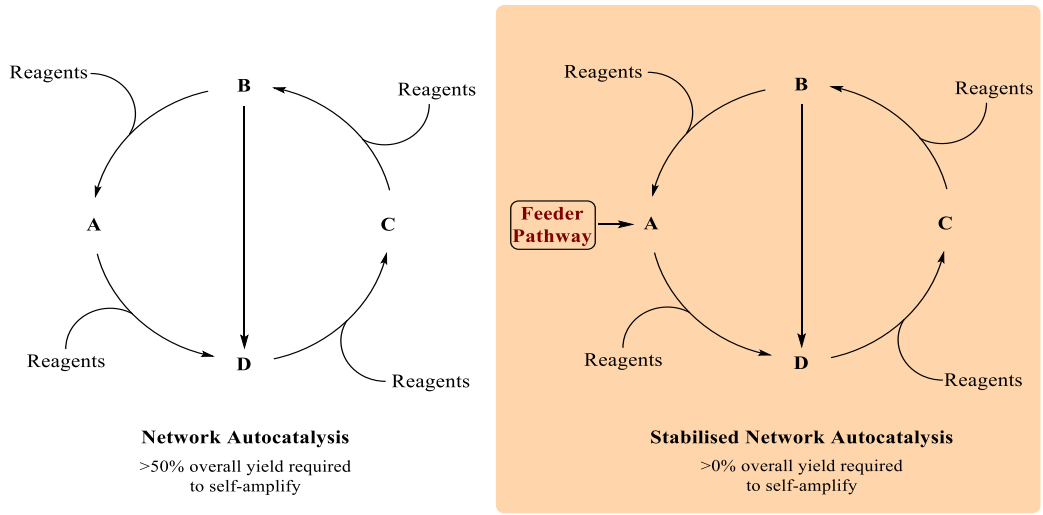


Figure 7. Autocatalytic models with and without feeder pathway.

In order to maintain a sustained supply of acetate to the network, this two-carbon molecule had to be synthesised prebiotically from environmentally available CO_2 through CO_2 fixation pathways.[60] There are seven such pathways known in biology: (1) Calvin cycle, (2) Wood-Ljungdahl (W-L) pathway (namely acetyl-CoA pathway), (3) rTCA cycle, (4) 3-hydroxypropionate bicycle, (5) dicarboxylate-hydroxybutyrate cycle, (6) 3-hydroxypropionate/4-hydroxybutyrate cycle and (7) the reductive glycine pathway. Autotrophic organisms use them to build up their organic molecules. Amongst these pathways, Calvin cycle is used by photoautotrophs, which, according to our previous discussion in chapter 1.2.2, most likely evolved after chemoautotrophs.[61] From a phylogenetic perspective, the anabolic generation of acetate through the Wood-Ljungdahl pathway was also found in organisms that possess a small number of genes which are common to all domains of life. Just as we discussed previously, these organisms were likely derived directly from LUCA. They feature Fe-Ni-S clusters, the second most abundant cofactor in biochemistry after ATP, which are engaged in the carbon fixation pathways. These figures made the W-L pathway the simplest and oldest CO_2 fixation pathway, and also one that could possibly fit in the prebiotic scenario as feeder pathway to sustain the autocatalytic rTCA cycle (Figure 8).[52]

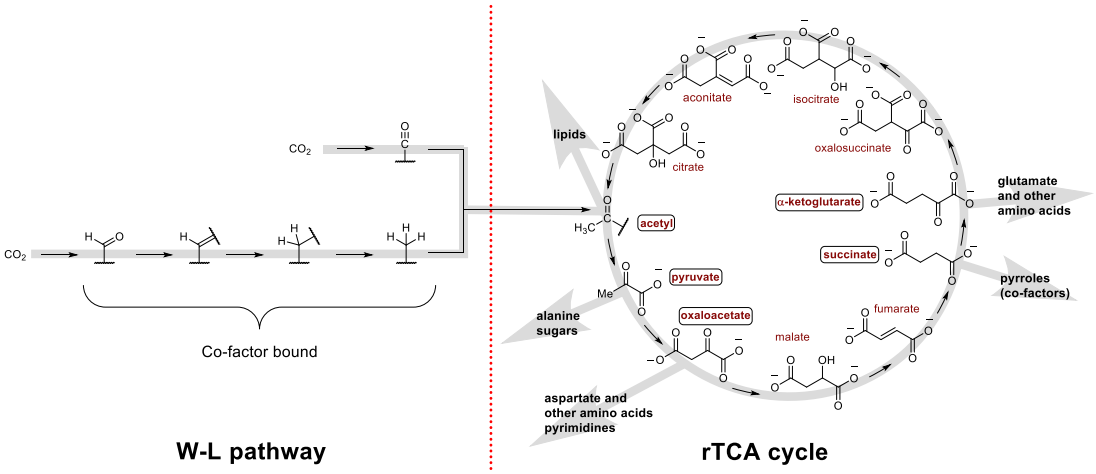


Figure 8. The autocatalytic rTCA cycle stabilized by the W-L pathway.

The above picture indicates that the W-L pathway involves two short and linear sequences of CO_2 reduction, the homologous products subsequently coalesce to form acetate which reacts subsequently with another CO_2 molecule to afford pyruvate. These two carboxylic acids are considered two of the five universal metabolic precursors. More detailed information about this

pathway will be described in the next chapter. The remaining five carbon fixation pathways share the autocatalytic property and also provide the five universal precursors just as the rTCA cycle does, but their reaction network is way more complicated than the W-L path and the rTCA cycle.

The next chapter will focus on the description of central carbon fixation pathways and relate them to the prebiotic scenario. Next, an extension of these metabolic pathways towards amino acids, sugars and pyrimidine nucleotides will also be mentioned, with the goal to show how prebiotic chemistry could have given rise to genetic precursors from CO_2.

2 The organization of life's metabolism: from CO_2 to RNA

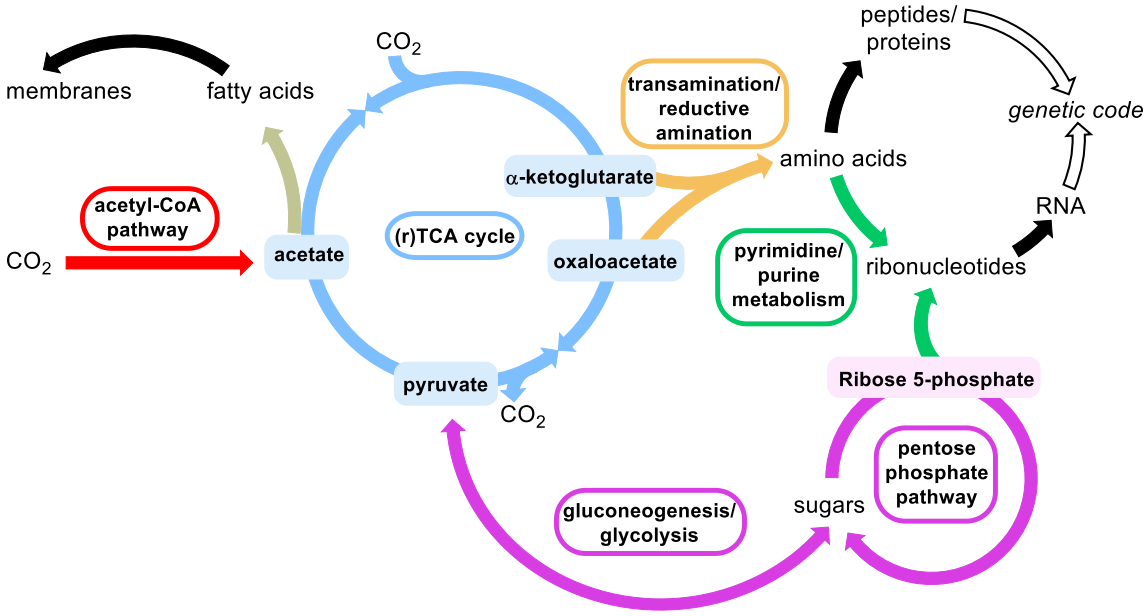


Figure 9. The summarized metabolic pathways that convert CO_2 to genetics. (Figure adapted from *Muchowska, 2019* [26])

Modern biochemistry is constrained by the laws of chemistry and physics. It combines several sub-systems of metabolic pathways that are highly interconnected.[62] The simultaneous cooperation of these pathways allows life to generate its necessary metabolites and to obtain the energy that sustains its function. From a prebiotic point of view, the emergence of such complex system could be considered in a more condensed way, that is, by studying the nonenzymatic chemistry of converting inorganics (sources of CHNOPS) into the molecules of

genetics. This evolution includes multiple intermediate steps that core metabolic reactions underwent while operating all at the same time. The building blocks that assembled into genetic molecules resulted from the five universal precursors, within several metabolic pathways such as sugar metabolism (to obtain ribose 5-phosphate) and amino acid biosynthesis (Figure 9). These core metabolic pathways will be described in the following paragraphs.

2.1 The acetyl-CoA pathway

2.1.1 Intoduction

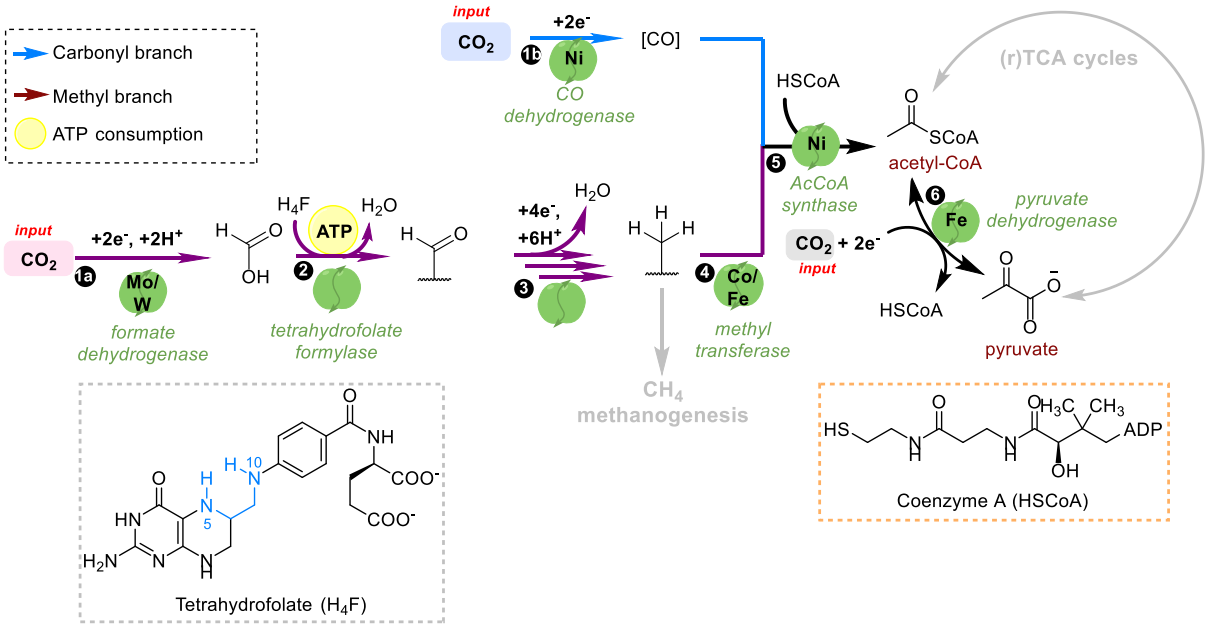


Figure 10. The metalloenzymes and cofactors of the acetyl-CoA pathway.

Among the seven autotrophic carbon fixation pathways in biochemistry, the acetyl-CoA pathway is considered to be the shortest and most dependant on transition metals, as well as the only linear pathway.[63, 64] Phylogenetic analyses have revealed that this pathway is mostly used by acetogens and methanogens, anaerobic organisms found at the roots of the tree of life.[65] Acetogenic organisms take both parallel reductive reaction pathways (methyl synthesis in Step I & carbonyl synthesis in Step III, Figure 10) and combine them to build up acetate and pyruvate from CO_2 (Step IV.[66] This process creates an ion gradient that facilitates adenosine 5'-triphosphate (ATP) synthesis. This is the most commonly used type of energy metabolism in biology.[67] Note that only the methyl synthesis reaction (Step I) is found in methanogenic organisms and is related to the methanogens-specific type of energy metabolism. For all these

reasons, the acetyl-CoA pathway is thought to be the most ancient CO_2 fixation pathway and is assumed to have its origins in prebiotic chemistry.[68]

The reaction pathway initially proceeds though two parallel mechanisms. The first one involves a reversible two-electron reduction of CO_2 to formyl group with the aid of multiple metalloenzymes containing Mo, W or Fe, and cofactors such as tetrahydrofolate (H_4F).[69] The presence of metals in enzymes is another reason why the W-L path is considered to have prebiotic origins. In acetogens, the second type of reduction proceeds in parallel, leading to carbonyl formation, and is promoted by a Ni-based metalloenzyme.[70] A possible electron bifurcation is surmised to enable these reductions which present a high energy barrier. The electrons from dihydrogen are stored in highly reduced Fe-based cofactors.[71] Afterwards, the Co-based enzyme transfers the formyl group from its cofactor to the active site of a second Ni-based enzyme (ACS) where it meets the carbonyl group. The association of these two groups yields an acetyl-Ni complex, which is subsequently trapped by coenzyme A (HSCoA). The formed acetyl-CoA thioester mostly undergoes enzymatic reductive carboxylation to yield pyruvate, but also acts as precursor for fatty acids synthesis.

2.1.2 Nonenzymatic versions of the acetyl-CoA pathway

Recently, several prebiotic works achieved the nonenzymatic chemistry of the W-L pathway. Many of them selected CO_2 or HCO_3^- as C1 compounds, or even more reduced versions such as CO and HCOOH.[72] External reductants are crucial in the first case which are able to act as electron donor, as well as catalyst to bypass the kinetic barrier thus improve the electron bifurcation like the metalloenzymes/cofactors.

The first nonenzymatic analogue of the acetyl-CoA pathway was reported in 1997 by Huber and Wachtershäuser using methyl thiol and CO at 100 °C, in which acetate was obtained after 7 days.[73] Their mechanistic studies indicated that thioacid or thioester were probably formed as activate intermediates (Table 2, entry 1). Similar results which applied harsh condtions (high pressure or concentration) were reported later but thought to be geologically unrealistic (entry 2 and 3).[74, 75]

In 2010, Feng *et al.* reported the first use of Fe-based nanoparticles to mimic the pathway under moderate conditions.[76] They ended up with very little amount of acetate and formate (entry 4). The advancement of electrochemistry in the following years also encouraged the prebiotic research of of CO_2 fixation, which lead to the efficient production of acetate (entry 5)[77] and

even pyruvate (entry 6)[78] under much milder reaction conditions compared to the previous results.

In the late 2010s, Moran and co-workers discovered that the use of zero valent metals (Fe, Ni or Co) resulted in the formation of acetate and pyruvate at 100 °C with moderate CO_2 pressure in aqueous solution, whereas Mn, Wo and W gave acetate only (entry 7)[79] These metals are highly reminiscent of the metalloenzymes/cofactors that sustain the biological acetyl-CoA pathway. A later study by Martin, Moran and Tüysüz incorporated H_2 as additional reductant revealed improved yields of acetate and pyruvate, the use of hydrothermal minerals (awaruite, magnetite and greigite) turned this result likely realistic under prebiotic conditions (entry 8).[80]

Table 2. The summarized results of nonenzymatic analogues of the acetyl-CoA pathway[a] (adapted from *Muchowska, 2019* [26])

Entry	Carbon source	Reducing agent /catalyst	Reaction conditions	Product Yields				Ref.
				Formate	Methanol	Acetate	Pyruvate	
1	CO (1 bar), CH_3SH (8 mM)	FeS-NiS (1:1)	pH 8, 100 °C, 7 d	-	-	3.28 mM (41%)[b]	-	73
2	HCOOH (110 mmol), thiol	FeS	500-2000 bar, 250 °C, 6 h	-	-	5.5 x 10^{-5} mmol (0.05%)	7.7 x 10^{-5} mmol (0.07%)	74
3	CO_2 (22.5 mmol)	FeNi (98.75:1.25)	300 °C, 6 h	1.4 mM (2.3%)	-	-	-	75
4	CO_2 (14 bar), H_2O	Fe^0 (5 mmol)	200 °C, 72 h	8.5 mM (0.0085%)[c]	-	3.5 mM (0.0035%)[c]	-	76
5	$NaHCO_3$ (0.5 M aq)	N-doped nanodiamond (3.68%) electrode, −0.55 to −1.30 V	rt, 1 h	1.2 mM (0.24% / 0.01%) [d]	-	16.1 mM (3.2% / 0.07%) [d]	-	77
6	CO_2 (bubbling)	Greigite electrode (−1.1 V)	pH 6.5, 120 h, rt	1.3 µmol (1.5%)[d]	0.35 µmol (1.2%)[d]	0.57 µmol (2.6%)[d]	0.48 µmol (2.8%)[d]	78
7	CO_2 (1-35 bar)	Fe^0	H_2O, 30-100 °C, 16 h	0.41 mM (0.014%)	0.12 mM (0.026%)	0.18 mM (0.054%)	0.03 mM (0.012%)	79
8	CO_2 (15-25 bar)	Ni_3Fe (also Fe_3O_4, or Fe_3S_4), H_2	H_2O, pH 4-10, 100 °C, 16 h	332 mM	0.12 mM	0.56 mM	10 µM	80

[a] Percentages in brackets correspond to yields calculated with respect to the carbon donor as the limiting reagent. [b] Thiol as the limiting reagent. [c] Metal as the limiting reagent. [d] Faradaic efficiency.

2.2 Reductive tricarboxylic acid cycles

2.2.1 Introduction

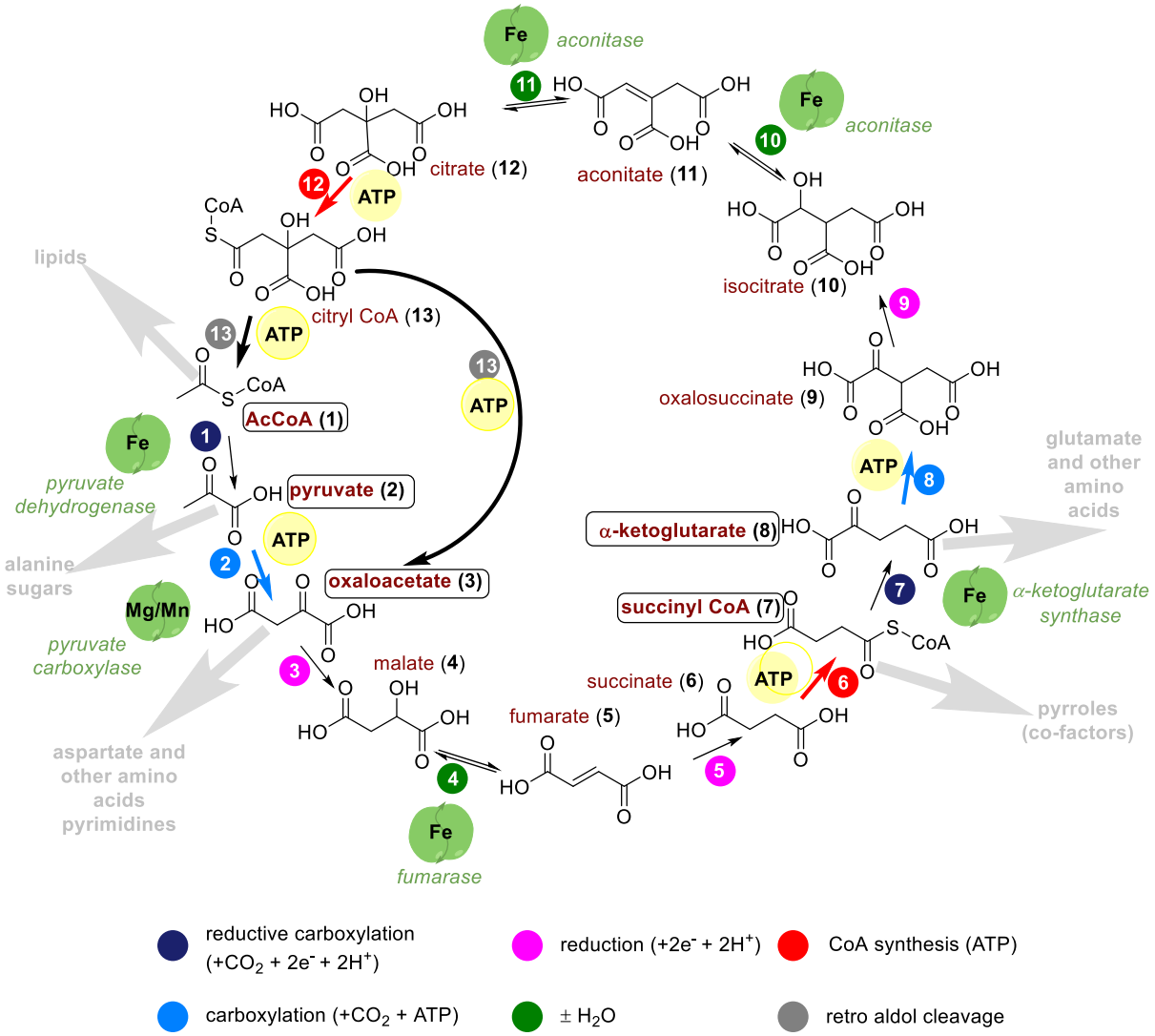


Figure 11. The reverse Tricarboxylic acid (rTCA) cycle. ATP-dependant reactions are shown in bold arrows. The five universal precursors are highlighted in boxes. (figure adapted from *Muchowska, 2017* [26]).

The rTCA cycle is also found to be used by some anaerobic autotrophs, and has been suggested as evolutionary ancestor of modern biochemistry due to its unique net anabolic function that provides constant supply of the five universal metabolic precursors, which is unpratical for other three CO_2 fixation pathways (see chapter 1.2.4).[81] However, the full rTCA cycle is found only in bacteria and not in archaea, which made this hypothesis debatable.[22] This highly autocatalytic cycle includes in total 10 carboxylic acids and 3 thioesters. These intermediates are interconvertible between them through six different types of reactions which are (1) two

reductive carboxylation reactions that form C-C bond through two-electron reduction of CO_2 (step 1, 7 in Figure 11), (2) two carboxylations via carboxyphosphates where C-C bond formation is ATP dependent (step 2, 8), (3) three reductions (step 3, 5, 9), (4) three reversible hydrations and dehydrations (step 4, 10, 11), (5) two thioesterifications which are also ATP dependent (STEP 6, 12), and (6) one ATP-dependent retro-aldol reaction that cleaves C-C bond of citryl-CoA (step 13). Enzymes mediating the rTCA cycle are mainly metal-based. Notably, enzymes that use Fe as cofactors promote reductive carboxylations and the reversible hydrations/dehydrations. This could lead to a hypothesis how naturally occurring mineral catalysts could enable the prebiotic versions of these reactions in the rTCA cycle.

Despite the fact that experimental studies to date have not reported the full rTCA cycle in the context of prebiotic chemistry, one reason is because the undesired parasitic reactions leading to off-cycle products occurred faster than the cycle itself to amplify.[82] A horseshoe rTCA sequence (from pyruvate to α-ketoglutarate, step 2 to 7) is therefore suggested to be more relevant to the prebiotic context as it is simpler and linear.[22] This alternative reaction pathway still contains all the five universal intermediates in its sequence and is proposed to be potentially boosted by mineral catalysts under primitive conditions.[82, 83, 84]

2.2.2 Prebiotic analogue of the rTCA cycle

Previously, single reactions such as nonenzymatic hydrations or dehydrations were reported on similar substrates as those of the rTCA cycle.[85, 86, 87] However, the first systematic studies on mimicking the primitve rTCA cycle published by Cody et al. only came up in 2001.[88] Their results highlighted the nonenzymatic decomposition of citrate via the retro-aldol reaction using FeS/NiS as catalyst at 200 °C. However, the competing parasitic reactions were not negligible under such extreme conditions. The following research efforts mainly focused on the reductions in the horseshoe sequence, namely the reduction of oxaloacetate to malate (step 3) and fumarate to succinate (step 5). Various conditions were reported for these reactions, such as the photochemical reduction using ZnS by Zhang and Martin,[89] hydrothermal reduction using FeS and H_2S by Wang *et al.*[90] and electrochemical reduction using FeS by Kitadai *et al.*[91]

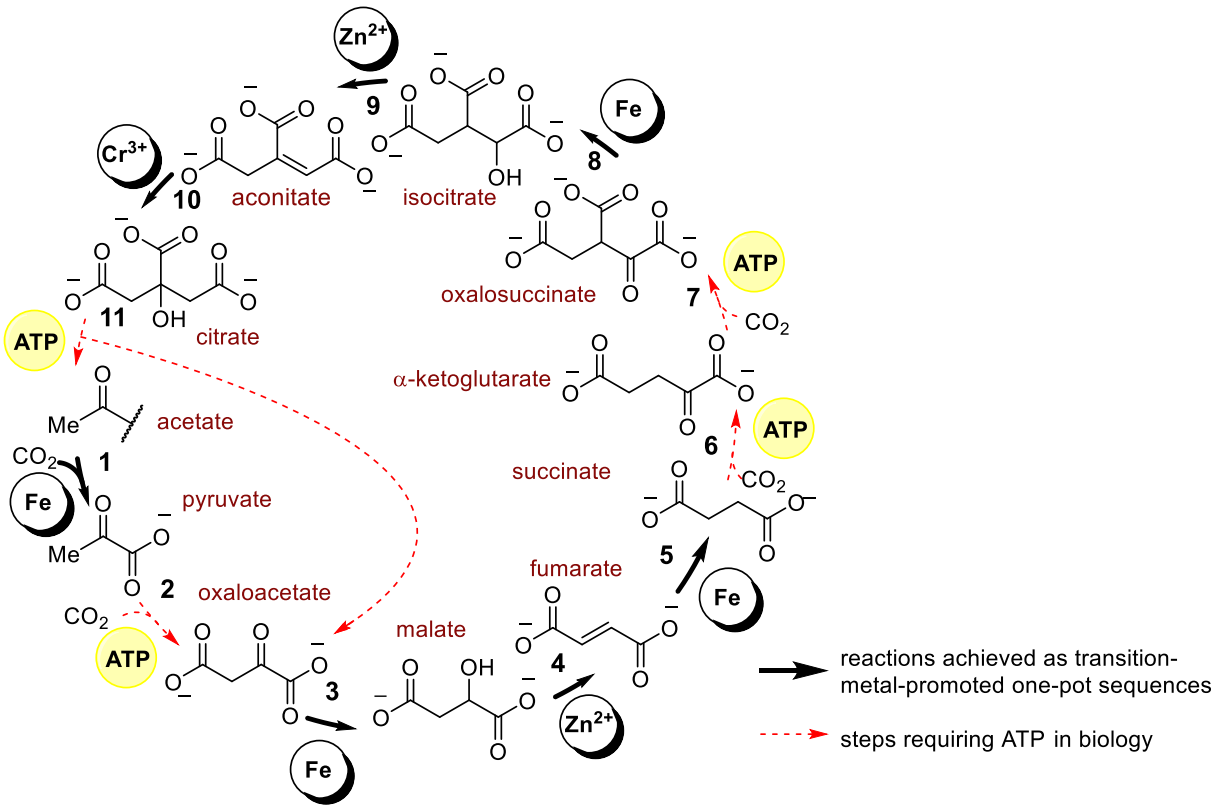


Figure 12. The ATP-independent reactions of the rTCA cycle promoted by transition metals (bold arrows). Reactions in dashed arrows were not succeeded. (figure adapted from *Muchowska, 2020* [26]).

Recent observations reported by Muchowska and co-workers (Figure 12) achieved the reactions that are ATP-independent in the rTCA cycle using transition metal (Fe^0, Zn^{2+}, and Cr^{3+}) catalysis. However, the steps that are ATP-dependent were missing. Their pursuing study on covering the ATP-dependent steps was suceeded by heating glyoxylate and pyruvate together with a Fe^{2+} salt. The created reaction network recaps 9 out of 11 metabolic intermediates of the rTCA and 7 out of 11 metabolic reactions of the TCA cycle.[92]

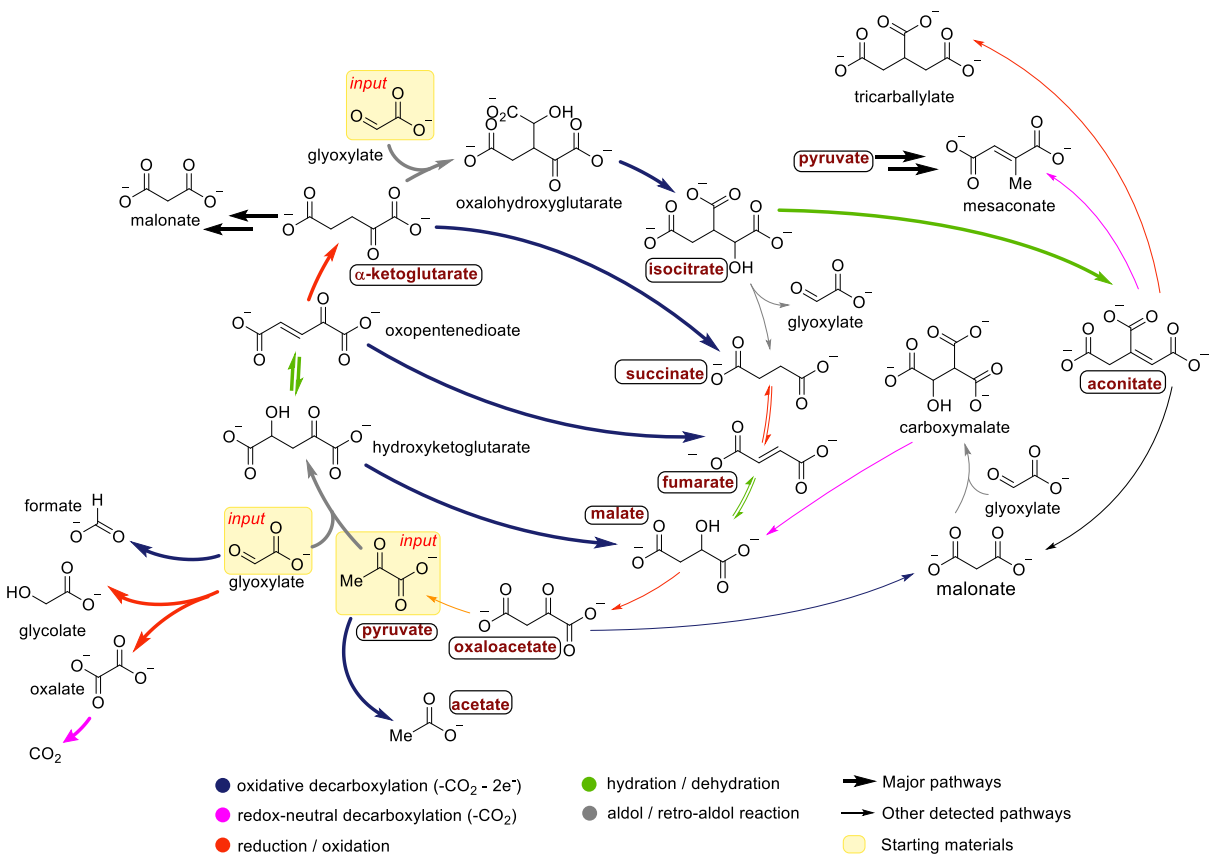


Figure 13. The Fe^{2+}-promoted reaction network arising from pyruvate and glyoxylate. Highlighted molecules are intermediates of the rTCA cycle. (figure adapted from *Muchowska, 2019* [92])

2.3 Amino acids syntheses

2.3.1 Introduction

Amino acids are the building blocks of peptides and make up 45% w/w of dry cellular mass.[93] The amino acids condensation results in polypeptides and the reaction is catalyzed by multiple enzymes in biology. Yet, biomolecules such as enzymes and ribonucleotides contain amino acid building blocks. This archetypal chicken-and-egg paradox remained unresolved and became one of the most popular research topics in prebiotic chemistry. Before looking at the prebiotic version of the amino acid syntheses, we first go through the biosynthesis of these central molecules. The shortest itinerary for a nitrogen atom to entermetabolism starts from N_2 or nitrate reduction.[94] The produced ammonia subsequently encounters α-ketoglutarate, an intermediate of the (r)TCA cycle, producing glutamate (GLU) via reductive amination (Figure 14A). The major function of glutamate is to transfer its amine group to other α-ketoacids via transamination.[95] The mechanism of this step is assisted by enzymes called *aminotransferases* or *transaminases* with the aid of a cofactor called pyridoxal phosphate (PLP). Another key role of glutamate is being the synthetic precursor for glutamine (GLN), an amino acid that acts as nitrogen donor in the pyrimidine nucleobases biosynthesis as well as in the synthesis of glutamate from α-ketoglutarate by transamination (*glutaminase* as enzyme). Thus, the glutamate/glutamine couple is the major hub for nitrogen metabolism in biochemistry.[96] Interestingly, 16 out of the 20 proteinogenic amino acids in nature are synthesized from three α-ketoacids (pyruvate, oxaloacetate and α-ketoglutarate) that are intermediates of the (r)TCA cycle (Figure 14B).[97]

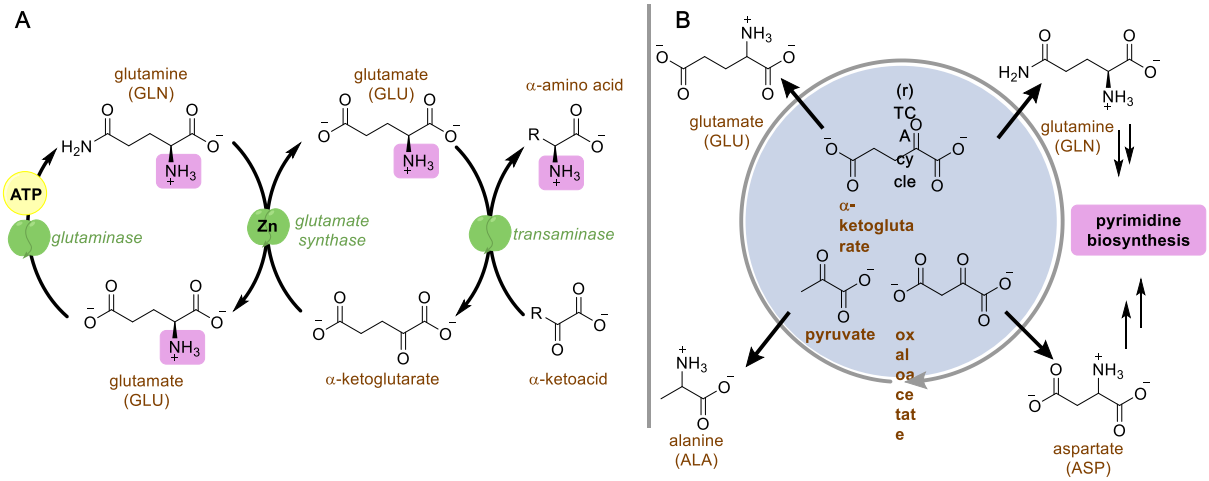


Figure 14. *a.* Enzymatic nitrogen mediation between glutamate/glutamine couple and other transamination processes. *b.* Biosynthesis of amino acids from α-ketoacids of the (r)TCA cycle.

2.3.2 Abiotic synthesis of amino acids

Even though multiple abiotic amino acids syntheses exist, such as the Miller-Urey experiment[6], Strecker synthesis[30] or the Bücherer-Bergs synthesis, [98] none of these resembles known biosynthetic pathways. A potentially feasible prebiotic synthesis could be the reductive amination or transamination of α-ketoacids, compounds which have been proved prebiotically accessible from the previous discussions.

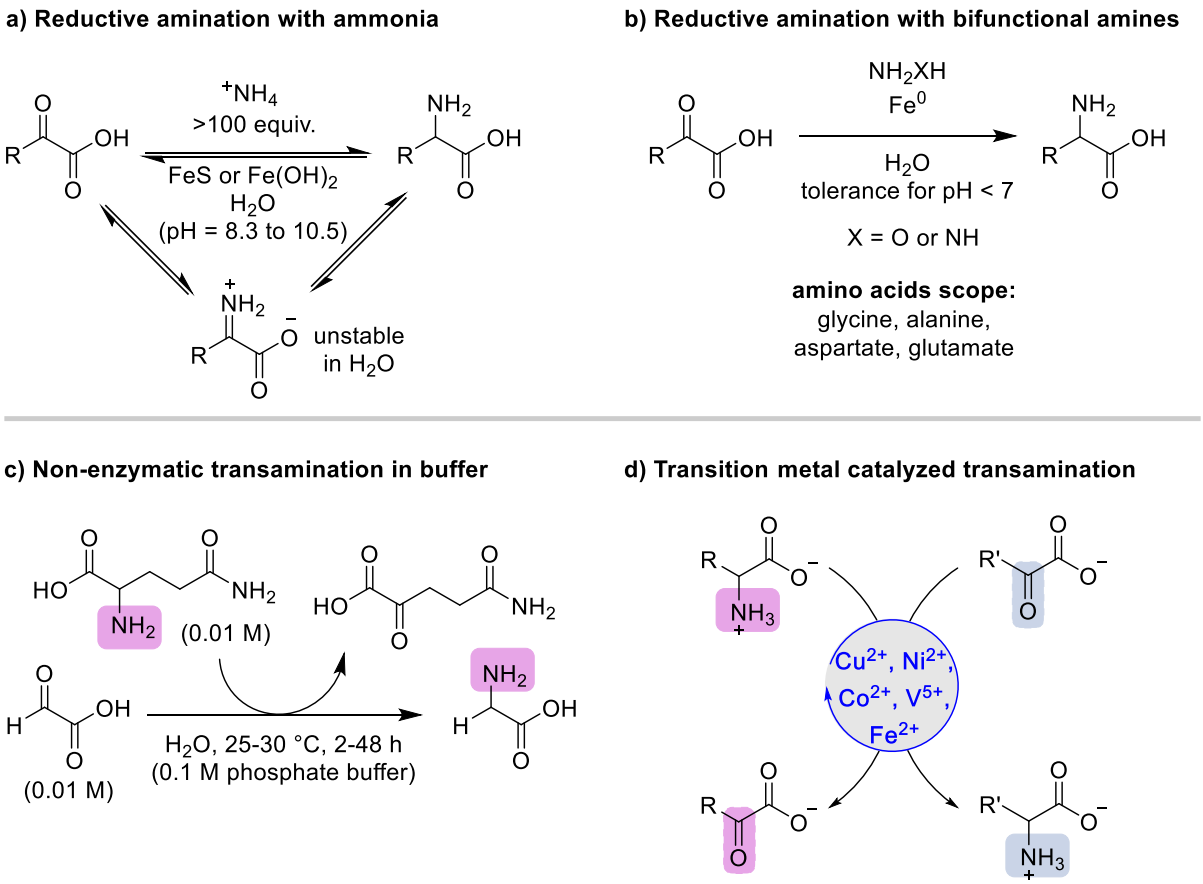


Figure 15. An overview of biological relevant prebiotic amino acids synthesis.

Nakajima and co-workers began their investigation with phenylpyruvate using the FeS-containing Schrauzer's complex and ammonia and ended up with great amount of phenylalanine by transamination. [99] This idea inspired the later work of Huber and Wächtershäuser. They adopted the previous method by using freshly prepared FeS and $Fe(OH)_2$ in ammonia solution. The medium with electron-rich metals favored the nonenzymatic reductive amination of many α-ketoacids except for oxaloacetate (Figure 15a). [100] This chemically unstable substrate was later included in the electro-reductive amination using FeS-Fe^0 by Kitadai.[91] However, the above mentioned reactions require a very large stoichiometric

excess of ammonia (typically ≥ 100-fold) to push the equilibrium towards the formation of the unstablized intermediate imine. This character raised doubts about the prebiotic feasibility of having such highly concentrated ammonia environment. Thus, an alternative nitrogen source has also been taken into account by other researchers. The Moran group recently explored the replacement of ammonia by stoichiometric hydrazine or hydroxylamine,[26] as hydroxylamine is a potential prebiotic nitrogen source (Figure 15b).[101, 102] Metallic iron was used as reductant for C=N bond of imine species and cleavage agent for N-N bond of the nitrogen donor. This prebiotic synthesis gave access to four biological amino acids which are glycine, alanine, aspartate and glutamate. However, the use of bifunctional amines (hydroxylamine and hydrazine) made the reductive amination blemished in a protometabolic context, because it is incompatible to use them with certain electrophilic functions, such as thioester and acyl phosphate, if they ever appear in the prebiotic mixture.

In addition to this, nonenzymatic transamination was very well studied in prebiotic chemistry. The first result was reported by Nakada and Weinhouse in 1953.[103] Glycine was obtained from glyoxylic acid using the glutamine/glutamate couple under buffered solution (Figure 15a). The transition metal catalyzed transamination was later investigated by Meisch *et al* (vanadium)[104] and Holanda et al (iron)[105], their reaction required heating to high temperature (70-100 °C) and the mechanism remained unclear. Recently, Mayer, Kaur and co-workers achieved the transamination with copper, nickel, vanadium and cobalt catalysis[106], under condtions that are much more relevant to biology (neutral pH and mild temperature from 20-50 °C) (Figure 15b). This result provided an expanded scope and improved reaction yield. The mechanistic studies uncovered the stabilization of imine intermediate by Ni^{2+}, Cu^{2+} and Co^{2+}, and the increasing of imine bond acidity by Ni^{2+} and V^{5+}. These results offered compelling evidence that amino acids are capable to be synthesized from α-ketoacids even in the absence of enzymes/cofactors under conditions that are biologically compatible.

2.4 Carbohydrate metabolism

2.4.1 Introduction

Sugar metabolism occurs in two directions. The anabolic pathway known as gluconeogenesis, synthesizes sugar from pyruvate and exhibits high uniformity across all life.[107] The catabolic pathways, namely glycolysis and the pentose phosphate pathway (PPP), break down sugars back to pyruvate. Both anabolic and catabolic pathways are present in archea, but the catabolic

pathways display a much higher diversification in different archaea species.[108,109] Intermediates from these pathways are mostly phosphorylated and often used as synthetic precursors that are incorporated into other metabolic pathways. For example, glycerol-3-phosphate is involved in lipid metabolism, phosphoenolpyruvate (PEP) – in aromatic amino acids synthesis, and ribose 5-phosphate (R5P) or phosphoribosyl pyrophosphate (PRPP) – in ribonucleotides synthesis. Here, I will selectively present the interconnected metabolic reaction sequence that allows for the formation of a ribonucleotide precursor (PRPP) from pyruvate, as this special pentose phosphate has its key role in metabolism that links the carbon fixation path and genetic molecules synthesis: its synthetic precursor (pyruvate) is produced from the carbon fixation pathway and itself works as primer for ribonucleotides synthesis. The participation of PRPP in ribonucleotide synthesis will be discussed in the next section (Chapter 2.5).

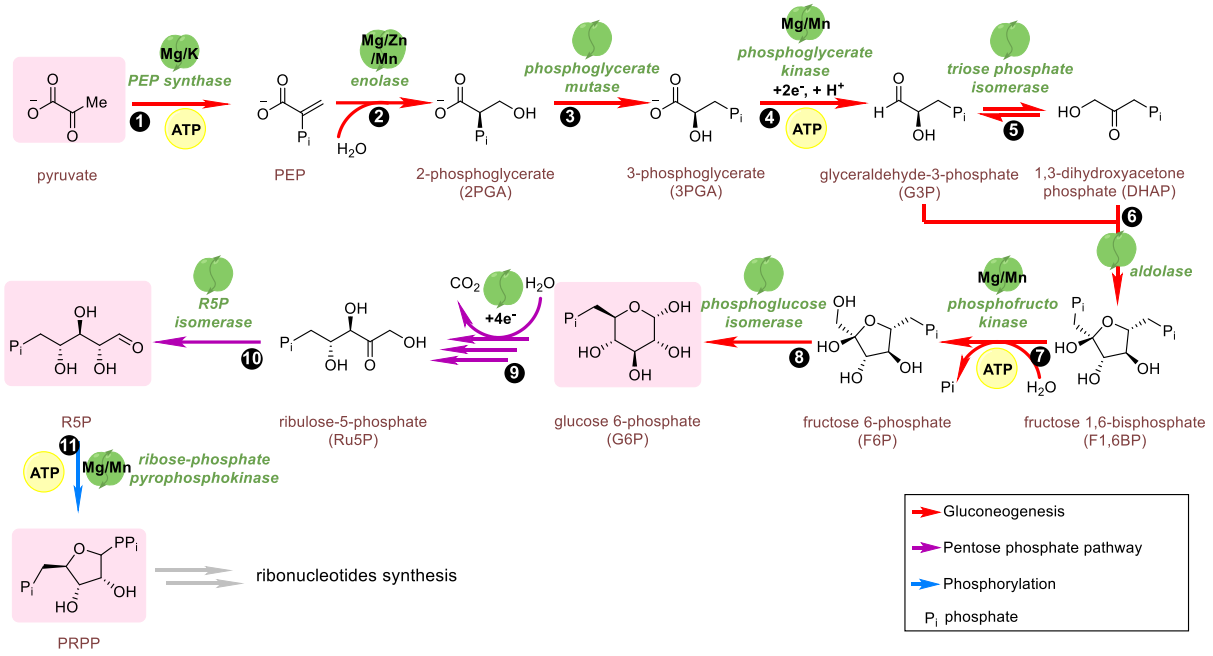


Figure 16. Overview of metabolic reactions from pyruvate to phosphoribosyl pyrophosphate (PRPP) and the enzymes which catylze these reactions. Every transitional intermediates are highlighted in pink boxes. (figure adapted from *Muchowska, 2020* [26])

PRPP is a major catabolic metabolite from sugar metabolism. It is a higly phosphorylated version of R5P at C1 position. In biology, living organisms first build up sugar phosphate from pyruvate by gluconeogenesis (red arrows in Figure 16). The resulted intermediate G6P in a pyranose form serves as a catabolic substrate for the pentose phosphate pathway (PPP) (purple arrows in Figure 16). The most interesting purpose of the PPP is the generation of various

pentose sugars, especially the R5P, which undergoes ATP phosphorylation towards PRPP (blue arrow in Figure 16). The process leading to PRPP includes two individual metabolic pathways, first anabolic then catabolic. The chemical mechanisms consist of different set of reactions containing phosphorylations, dephosphorylations, isomerizations, hydrations/dehydrations, and aldolisations. [110]

2.4.2 Prebiotic sugar metabolism

Unlike the formose reaction[111] or the Kiliani−Fischer homologation[112] which are much simpler mechanisms to give rise to sugars, the biochemistry of sugar metabolism is quite complex, it employs its distinct enzymatic catalysis and each of its intermediate has their own role in biology. The evolutionary origin of this complex system has been a mystery for a long time. One hypothesis was made to suggest that phosphate binding would be the key driving force for the evolution that predates enzymes.[113] Like other protometabolic ractions described above, the biological sugar metabolism should too have its own nonenzymatic version. Although studies on the nonenzymatic analogues of gluconeogenesis, glycolysis and PPP remain scarce, some small subsets of reactions of these pathways promoted under prebiotic conditions have been reported.

The aldol reaction that makes the C-C bond is very well explored in prebiotic research (step 6 in Figure 16). Ralser and co-workers achieved this step by freezing (-20 °C) G3P and DHAP mixture in glycine/$NaHCO_3$ solution for 8 days.[114] The use of metal ions[115] and surface-templated minerals[116,117] in the similar type of aldolisation was found to accelerate the reaction kinetic. The templating effect on the silica surface led to the pre-organization of substrates and also catalyzed the phosphorylation of ribose (including R5P) to PRPP after multiple wet-dry cycles.[118] The choice of activating and phosphorylating agents is also a big challenge in prebiotic research. An ideal prebiotic activating or phosphorylating agent is required to be simple, reactive and accessible in an early Earth's environment, to mimick the role that ATP plays in current biology. Simple thioacids, acyl phosphates, inorganic polyphosphates or cyclic polyphosphates are recent candidates to do so, even though hydrolysis of these presents major issue during the reaction in solution.[119,120,121] To summarize, much remains to be done before a prebiotic analog of gluconeogenesis is accomplished nonenaymatically.

In contrast to gluconeogenesis, more systematic studies have been performed on sugar catabolic pathways. The most comprehensive approach was published by Ralser and co-workers.[122,123]

The team achieved 29 interconversions present in glycolysis and the PPP by using soluble ferrous salts in pure water at 70 °C. The breakdown of 6-phosphogluconate (oxidized form of G6P) generates R5P (Figure 17), the synthetic precursor for ribonucleotides. Here, after demonstrating that aspartate and glycine could be obtained under plausible prebiotic conditions, R5P, another key building block for ribonucleotides synthesis, was proved as prebiotically available precursor that was obtained from nonenzymatic reactions.

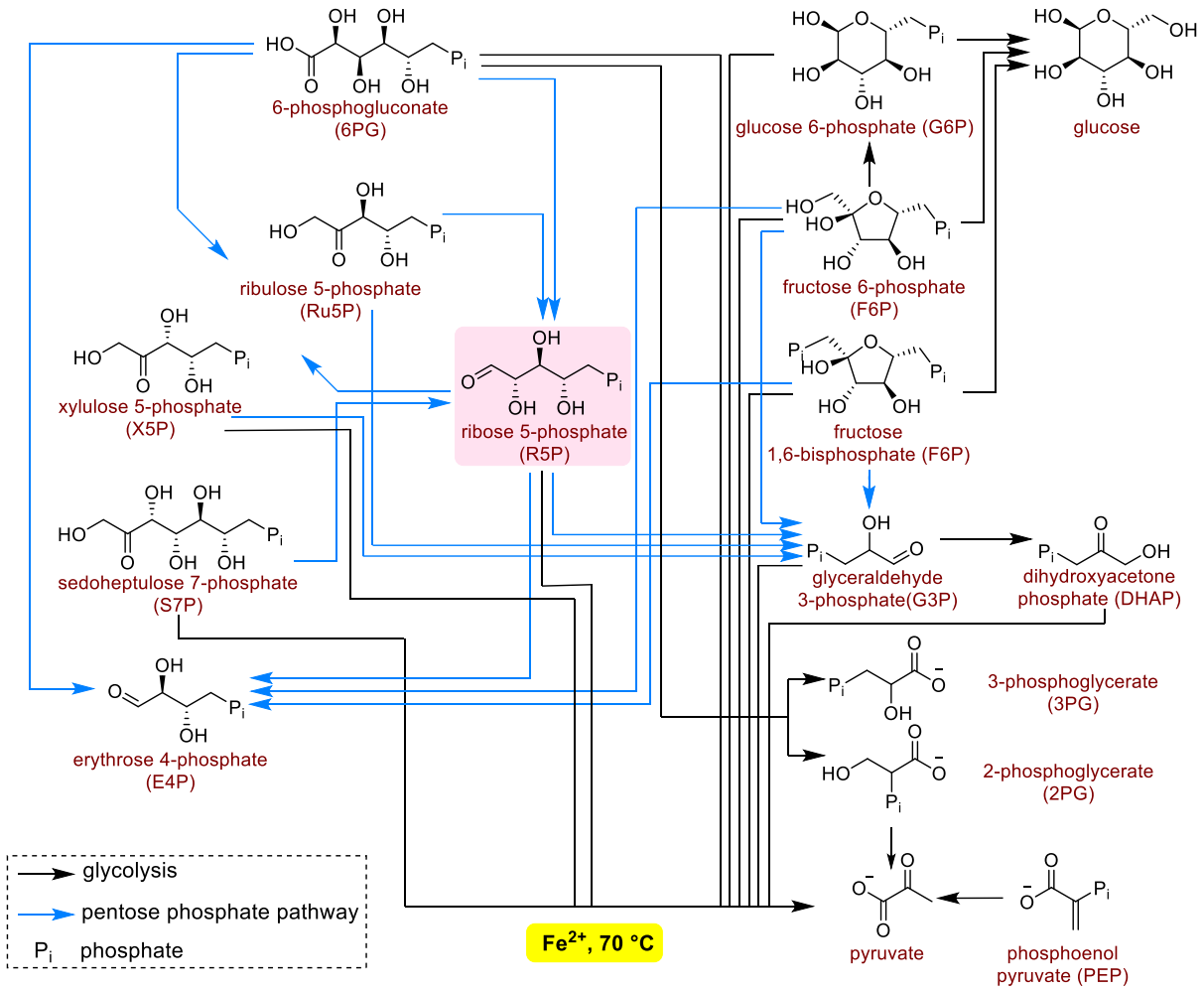


Figure 17. Fe-promoted glycolysis (black arrows) and pentose phosphate pathway reported by Ralser. (figure adapted from *Muchowska, 2020* [26])

2.5 Ribonucleotides metabolism

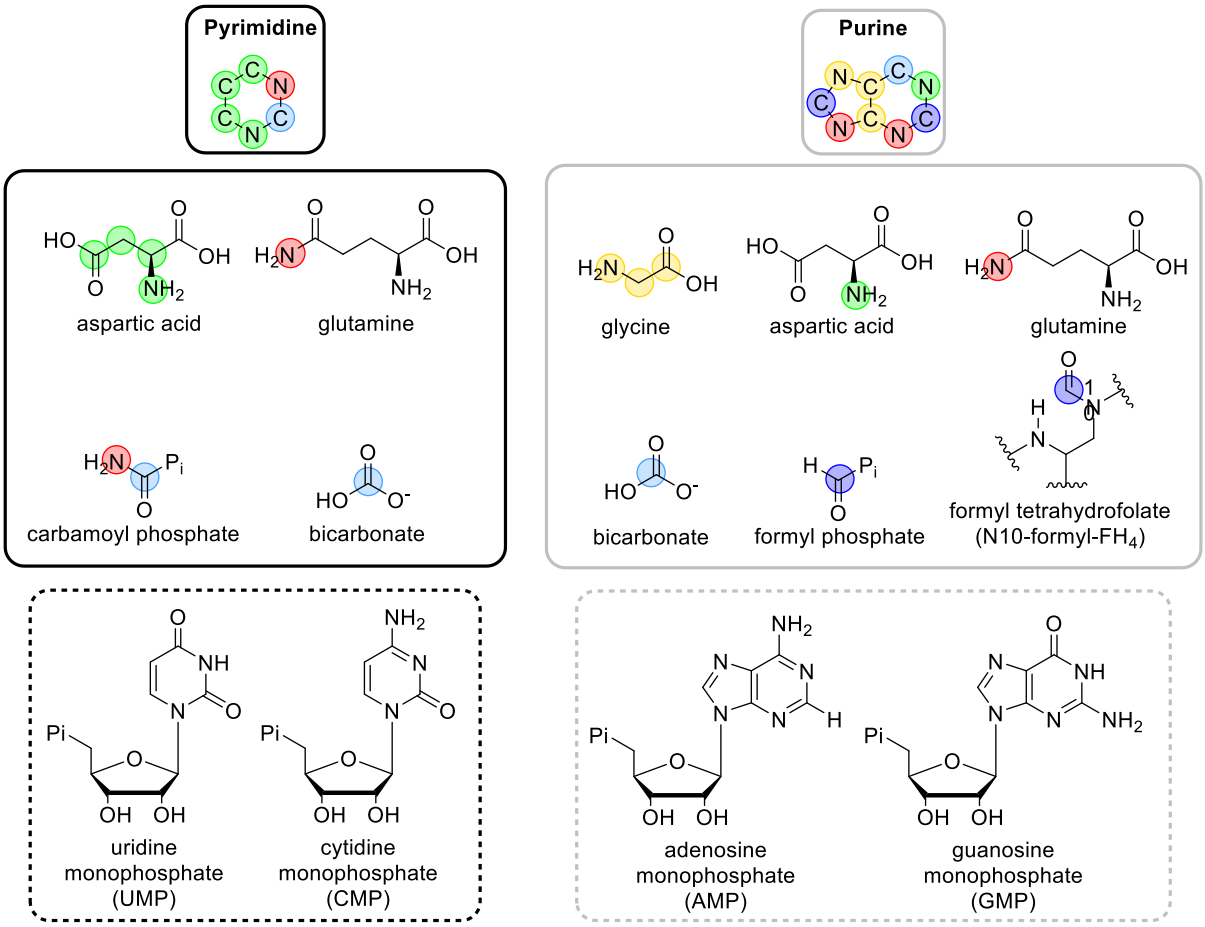


Figure 18. Biosynthetic origins of pyrimidine and purine nucleobase scaffolds, and their ribonucleotides.

Ribonucleotides are one of the most important biochemical intermediates in life because, at first, they are the main building blocks of the genetic polymers (RNA, DNA); secondly, the triphosphorylated ribonucleotides (ATP, GTP[124], UTP[125] and CTP[126]) are direct energy stock in biochemistry as the P-O bond hydrolysis releases energy to sustain the energy-dependent metabolic transformations; thirdly, their derivatives are included in cofactor structures, such as folates and flavins that act as electron shuttles in biochemical redox reactions. In biochemistry, ribonucleotide oligomers can be divided into two big families: pyrimidine ribonucleotides and purine ribonucleotides. The key feature that discriminates these structures is based on their base scaffolds, whether they are monocyclic pyrimidines or bicyclic purines (Figure 18). The biochemical synthesis of these ribonucleotides involves life metabolite intermediates, such as amino acids (glycine, aspartate and glutamine) and pentose phosphate pathway intermediates (R5P or PRPP, R1P), that react together with the aid of different enzymes/cofactors and ATP.[43]

In RNA polymers, pyrimidines are two of the four key nucleobases and include uridine (U) and cytidine (C) (black dashed box in Figure 18). The other two key structures are purines: adenosine (A) and guanosine (G) (grey dashed box in Figure 18).

2.5.1 Biosynthesis of ribonucleotides

Without overlooking the undeniable importance of ribonucleotides in the origin of life, especially in the RNA world hypothesis, the way to produce these molecules and the way they participate in the emergence of early life remains unclear in prebiotic chemistry. A plausible solution to interpret this puzzle would refer to the protometabolic model (in chapter 1.2.1), in which was described the emergence of the spontaneous self-organized reaction networks by nonenzymatic catalysis under primitive environment. As suggested by Harrison and Lane, the way that life produces ribonucleotides today could be considered as guide for the prebiotic synthesis of these molecules.[127] Instead of searching in promiscuous chemical reactions that were possibly occurred in the origin of life, the biosynthesis provides a more logic and guided conception that life would undertake from the beginning.

2.5.1.1 Purine synthesis

Purines are made from the coupling of PRPP and amino acids (GLY and ASP), followed by several often chemically similar steps: folate-assisted formylations (step 3, 9 in Figure 19), aminations (step 1, 2, 4, 7, 10; some of them are ATP-dependent, cyclization (step 5), carboxylation (step 6) and elimination (step 8). Inosine monophosphate (IMP) formed in this pathway serves as synthetic precursor to other purine ribonucleotides. The divergent synthetic route to adenosine and guanosine nucleotides from IMP involves a two-step GTP-assisted amination (step 11), and two-steps redox-dependent amination (step 12) respectively.

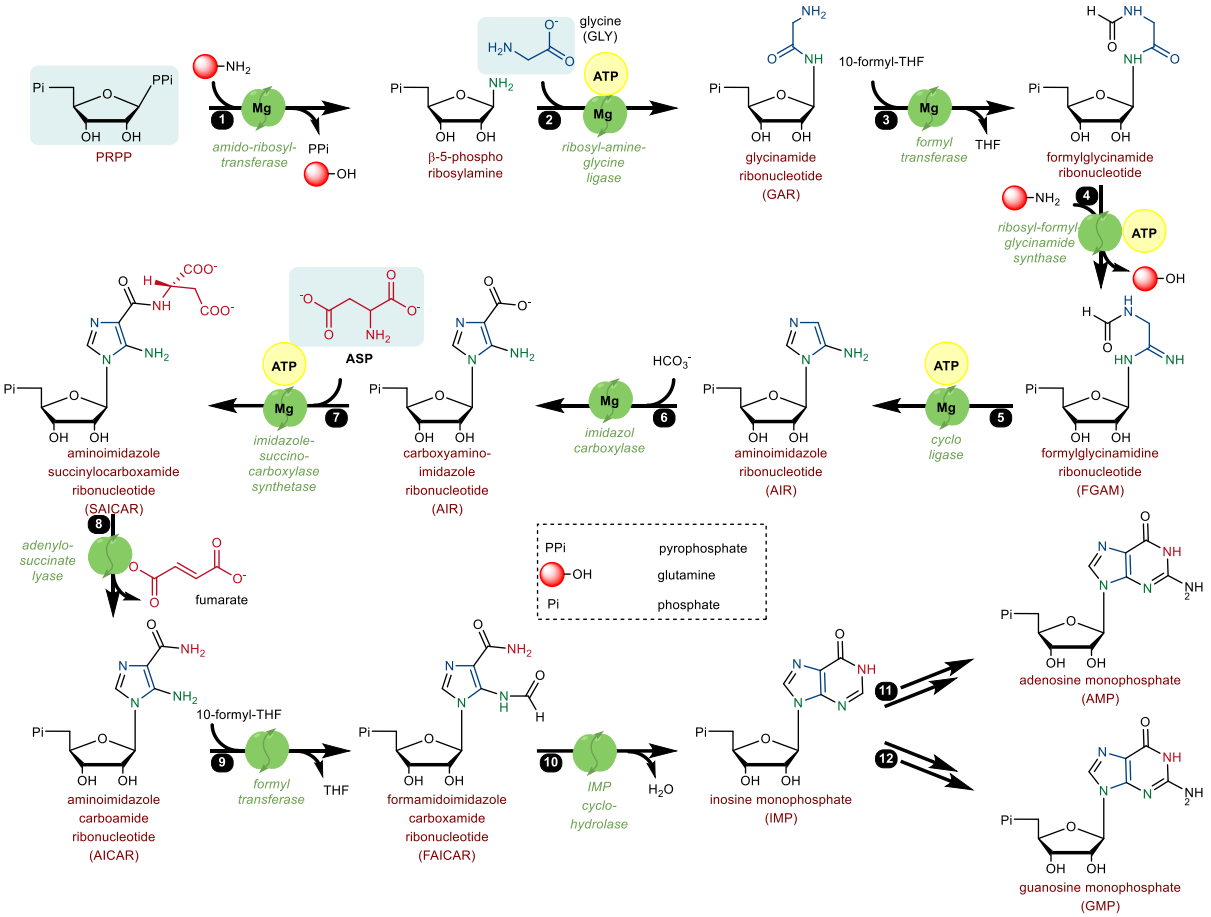


Figure 19. The *de novo* biosynthesis of purine ribonucleotides

2.5.1.2 Pyrimidine synthesis

The biosynthetic pathway towards pyrimidine ribonucleotides is shorter than the one towards the purines. It includes two distinct processes: the *de novo* biosynthetic pathway (black arrows, Figure 20) and the salvage pathway (grey arrows). The *de novo* biosynthesis of ribonucleotides starts from simple amino acids and includes multiple steps: (1) glutamine transfers its ammonia moiety to the free C1 compound (bicarbonate) to yield carbamoyl phosphate (CAP, an active phosphorylated intermediate from urea cycle in metabolism) (step 1); (2) aspartate is then carbamoylated with the high energy phosphorylated molecule to give carbamoyl aspartate (CAA) (step 2); (3) the dehydrative cyclization (step 3) of the CAA leads to the formation of the six-membered ring dihydroorotate (DHO); (4) the oxidation of DHO (step 4) to the nucleobase orotate (ORO); (5) the N-C bond coupling between the ORO and PRPP (step 5a); (6) the decarboxylation of OMP gives rise to UMP (step 6), which is followed by multiple phosphorylations and the capture of ammonia yielding UTP and CTP, respectively (step 7-8).

On the other hand, the salvage pathway allows life to regenerate its ribonucleotides by recycling free nucleobases (uracil and cytosine) from the catabolism of the corresponding ribonucleotides, and further coupling with the pentose phosphate (step 5b, 5c). The difference here compared to the *de novo* path is the use of R1P as pentose donor instead of using PRPP. However, both pathways provide feedstocks to genetic molecules.

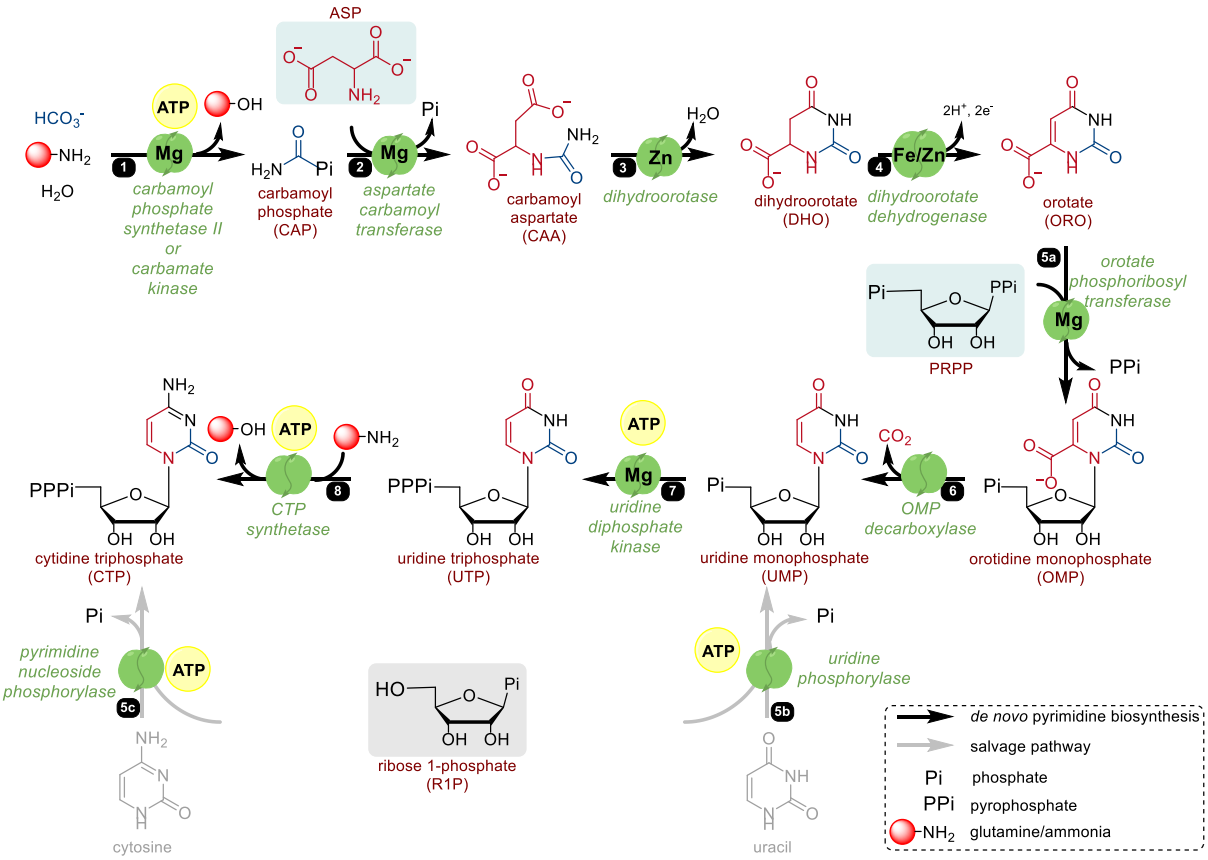


Figure 20. Biosynthesis of pyrimidine ribonucleotides including *de novo* biosynthesis (black arrows) and salvage pathway (gray arrows).

Considering the prebiotic context, the *de novo* pathway of ribonucleotides synthesis seems to appear before the salvage pathway. This is because the building blocks of the *de novo* pathway are much simpler and considered prebiotically accessible, meaning that they could have been supplied by the various proto-metabolic pathways described above (AcCoA pathway, rTCA, reductive amination/transamination and sugar metabolism).

2.5.2 Prebiotic analogue of ribonucleotides synthesis

2.5.2.1 Chemical synthesis of ribonucleotides

Ever since the "RNA world" hypothesis was made, multiple chemistries have been proposed over the past decades in attempts to corroborate this hypothesis experimentally, using mostly HCN or its derivatives as starting material. Several examples will be discussed below.

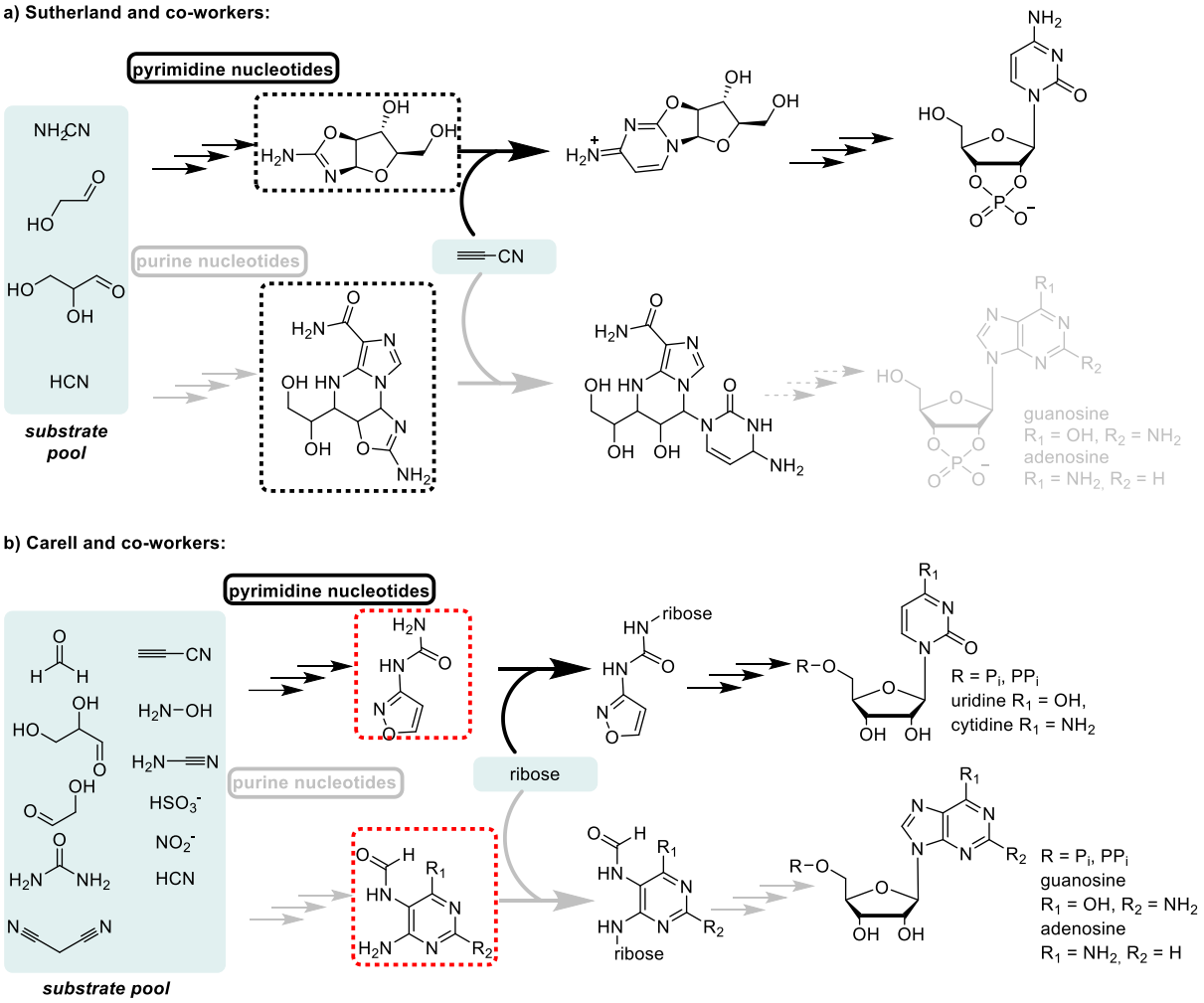


Figure 21. *a*. HCN-based syntheses of ribonucleotides that bypass ribose by Sutherland and co-workers.[31,128] (solid arrows are reported chemistry, dashed arrows are proposed reactions) *b*. Unified chemical synthesis of RNA ribonucleotides by Carell and co-workers.[129]

Sutherland and co-workers made the pyrimidine nucleotides and purine precursors from HCN, glyceraldehyde, cyanamide, glycoaldehyde and excess of isocyanomethane (Figure 21a).[31,128] Unlike in biosynthesis, these purely chemical syntheses involve several reaction steps that are all distinct from each other (reagents, solvent, pH, temperature, catalysts). Instead of the

biochemistry-like N-C coupling between the bases and the pentose phosphate, Sutherland and co-workers access ribonucleotides via the intermediacy of pentose amino-oxazoline (molecules in black dashed boxes). This approach bypasses the use of ribose. In his experiments, sequential changes to the reaction conditions are needed, which requires well-timed monitoring and even isolation of certain intermediates to pursue the next step. The utilization of UV irradiation as energy source can also cause criticism, as most biological and geological evidence support a later emergence of photosynthesis compared to chemosynthesis,[55] driven by inorganic catalysts such as minerals. Although the Sutherland team's experiments produce both pyrimindine and purine ribonucleotides, the mismatch between their chemistry and the metabolic reactions used by life today makes this approach the field of origin of life.

Another good example of an experimental approach towards ribonucleotides was published by Carell and co-workers.[129] In this case, four of the RNA nucleotides (A, U, G, C) are accessible and the chemistries between the pyrimidines and purines have more similarities (Figure 21b). The difference between the Carell approach and the Sutheland approach is the intermediacy of isoxazolylurea and formamidopyrimidines in the former (compounds in red dashed boxes). These compounds further react with ribose to form pyrimidine and purine ribonucleotides, respectively. However, these reaction sequences still needs significant occasional human intervention[130] to change conditions and even solvents, making it more similar to a human-led total synthesis of a natural product than to the ribonucleotide biosynthesis, where molecules are produced and react between one another in a continuous way.

These two examples provide evidence that ribonucleotides could be synthesized from purely chemical synthesis which is not similar to modern biochemistry. However, it still remains an open question how a promiscuous chemical system would have evolved to the highly self-organized metabolic networks we know today. How does primitive life adapt arise from cyano-based compounds and then reinvent itself, leaving behind no trace of its original chemistry? A long evolutionary time is often invoked to rationalize such a transition but it is still a poor justification.[131]

2.5.2.2 "Biology-like" syntheses of ribonucleotides

An alternative explanation for the emergence of ribonucelotides could refer to the continuity with modern biochemistry. As recently suggested by Harrison and Lane that "*to understand the origin of life, we would be foolish to ignore life as a guide*", meaning that a prebiotic ribonucleotide synthesis could have been a close mimic of its biosynthesis.[132] However, much prebiotic chemistry has been described that conceptually resembles ribonucleotides biosynthesis[133], but very few examples involve precisely the same reactions and substrates.

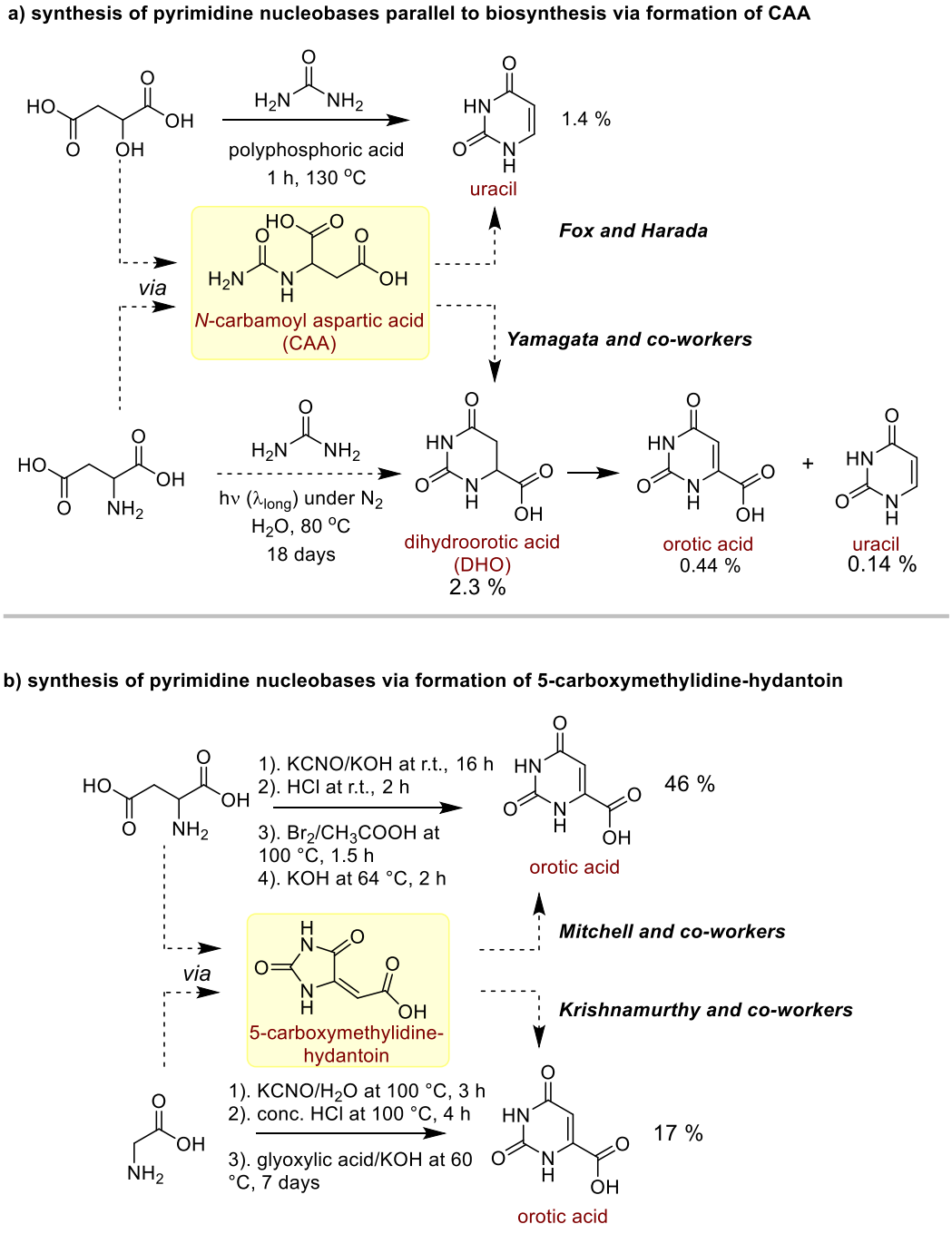


Figure 22. Prebiotic analogues of pyrimidine nucleobases syntheses

For purine synthesis, experimental studies leading to adenine were done by Lagoja and Herdewijn, starting from the condensation of glycinamide and *N,N'*-diformylurea (first step in Figure 23a).[134, 135] This reaction involves three distinct steps. The first step was carried out in an organic solvent, dimethylformamide (DMF), at a very high temperature. The second one was carried out in a heterogeneous medium using organic phosphorylating agent trimethylphosphate. The starting materials could be considered asprebiotically accessible: glycinamide and *N,N'*-diformylurea could be obtained from glycine, ammonia and CO_2. The use of organic solvent and organic phosphate, however, makes the prebiotic plausibility of these conditions rather unikely. Other experiments, such as the purine synthesis from glycine by Springsteen and co-workers, are also quite debatable in terms of their prebiotic relevance because of the dependence on nitrile chemistry.[136]

At the same time, the research focusing on the "biology-like" pyrimidine synthesis was also very well investigated. Fox and Harada obtained uracil from a mixture of malic acid and urea in polyphosphoric acid solution (Figure 22a, top equation).[137] Although the substrates are not identical to those in the biological pathway, the intermediate *N*-carbamoyl aspartic acid, CAA, corresponds to the biological metabolite. These substrates are accessible from other prebiotic reactions. A similar strategy was proposed by Yamagata and co-workers (Figure 22a, bottom reaction).[138] The study focused on the use of aspartic acid and urea in buffered solution. After thermal photo-irradiation, orotic acid (ORO) and uracil (URA) were obtained in a synthesis analogous to the *de novo* pathway but without enzymes. Another prebiotic analogue of pyrimidine nucleobase synthesis of ORO was achieved by Mitchell and co-workers (Figure 22b, top equation).[139] Unlike the previous results, this experiment involves multi-step synthesis starting from aspartic acid and potassium cyanate (KOCN) and requires crystallization of each of the intermediates. The resulting five-membered ring intermediate 5-carboxymethylidine-hydantoin replaces the role of dihydroorotic acid (DHO) in the synthetic pathway, which differs from the *de novo* pathway. Another study was carried out by Russell and co-workers using a similar chemical pathway.[140] A related pathway based on cyanate-dependent chemistry was highlighted recently by Krishnamurthy and co-workers. They use glycine, KOCN and glyoxylic acid as building blocks (Figure 22b, bottom equation).[141] The reaction also resulted in ORO via formation of 5-carboxymethylidine-hydantoin.

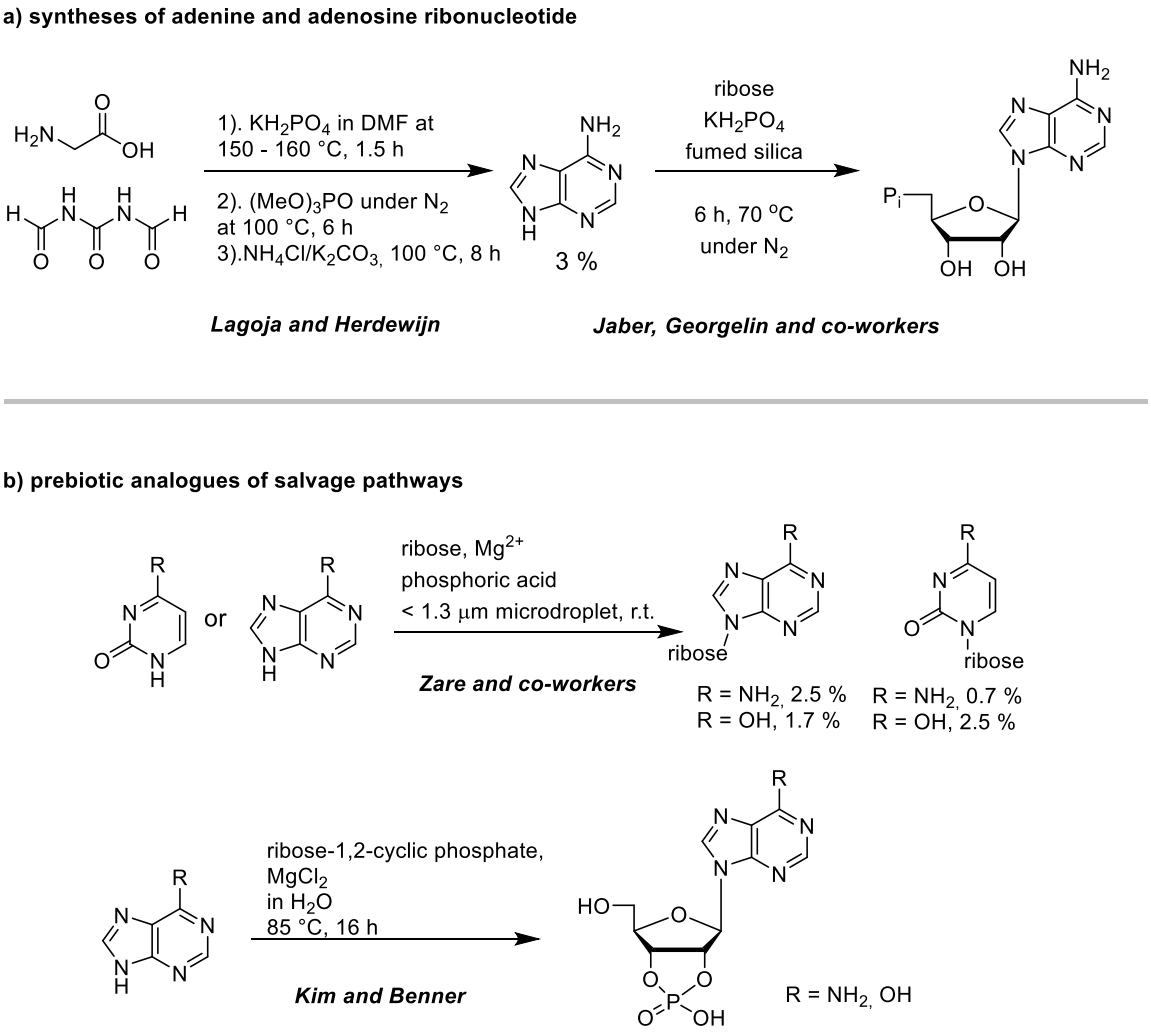


Figure 23. Prebiotic analogues of salvage pathways

Although the prebiotic syntheses of nucleobases are very well studied, the major challenge in the prebiotic ribonucleotide synthesis remains to be the *N*-glycosylation between the base and sugar and the following steps (phosphorylation, amination). Studies on prebiotic analogues of salvage pathways have been performed in order to link these two building blocks. Jaber, Georgelin and co-workers reported the silica-supported N-C coupling reaction between adenine and ribose under buffered thermal conditions (second reaction in Figure 23a).[142] Other purine and pyrimidine nucleobases were further examined using amidotriphosphate as activating agent[143, 144] or with the assistance of compartimentalization in aqueous microdroplets[145] (Figure 23b).

Aim of this thesis

Although the genetic-first hypothesis argues that the prebiotic synthesis of ribonucleotides should have been distinct from their biosynthesis, the metabolism-first origin of genetics remains highly appealing and presents numerous strong advantages compared to the former. In contrast to the above-described literature results which relied on the photochemistry or cyanide-based chemistry, the catalytic potential of minerals, clays and metals has been much less explored. To fill this gap, the research I have performed during my doctoral studies envisaged to achieve the prebiotic synthesis of pyrimidine ribonucleotides in a way that parallels their biosynthesis, using metal catalysts. Consequently, this thesis includes two major experimental parts:

1. The nonenzymatic analogue of *de novo* pyrimidine nucleobase synthesis (ORO, URA), including first investigation of the individual steps which are then optimized and step-by-step combined into a "one-pot" synthesis. These results have been recently published in a journal.[149]

2. The nonenzymatic coupling of the nucleobase ORO and phosphorylated pentose (R5P, PRPP). This part remained challenging and therefore only preliminary results will be shown here.

PART II

3 Towards nonenzymatic synthesis of pyrimidine nucleobases

3.1 *N*-Carbamoylation

We began our investigation from the first step of the *de novo* pathway. (Figure 24) This N-carbamoylation includes the reaction between the ASP (from nonenzymatic reductive amination of oxaloacetate[43]), bicarbonate and an ammonia source (ammonium or glutamine). The combination of the last two reagents leads to the formation of the intermediate CAP, which serves as carbamoyl-donor reagent.

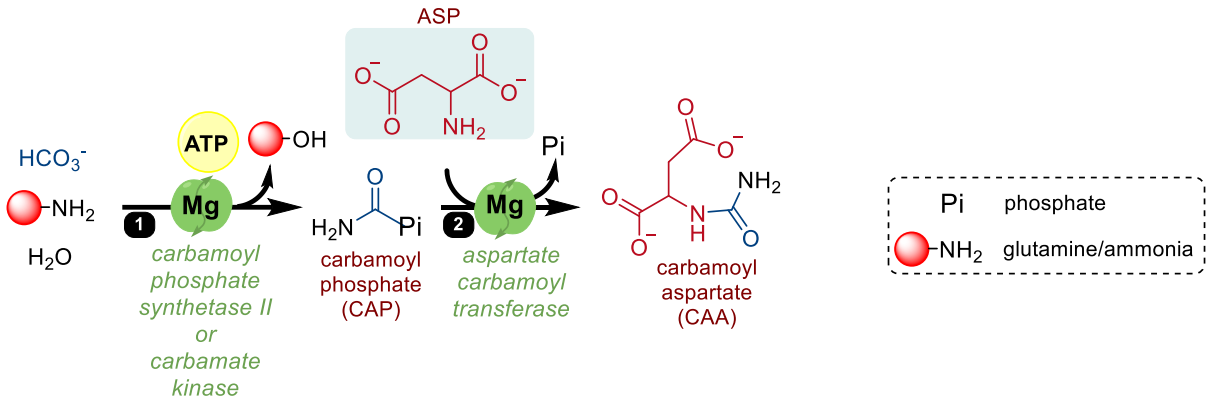


Figure 24. *N*-Carbamoylation of ASP in the *de novo* pyrimidine synthesis via formation of intermediate CAP.

3.1.1 C1 compound, ammonia source and phosphate screening

The prebiotic availability of carbamoyl phosphate (CAP) is uncertain because of its instability in water. It is known that the highly activated intermediate CAP could be obtained from the urea cycle in biological metabolism, where glutamine/ammonium transfers its ammonia group to the bicarbonate species with the aid of ATP. A chemical equilibrium between the CAP and cyanate in the presence of excess phosphate was also reported in some prebiotic publications.[146] Here, we began the first set of screening reactions with different combinations of substrates containing the main elements of C, N, P, which react with aspartate at different temperature for 16 h (Figure 25a).

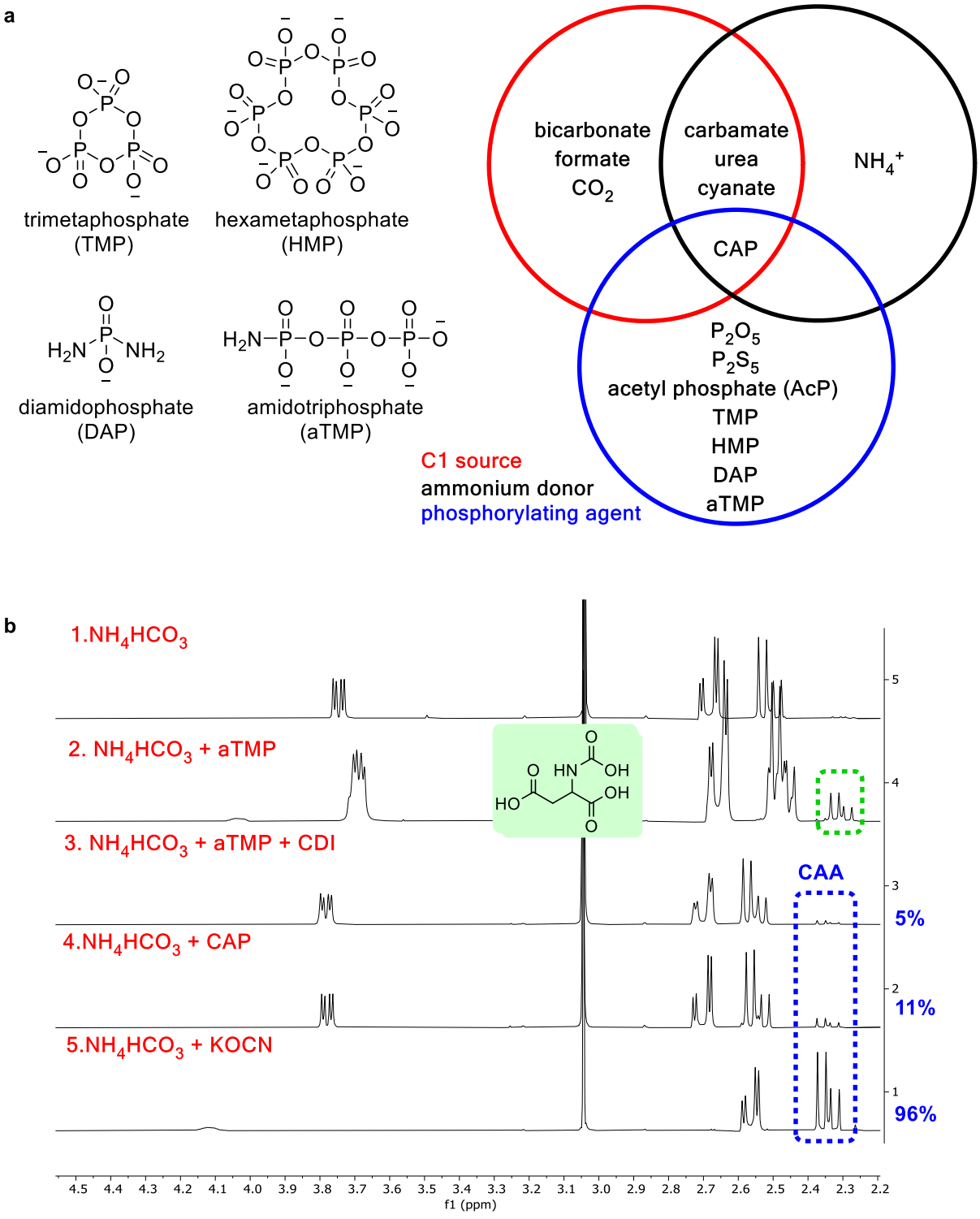


Figure 25. *a*. Key building blocks for the carbamoylation step; each combination of substrates was screened at different temperatures. *b*. ^{1}H NMR result of screened conditions which showed the formation of new compounds. In spectra 1-4, the non-specified signals correspond to ASP and internal standard dimethylsulfone (DMS, ~3.05 ppm). Reaction condition: 8.5 mM of D,L-aspartic acid, 10 equiv. of NH_4HCO_3, 1 equiv. of phosphate (aTMP or CAP) or KOCN, 1 equiv. of CDI in 1mL of MilliQ water at 60 °C for 16 h.

Amongst these screened reactions, most gave negative results, with only a few of them producing interesting products whose ^{1}H NMR spectrum is shown above (Figure 25b). The reaction of NH_4HCO_3 with ASP at 60 °C provides no results (spectrum 1). When aTMP was added to the mixture, a new compound appeared but could not be identified as CAA (spectrum 2). After comparing to different authentic samples, we suggested the structure as *N*-carboxyaspartic acid (highlighted in green dashed box). This undesired carboxy amino acid could be further converted to CAA (blue dashed box) with the addition of an additional activating reagent, carbonyl diimidazole (CDI), which is commonly used in chemical synthesis to activate the carboxylic acid by increasing its electrophilicity (spectrum 3).[147] The proposed reaction mechanism is described below (Figure 26). A similar result was also observed when using formate as C1 source. Commercial CAP was also assayed, but the resulted formation of CAA was observed with a relatively low conversion (~ 11%).

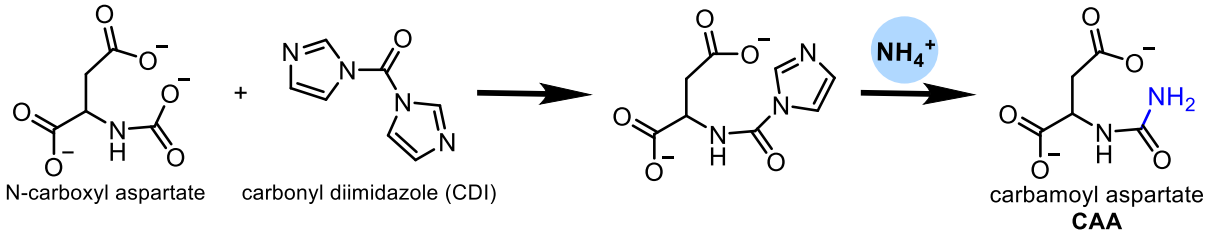


Figure 26. The plausible mechanism of CAA formation via formation of carboxy aspartate assisted by CDI.

The best result was achieved using KOCN, just as reported before in Mitchell's paper[139]. The reaction between ASP and KCON gave nearly quantitative yield of CAA. However, unlike the cyanate chemistry, the reaction with CAP to prepare CAA, as in the biological pathway, was never explored under prebiotic conditions. The goal of our study was to see whether the biological pathway could have had nonenzymatic origins, and not to find the most direct synthetic route. So, whether cyanate is an intermediate in the reaction starting from CAP does not diminish the importance of the observation that the reaction works starting from CAP. The following experiments will therefore focus on the use of CAP as a reagent to transfer its carbamoyl group onto ASP.

Due to the high resemblance of the chemical shift between CAA and *N*-carboxyaspartate, an extra analysis by LC-QTOF-MS was done to confirm the structure of products. CAA is compared with its authentic sample, while *N*-carboxyaspartate is not visible when analysed using this technique.

3.1.2 Reaction optimization

The reaction between ASP and CAP was firstly assayed over a pH range of 3-9 (Figure 27). The results indicate a higher reactivity/stability of CAP under mild basic solution. The acidic pH was adjusted with a dilute solution of HCl, whilst pH ≥ 7 was obtained with the addition of different solid salts. The optimal yield was obtained at pH 8 with the addition of 2 equiv. of solid $NaHCO_3$. However, the presence of ammonium salt or phosphate salts at the same pH inhibit the carbamoylation reaction. From a mechanistic standpoint, the nucleophilic character of ammonium salt would trap the reactivity of CAP which degrades to urea, thus decrease the yield. Another thing is that cyanate is known to be produced, along with phosphate, by thermal fragmentation of CAP and is therefore a likely intermediate in the reaction.[146] In line with this proposal, the use of pH 8 phosphate buffer, rather than water adjusted to pH 8 with bicarbonate, decreased the yield of CAA to 10%, which might be explained by a less favorable equilibrium between CAP and cyanate in the presence of excess phosphate.

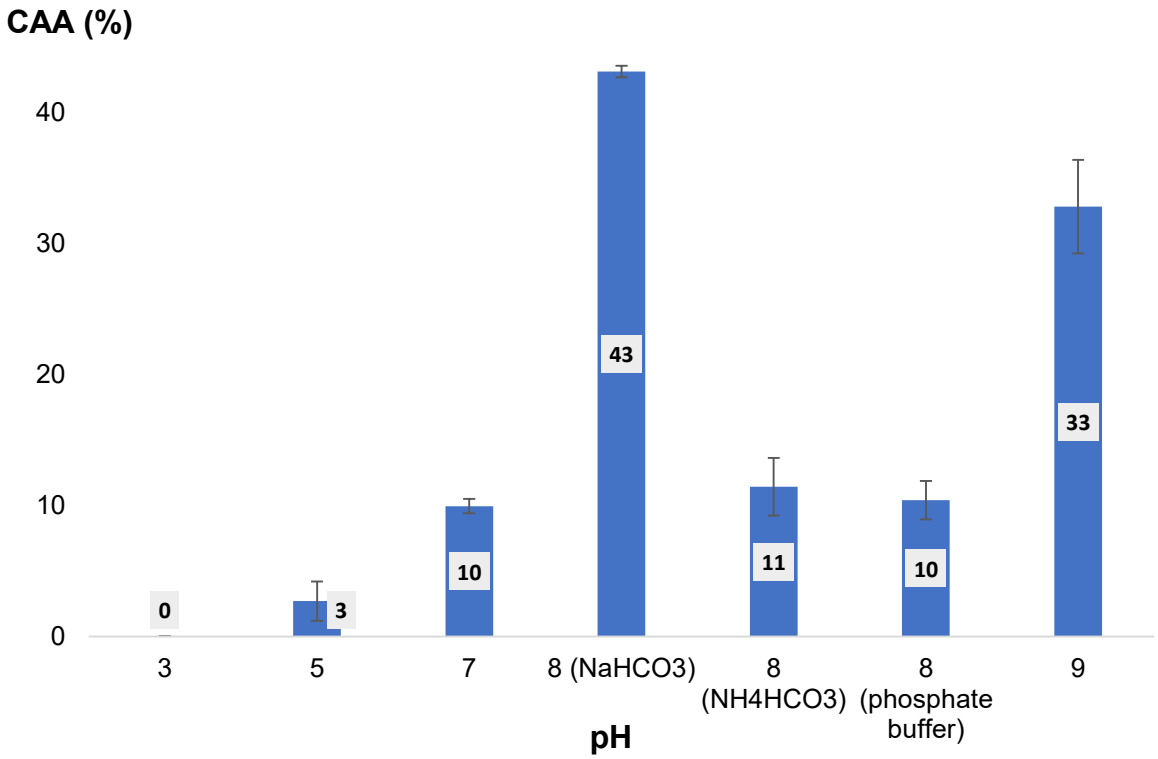


Figure 27. The effect of pH and inorganic salt on the carbamoylation reaction between ASP and CAP. Reaction condition: 8.5 mM ASP, 1 equiv. of CAP, 2 equiv. of bicarbonate salts (pH = 8) or 10 equiv. of $NaHCO_3$ salt (pH = 9) in 1 mL of MilliQ water or in 1 mL of 1 M phosphate buffer (pH = 8), 60 °C for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

The carbamoylation reaction was then assayed over different metals (Figure 28). A big range of alkaline metals and transition metals were screened, including most of the first and second rows of metals in the periodic table. However, the results of these metals did not exhibit any promoting effect to this reaction, amongst which have presented negative effects on this reaction (eg. Mn^{2+}, $Cu2^{+}$, Fe^{3+}, Co^{2+}, Ni^{2+} salts kill the reaction). A plausible reason would be the degradation of CAP in the presence of such metals but more evidence is needed to confirm this, which lies beyond the scope of this project.

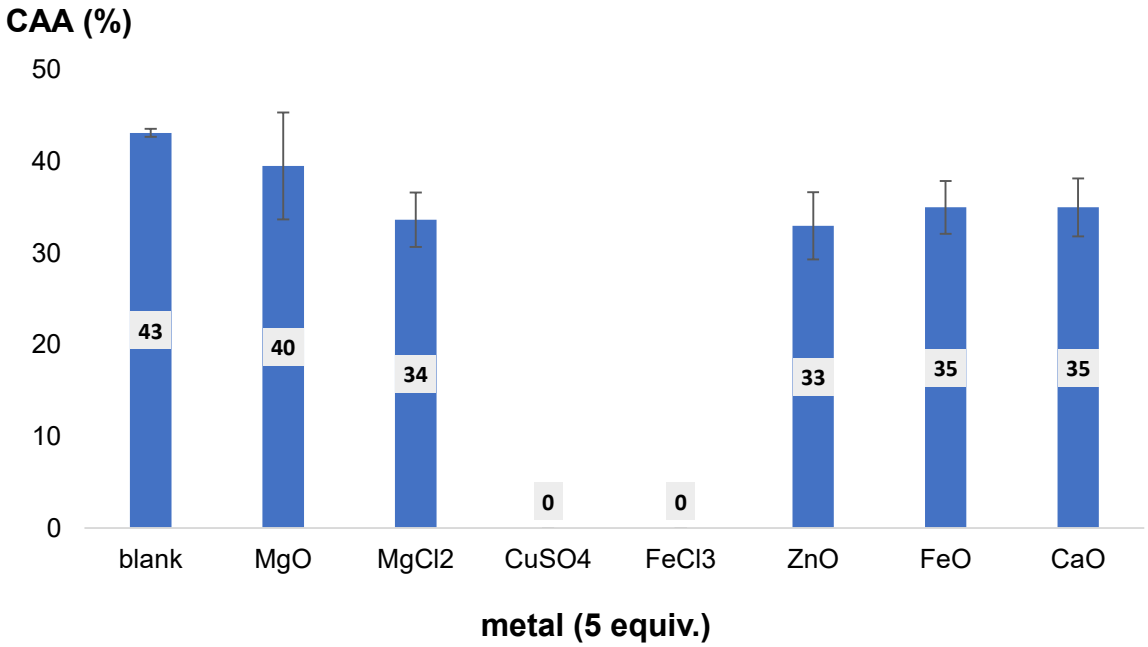


Figure 28. The effect of metals on the carbamoylation reaction between ASP and CAP. Reaction condition: 8.5 mM ASP, 1 equiv. of CAP, 2 equiv. of $NaHCO_3$ salt (pH = 8) and 5 equiv. of metal in 1 mL of MilliQ water, 60 °C for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

As commercial CAP is quite expensive (373 euros/1 g from Sigma Aldrich®), we began the screening reactions using only 1 equiv. of this reagent. However, with the observation of its chemical instability in aqueous solution, we can investigate a higher loading of CAP in the reaction to compensate its degradation/hydrolysis. Therefore, we increased its amount to 2 equiv. and noticed an enhanced conversion of ASP to CAA (Figure 29 left). Another factor that might impact the reaction yield would be the concentration of substrate. Knowing that the solubility of ASP in neutral water is near 40 mM, we next examined a higher concentration of ASP and realized that at 21 mM, the yield of CAA reached a maximimum in the presence of 2 equiv. of CAP (Figure 29 right).

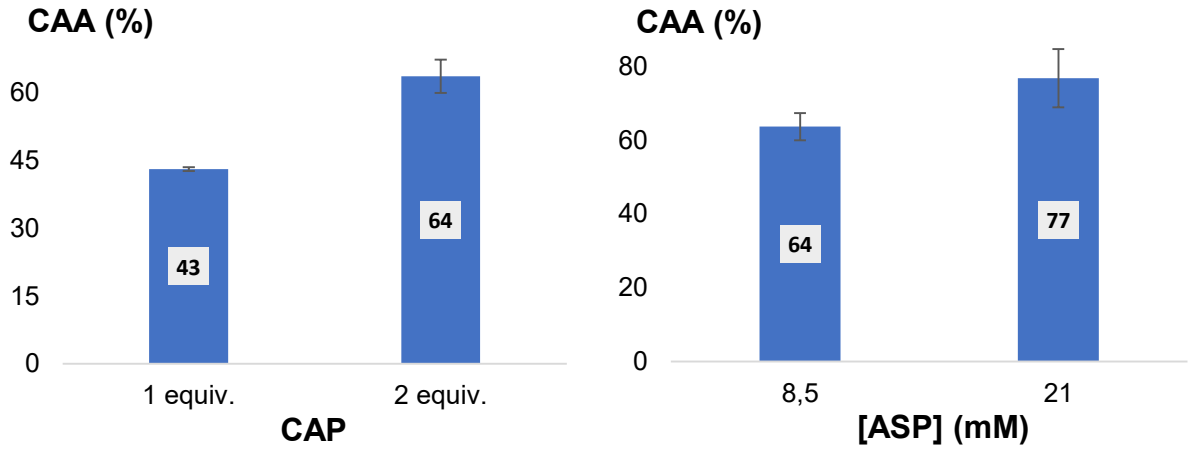


Figure 29. *Left*. The effect of CAP quantity on the carbamoylation reaction. Reaction condition: 8.5 mM ASP, CAP and 2 equiv. of $NaHCO_3$ salt (pH = 8) in 1 mL of MilliQ water, 60 °C for 16 h. *Right*. The effect of substrate concentration on the carbamoylation reaction. Reaction condition: ASP, 2 equiv. of CAP and 2 equiv. of $NaHCO_3$ salt (pH = 8) in 1 mL of MilliQ water, 60 °C for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

The last step was the evaluation of temperature effect on the carbamoylation reaction (Figure 30). We started from mild heating at 60 ºC and tried to avoid high temperature in this project, because the next stage will involve ribose-5-phosphate, which decomposes rapidly at high temperature.[148] We have screened four temperature values, from 0 to 60 ºC, in combination with the previously found conditions. The result at 0 ºC and 20 ºC showed either no reaction or very low reactivity occurred. Whilst above 40 ºC, the reaction rate started accelerating. The yield got nearly doubled when the temperature jumped from 40 ºC to 60 ºC.

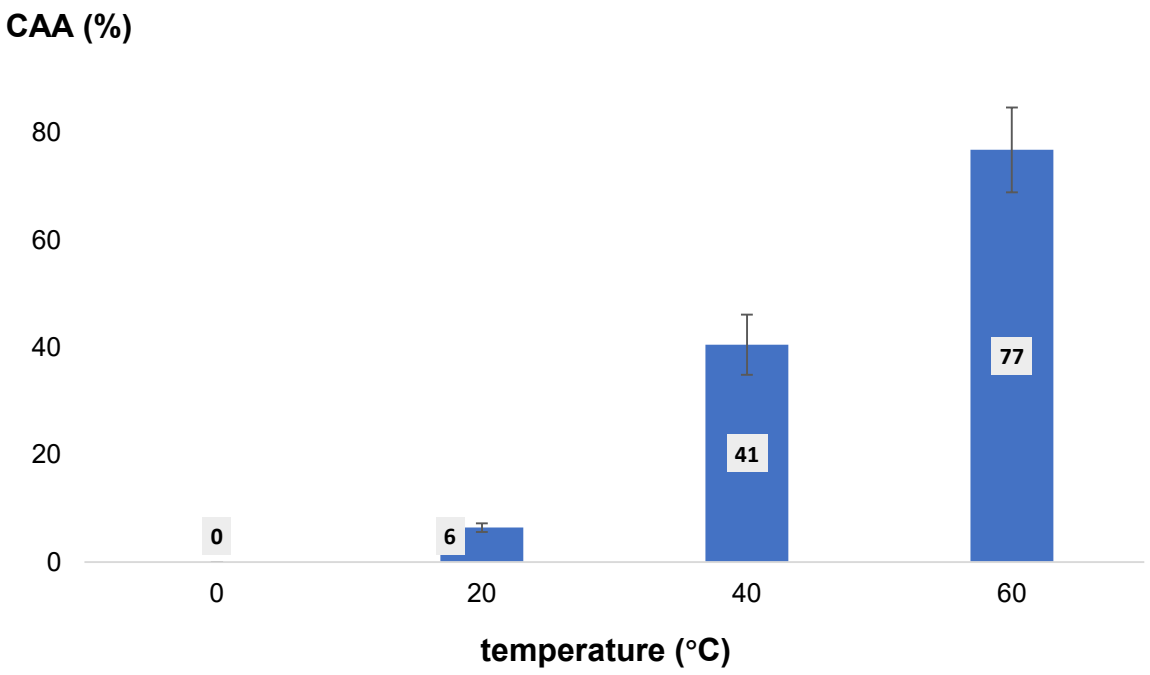


Figure 30. The temperature effect on the carbamoylation reaction between ASP and CAP. Reaction condition: 21 mM ASP, 2 equiv. of CAP and 2 equiv. of $NaHCO_3$ salt (pH = 8) in 1 mL of MilliQ water for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

In the end, we obtained the optimal reaction condition for the carbamoylation using a 2:1 ratio of CAP to ASP at pH 8 with 2 equiv. of $NaHCO_3$ at 60 ºC for 16 h (Figure 31). This condition will be further used to achieve the one-pot synthesis of pyrimidine nucleobases.

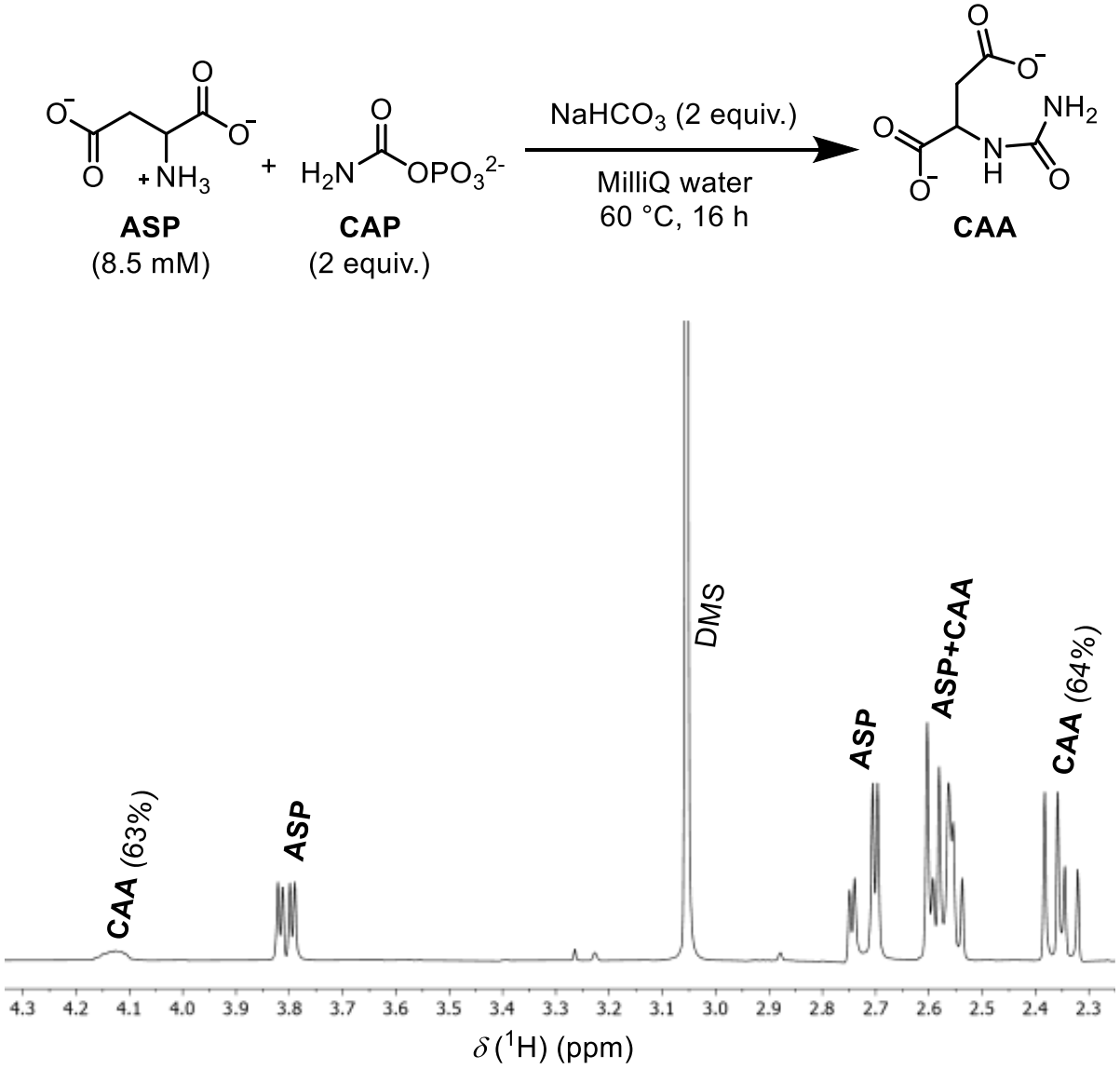


Figure 31. ^{1}H NMR spectrum of the products of reaction of aspartate ASP and carbamoyl phosphate CAP using optimal reaction conditions. δ, chemical shift. Yields were determined by ^{1}H NMR with water suppression techniques, integrating against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve. (Figure reproduced from *Yi, 2022*[149])

3.1.3 Competitive carbamoylation between different amino acids

The above-mentioned work can be viewed as a first experimental step to investigate the chemical reactivity between CAP and ASP in the nonenzymatic context. However, while prebiotic chemistry likely made ASP from reductive amination of oxaloacetate, it also produced other amino acids from their ketoacids such as glycine, alanine and glutamate that are derived from the glyoxlate cyle or rTCA cycle. Currently, there's no mechanism to separate ASP from other amino acids nonenzymatically. Thus, the implications of these amino acids on the carbamoylation of ASP must be addressed.

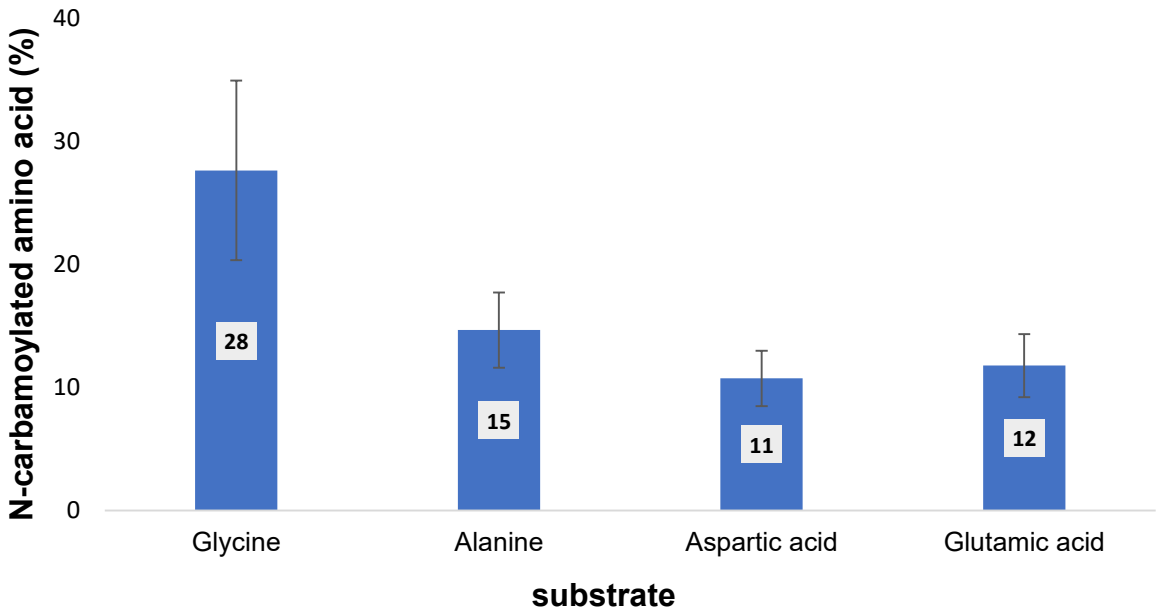


Figure 32. Reactivity of four different amino acids (glycine, alanine, ASP, glutamic acid) with CAP. Products are the *N*-carbamoylated amino acids. Reaction condition: 8.5 mM of each amino acid, 8.5 mM of CAP and 17 mM of $NaHCO_3$ salt (pH = 8) in 1 mL of MilliQ water at 60 ºC for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

Preliminary experiments indicate that the reaction of ASP with CAP is not selective over the other three amino acids, as we anticipated (Figure 32). However, we must agree that it is not necessary for the pathway to be favoured or to be relevant, but only to exist, in order for it to be later selected by evolution. The expectation that the biological pathway may have been part of a larger reaction network does not preclude other redundant pathways from having co-existed with it. But that point aside, there are multiple reasons why selectivity might not matter so

much. The first reason is that we do not know how many other amino acids would have been produced by the larger system at the time this pathway first emerged. Pyrimidine biosynthesis requires only aspartate as a structural building block, which comes from oxaloacetate. Purine biosynthesis requires only glycine for its carbon backbone, which comes from glyoxylate. The ribose-5-phosphate required to make both of them could come from pyruvate or from oxaloacetate. We do not know which pathway came first or whether both occurred simultaneously. Ongoing work in our lab (unpublished) suggests that the reductive amination of the various ketoacids in core metabolism is selective during competition experiments and that the production of glycine and ASP are favored. Therefore, even if more than one of glyoxylate/pyruvate/oxaloacetate/ketoglutarate were present, it is possible that only one (ASP) or two (ASP and glycine) amino acids were around at this stage of the game. A second reason is that, as published by Springsteen and co-workers[141], there are nonenzymatic processes which convert *N*-carbamoyl glycine to ORO. Therefore, even if both glycine and ASP were around and became carbamoylated, this would not necessarily be detrimental to the system's production of ORO. The third reason why this is not necessarily a problem is that hydrolysis of the *N*-carbamoyl amino acids can also occur, so just because some is formed, does not mean the other amino acids cannot exist in their free forms.

To conclude, the potential implications of competitive carbamoylations between different amino acids to a self-organized prebiotic chemistry depends on many factors which are at present difficult to constrain.

3.2 Dehydrative cyclization

In the *de novo* pyrimidine biosynthesis, the cyclization of CAA to DHO is independent from ATP and is catalyzed by the Zn-based metalloenzyme *dihydroorotase* (Figure 33, step a). However, it should be noted that the cyclization reaction is endergonic and at pH 7.4 at 37 ºC the enzyme catalyzed process favors CAA over DHO by a factor of 16.6 at equilibrium.[150] Thus, low yields of DHO are to be expected under nonenzymatic conditions. Nevertheless, one species of bacteria contains *5-carboxymethylhydantoinase* which catalyzes the equilibrium between CAA and HAA (Figure 33, step b), a molecule which was the major intermediate in the total synthesis of ORO reported by Nyc and Mitchell.[138, 139] Therefore, this five-membered ring intermediate may have been a great alternative biological process at one point that took on a secondary role compared to the *de novo* path via formation of DHO.

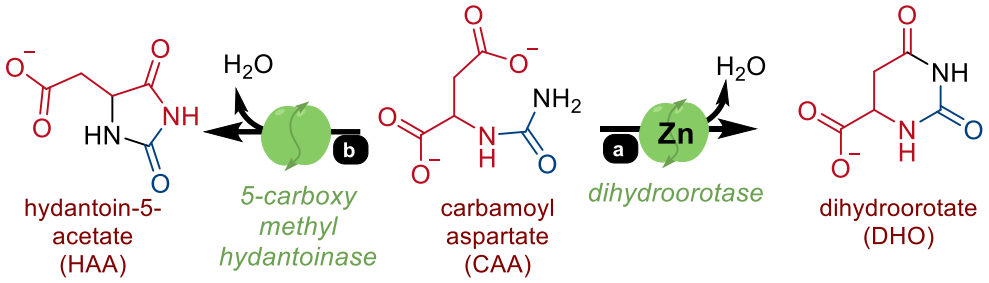


Figure 33. Dehydrative cyclization of CAA: *right*: in the *de novo* pyrimidine synthesis towards DHO; *left*: towards enzymatic synthesis of HAA.

3.2.1 Optimization

Encouraged by the nonenzyamtic carbamoylation results, we began our experimental exploration on the nonenzymatic cyclization using conditions at 60 ºC for 16 h. The first evaluation was the pH effect (from pH 1-8, adjusted with HCl, $NaHCO_3$ or phosphate buffers) in the absence of metallic ions/catalysts. We noticed that there's no DHO formation observed, only few amount of HAA was obtained at pH <2. This result is similar to that observed in a non-prebiotic context during the total synthesis of ORO.[139] Above that pH value, none of the cyclized product was observed.

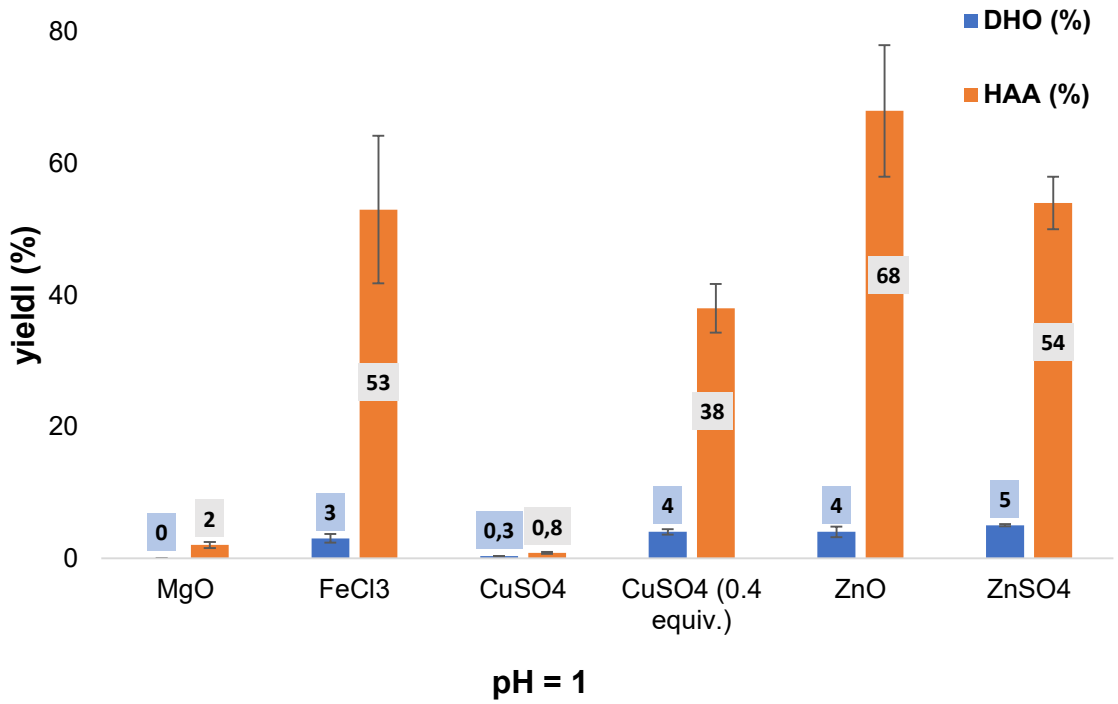


Figure 34. The effect of metals on the cyclization reaction of CAA to DHO/HAA at pH = 1. Reaction condition: 21 mM of CAA, 5 equiv. of metal in 1 mL of 0.1 M HCl at 60 ºC for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

A new screen was therefore carried out across the same range of conditions but in the presence of metal salts and oxides (Figure 34), which were also used for screening in the step of carbamoylation (alkalins, transition metals). We firstly essayed the reactions at pH 1 using HCl. Alkalin metals did not present any catalytic effect. Zn^{2+}, Cu^{2+} and Fe^{3+} produced a small amount of DHO, but the chemistry was selective over HAA. Interestingly, Zn-based metal salt/oxide afforded the best conversion of substrate CAA to cyclized products, which was mimicking the

metalloenzyme in this step. Fe(III) salt was doing a similar job. However, the addition of the same quantity of Cu(II) salt inhibited the cyclization reaction. Instead, a promoting effect was observed when the quantity of Cu(II) salt is decreased nearly 12 times (0.4 equiv.). This promoting effect did not show up in the case of other metal salts/oxide while decreasing their loading.

In line with the carbamoylation step, we next repeated the same set of screening with minimum amount of metal to promote CAA cyclization at pH 8, adjusted with the addition of $NaHCO_3$ solid (Figure 35). The overall conversion of the CAA to HAA is therefore decreased with the addition of base. The yield of DHO is less touched. However, at pH 8, the selective cyclization towards DHO is present with the use of Cu(II) salt, which was not the case for other metals.

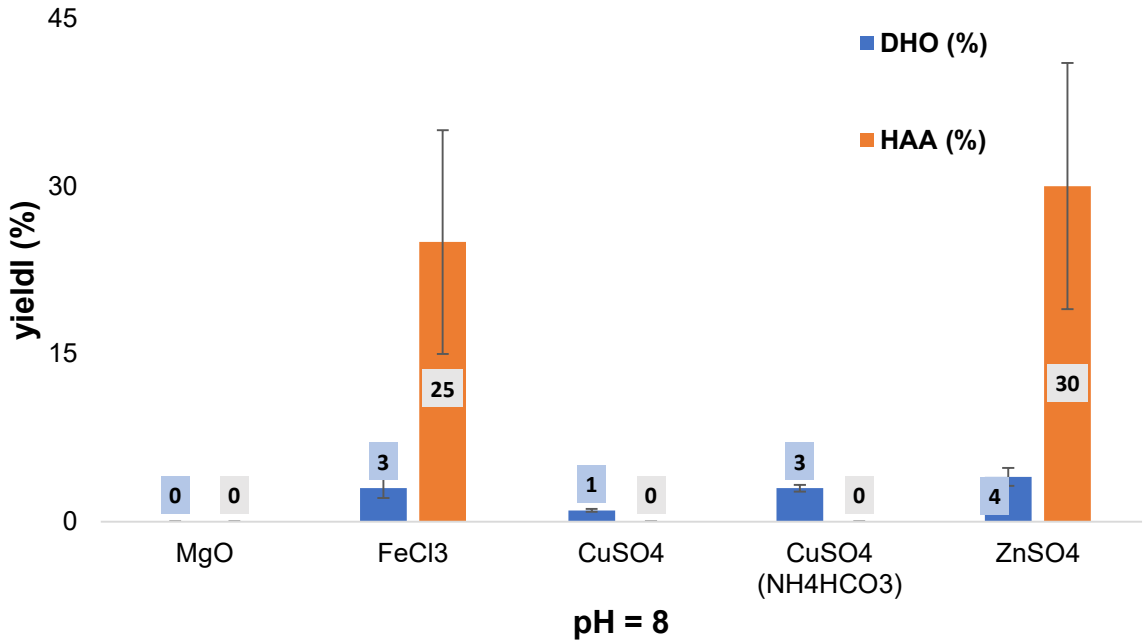


Figure 35. The effect of metals on the cyclization reaction of CAA to DHO/HAA at pH = 8. Reaction condition: 21 mM of CAA, 0.4 equiv. of metal, in 1 mL MilliQ water at 60 ºC for 16 h. $NaHCO_3$ or NH_4HCO_3 solid was used to adjust pH. Yields are the average of at least three independent runs. Error bars correspond to standard deviation

3.2.2 Cyclization using activating agents

Although the result of DHO formation in aqueous solution is improved in our experiment, compared to < 1% yield under direct UV irradiation[138]. The window of opportunity to look for possible activating reagents to increase the CAA cyclization towards DHO, is still open for us. One strategy was related to the activation chemistry of the carboxylic groups in CAA. Inspired by the results of Jencks and co-workers[151], the active phosphate species would act as promoter to react with carboxylates. The process produces the corresponding carboxy phosphate, which therefore facilitates the intramolecular nucleophilic substitution reaction by the urea moiety of CAA. The mechanism of this activation could be suggested as below (Figure 36):

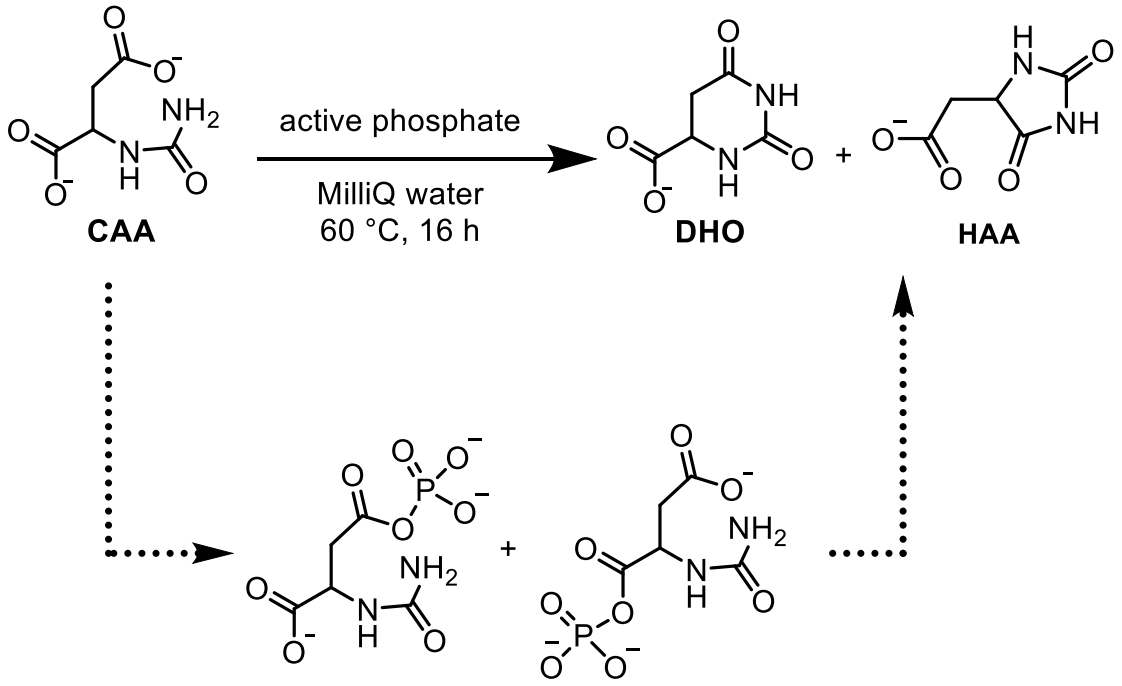


Figure 36. The plausible mechanism of active phosphate promoted cyclization of CAA via formation of carboxy phosphate species.

We have selected among various phosphate species, including the list of inorganic phosphate presented in chapter 3.1.1. Some organic phosphates were also tested, such as acetyl phosphate, CAP, $POCl_3$, succinyl phosphate and pyridinium phosphate[152] that are considered as strong phosphorylating agents. Among these reagents, only $POCl_3$ was able to trigger the cyclization of CAA (Figure 37). As suggested by Jencks[153], the combination of active phosphate with divalent metal salts such as Mg(II) would increase the phosphate-transfer reaction rate. So we incorporated the use of Zn(II) and Mg(II) salts and obtained ~100 % conversion of CAA to cyclized products DHO/HAA at 60 °C. The conversion became lower when temperature decreased to ambient temperature. Interestingly, the yield of DHO (15% in reaction 2, Figure

37) is much higher than the cyclization reaction at pH 1 in HCl with Zn salt (5%). Therefore, the $POCl_3$ must be activating agent instead of acidifier in aqueous medium. However, the mechanism of how $POCl_3$ reacts to the system remained unclear and will need more investigations on it. Further studies are currently ongoing in the Moran lab in order to provide more mechanistic insights and to find a plausible prebiotic analogue of $POCl_3$ that could lead to a comparable outcome.

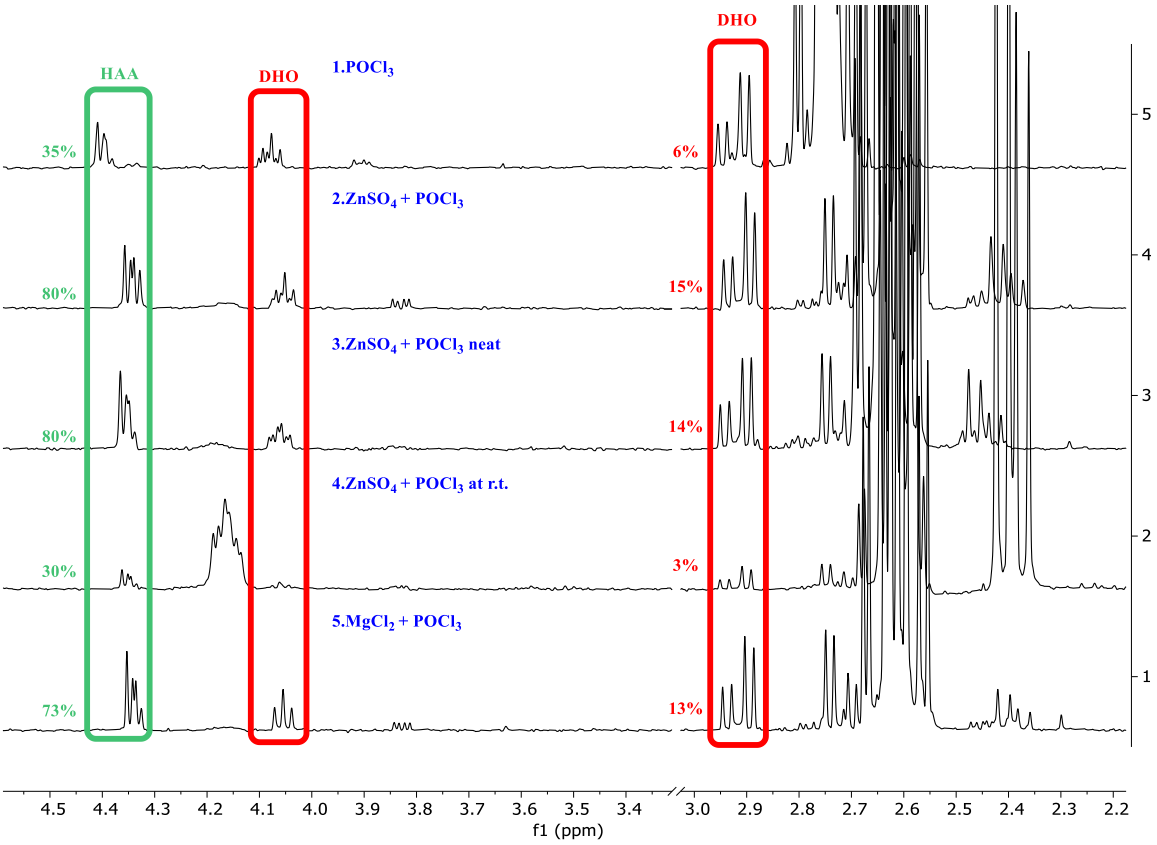


Figure 37. ^{1}H NMR spectrum of active phosphate promoted cyclization of CAA. Reaction conditions: 21 mM of CAA, 5 equiv. of phosphate species, 5 equiv. of metal in 1 mL MilliQ water at 60 ºC for 16 h. Yields were determined by ^{1}H NMR with water suppression techniques, integrating against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve.

3.3 Oxidation

The oxidation of DHO to ORO in the *de novo* biosynthesis uses NAD^+ as oxidant. A nonenzymatic version of this reaction had only been observed in < 1% under light irradiation[138], or in higher yield using conditions such as Br_2 in glacial acetic acid[139] which certainly has no prebiotic relevance. In our project, we tried to add complementary experiments to improve the photochemical oxidation using milder wavelengths, also to uncover the role of metallic oxidants that are able to convert DHO to ORO.

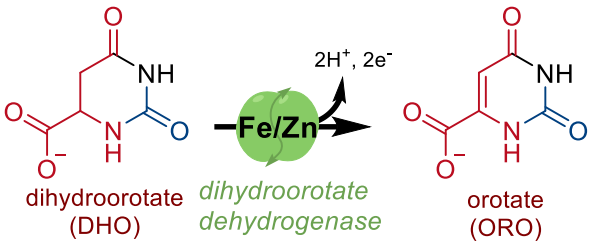


Figure 38. DHO oxidation to ORO in the *de novo* pyrimidine synthesis.

3.3.1 Photochemical oxidation

Inspired by Lu and co-worker's photo-reduction of bicarbonate by elemental sulfur (S^0)[154], we suggested that the creation of photoelectrons during the irradiation process might provide important implications for prebiotic synthesis of ORO from DHO (Figure 39).

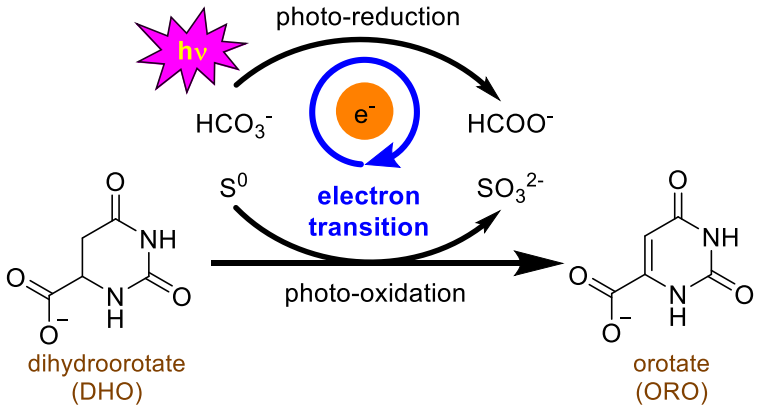


Figure 39. Plausible mechanism for photochemical oxidation of DHO to ORO in the presence of elemental sulfur and bicarbonate.

Exposing DHO to S^0 in degassed saturated $NaHCO_3$ solution resulted in its oxidation to 14% of ORO and decarboxylation to 1% of uracil (URA) at room temperature for 16 h under UV irradiation (UV A+B+C), among which 2% of ORO and the formation of URA were originated from background reaction when sulfur was absent (Figure 40a). Dihydrouracil (DHU) and formate were also observed from the decarboxylation of DHO and from the reduction of bicarbonate, respectively. At the same time, other reducing sulfur salts, NaHS (sodium hydrosulfide) and $Na_2S_2O_4$ (sodium hydrosulfite), were also examined in the photo-oxidation of DHO to ORO. NaHS did not result in any oxidation. Whereas $Na_2S_2O_4$ produced the highest yield of ORO in saturated $NaHCO_3$ solution, this reaction is identified as the optimal condition for photo-oxidation of DHO (spectral result in Figure 40b). Due to its ambivalent redox properties which exhibit complex pH-dependence, $Na_2S_2O_4$ can act simultaneously as reductant and oxidant.[155] The result using this sulfur salt in the absence of bicarbonate explained very well this special character with the production of 15% of ORO.

Control experiments were carried out with the use of $Na_2S_2O_4$, instead of UV light, and the reaction was performed under thermal conditions at 60 ºC for 16 h. However, only trace amount of ORO (< 1%) was observed in this reaction.

Another question is the origin of URA in this reaction, and whether it is formed by decarboxylation of ORO, or from the oxidation of DHU. Experimental results indicated a route via decarboxylation instead of oxidation of DHU, because ORO was decarboxylated to URA under the optimal conditions, whilst no oxidation was observed with DHU.

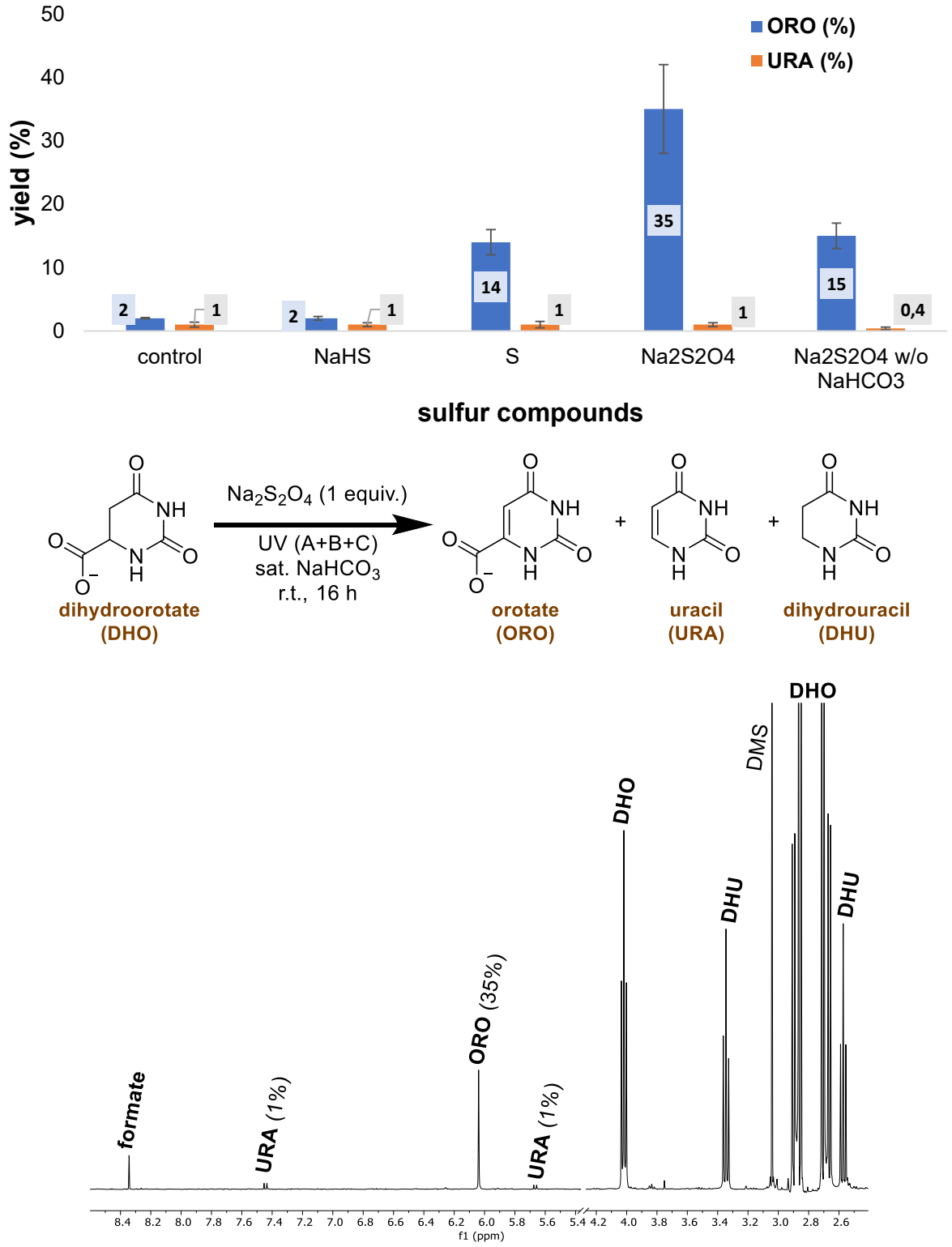


Figure 40. *a*. The effect of using different sulfur salts in the oxidation of DHO to ORO. Reaction condition: 8.5 mM of DHO, 1 equiv. of sulfur salt, in 1 mL of saturate $NaHCO_3$ solution, the mixture was degassed and flushed with argon, stirred at r.t. for 16 h with UV (A+B+C) irradiation. *b.* The optimized photo-oxidation and the ^1H NMR result. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

3.3.2 Thermal oxidation

3.3.2.1 DHO oxidation

The oxidation of DHO was never studied using earth abundant metals. Here, we first performed a screen with different oxidants at 60 ºC, including transition metal salts/oxides, and some other inorganic oxidants such as peroxides, nitrates, iodine, chloro-based salts and sulfur-based salts (Figure 41). Two optimal oxidants were found to be H_2O_2 and MnO_2, the prebiotic plausibility of which was recently discussed in the literature.[156, 157] We have also studied their reactivity under different pH values (from pH 1-14), H_2O_2 worked best at pH 9 (34%) when saturate $NaHCO_3$ was used as solvent. In the case of MnO_2 as oxidant, at pH 4 without any base/buffer, ORO was obtained in 18% as the best yield together with 2% of URA after 16 h. Hydrolysis of DHO back to CAA was also observed when pH >8. At extreme pH 14 when 0.1 M NaOH was used as solvent, DHO was nearly fully hydrolyzed and only <1% of ORO was formed in both oxidations.

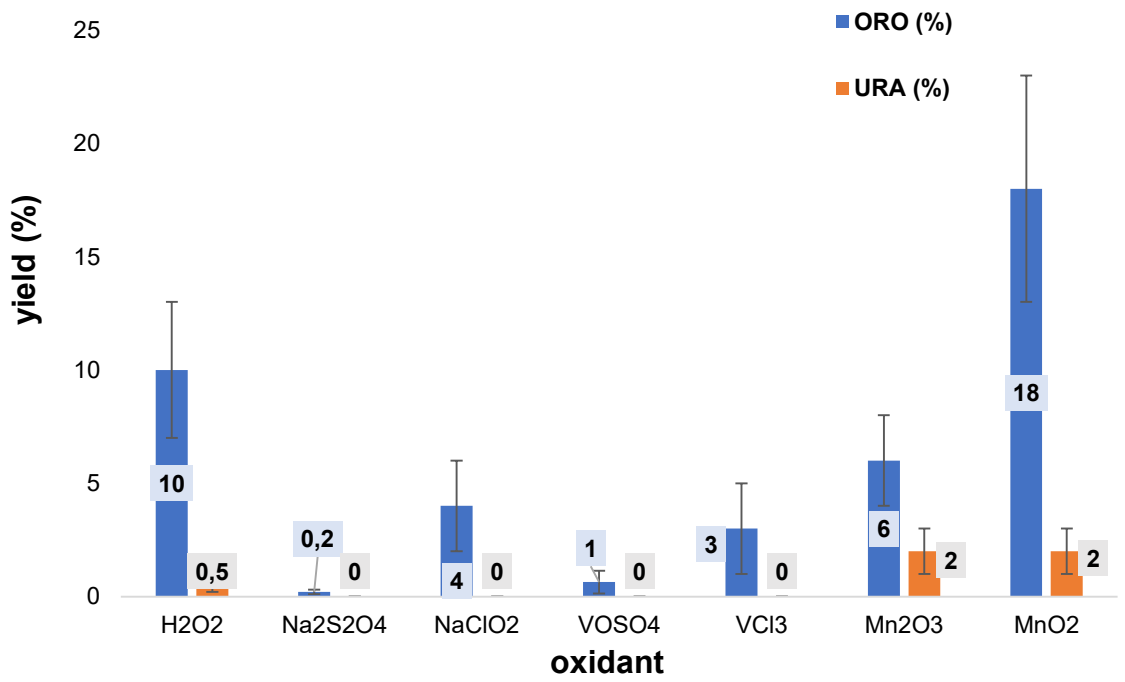


Figure 41. The effect of oxidants on the oxidation of DHO at pH = 4. Reaction condition: 8.5 mM of DHO, 5 equiv. of oxidant (20 equiv. of H_2O_2) in 1 mL of MilliQ water at 60 ºC for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

Next, we studied the influence of metal additives on the oxidation. Divalent metal salts like alkalines, Fe(II), Co(II), Ni(II), Mn(II), Cu(II) and even Pd(II) were screened together with two optimal oxidants. It was found that the presence of metal ions inhibited the H_2O_2 promoted oxidation of DHO. For example, the addition of Cu(II) ion reduced the ORO yield to 3%, but with slightly higher yield of URA (4%) at pH 9 with heating to 60 ºC for 16 h. However, in the case of the MnO_2 promoted oxidation at 60 ºC without pH adjustment, Cu(II) additive enhanced the oxidative outcome towards ORO by a factor of 2.8, whilst other divalent metals did not exhibit the same enhancement effect (Figure 42). The decarboxylation to URA was not changed.

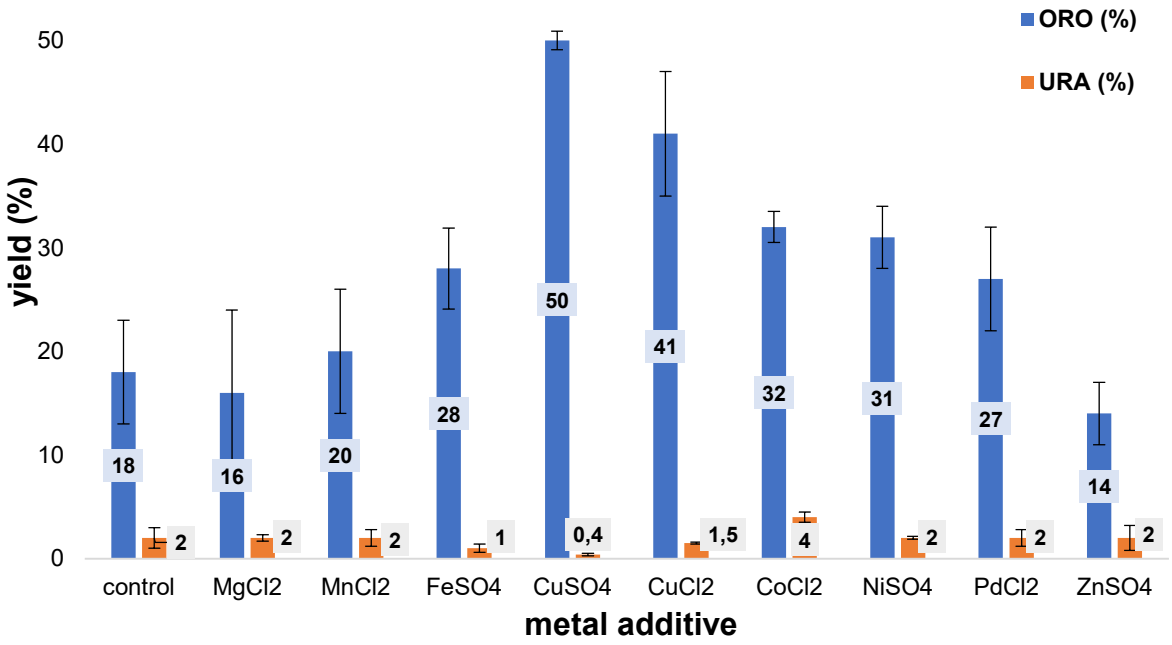


Figure 42. The effect of metal additives on the MnO_2 oxidation of DHO at pH = 4. Reaction condition: 8.5 mM of DHO, 5 equiv. of MnO_2, 0.4 equiv. of divalent metal salt in 1 mL of MilliQ water at 60 ºC for 16 h. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

Cu(II) can sometimes be considered as problematic in the origin of life field due to its cytotoxicity. Despite this, Cu(II) is routinely used as promising catalyst in a number of different prebiotic contexts.[106, 158, 159] In response to the question about the feasibility of changing Cu(II) concentration in the oxidation, we have examined three different concentrations of Cu(II) salt and the results the reactions yielded over longer timescales (Figure 43). At low Cu(II) concentration, oxidative reactions could be improved by heating with longer reaction time. However, 0.4 equiv. of $CuSO_4$ gave the highest yield of ORO after 16 h. In view of the concentration of DHO, increasing it to 21 mM did not display any enhancing effects.

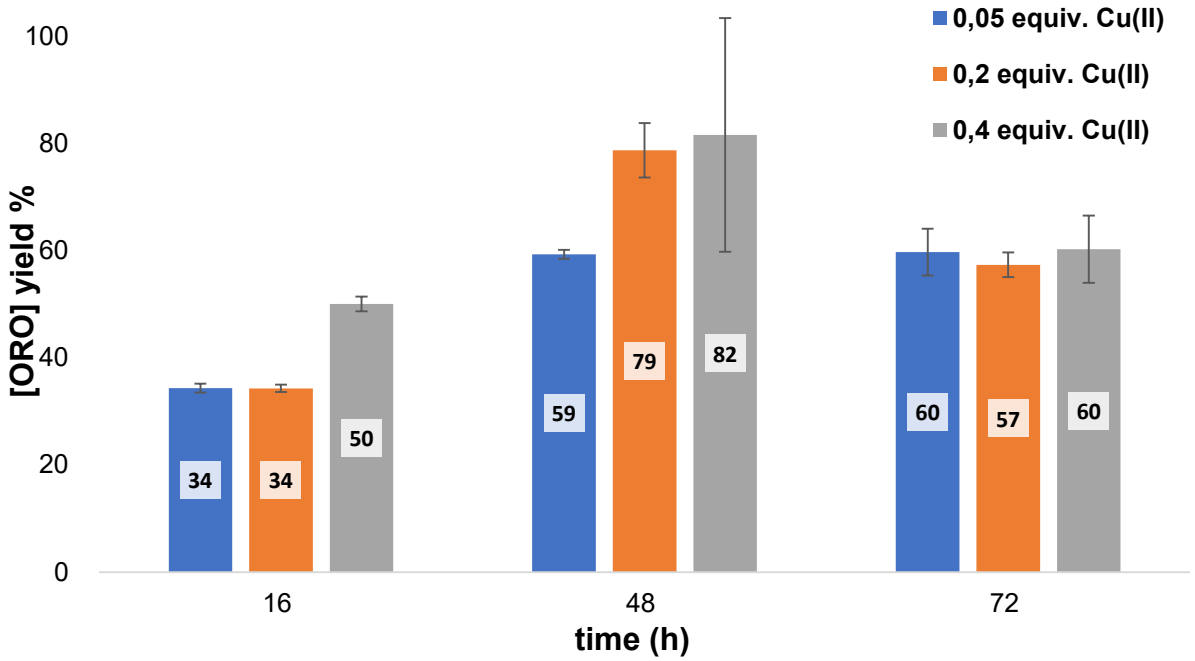


Figure 43. Effect of $CuSO_4$ concentration in the oxidation of DHO to ORO at different times. Reaction condition: 8.5 mM of DHO in 1 mL of water, 5 equiv. of MnO_2, the reaction was heated at 60 °C. Yields are quantified by ^1H qNMR on Bruker 400 MHz at ambient temperature. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

The next parameter to evaluate the oxidation reaction was the temperature effect. Thus, we carried out the MnO_2 promoted oxidation in addition of 0.4 equiv. of $CuSO_4$ at different temperature ranging from 0-60 °C for 16 h (Figure 44 left). The best temperature for the oxidation was identified at 60 °C. The quantity of MnO_2 was also essayed for the DHO oxidation to ORO using the optimal temperature and Cu(II) additive (Figure 44 right). The oxidation was found to be more effective when higher quantity of oxidant was used.

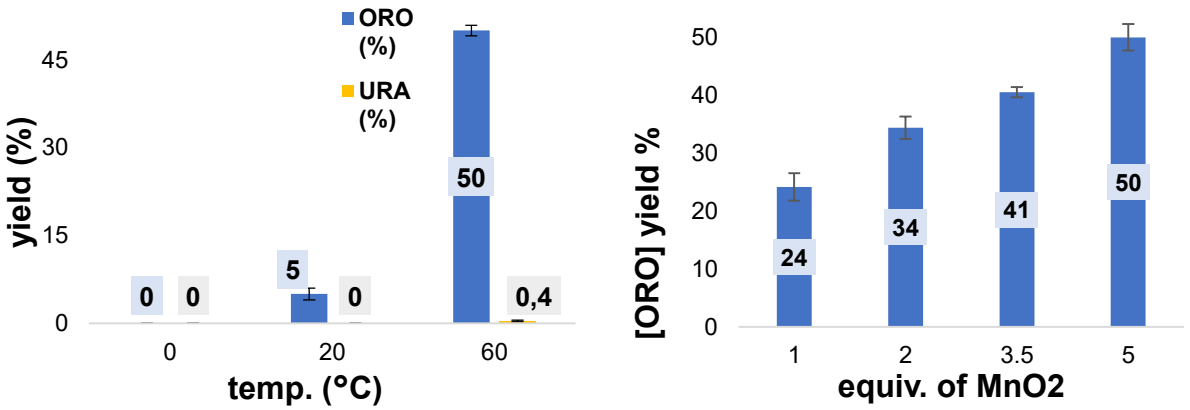


Figure 44. *left*. Effect of temperature in the oxidation of DHO to ORO. Reaction condition: 8.5 mM of DHO in 1 mL of water, 5 equiv. of MnO_2, 0.4 equiv. of $CuSO_4$ for 16 h. *right*. Effect of MnO_2 in the oxidation of DHO to ORO. Reaction condition: 8.5 mM of DHO in 1 mL of water, 0.4 equiv. of $CuSO_4$ at 60 °C for 16 h. Yields are quantified by ^{1}H qNMR on Bruker 400 MHz at ambient temperature. Yields are the average of at least three independent runs. Error bars correspond to standard deviation.

The use of MnO_2 in combination with $CuSO_4$ was therefore chosen as the best oxidant couple at 60 °C for 16 h (spectral confirmation in Figure 45). The experiments uncovered that Cu(II) is capable of influencing both the cyclization of CAA and oxidation of DHO on the non enzymatic version of the biological path to ORO.

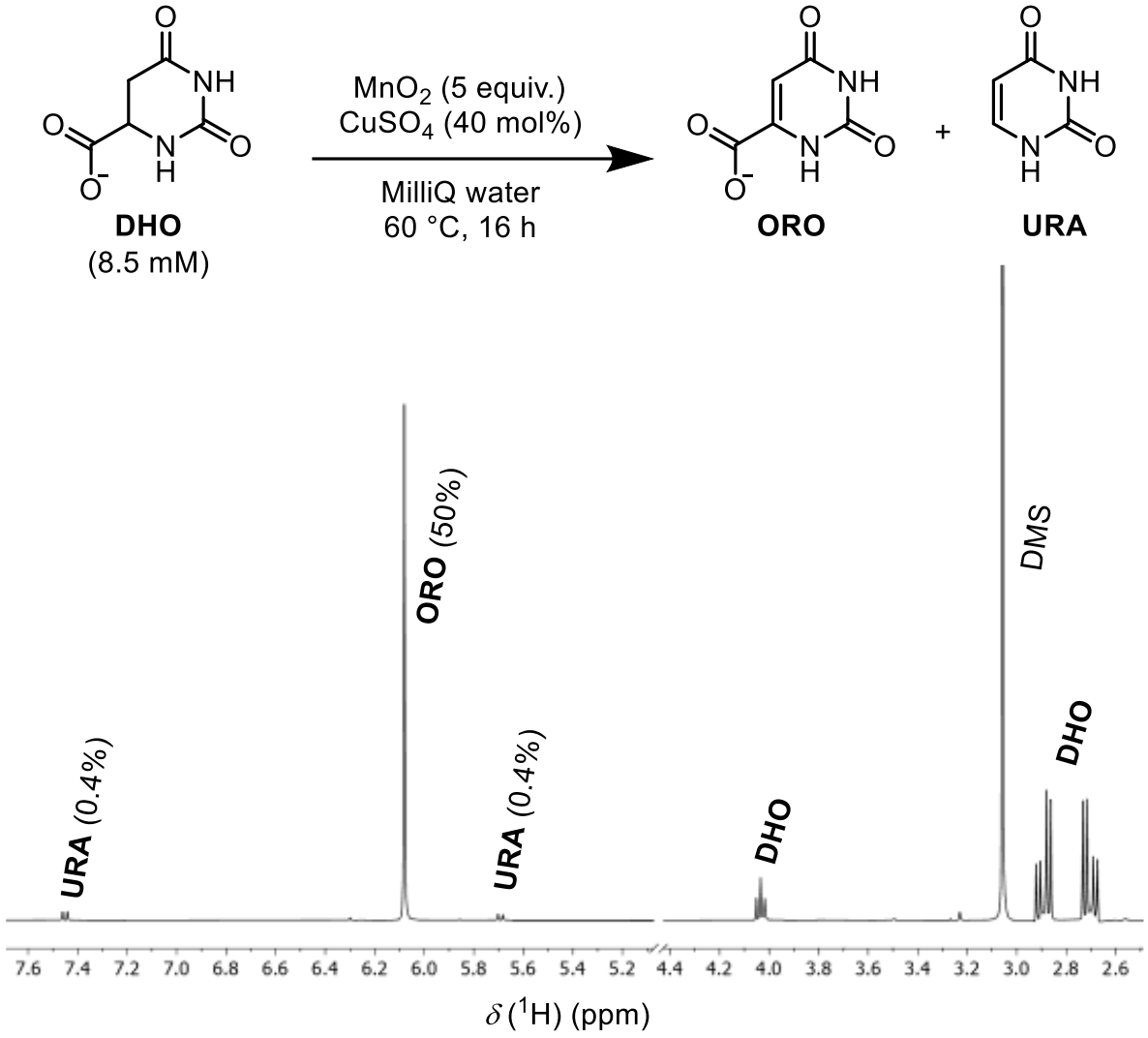


Figure 45. ^{1}H NMR spectrum of the thermal oxidation of DHO using optimal reaction conditions. δ, chemical shift. Yields were determined by ^{1}H NMR with water suppression techniques, integrating against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve. (Figure reproduced from *Yi, 2022*[149])

3.3.2.2 HAA oxidation

Given that the cyclization of CAA to DHO occurred in highest yield at lower pH, where it is accompanied by HAA as major product of the cyclization. Thus, we wondered if HAA might also be converted to ORO under similar oxidative conditions. Could HAA be oxidized to 5-carboxymethylidine hydantoin (see structure in Figure 22b) under plausibly prebiotic conditions, and could it then undergo a known rearrangement, as described by Mitchell[139], under alkaline aqueous conditions to give ORO?

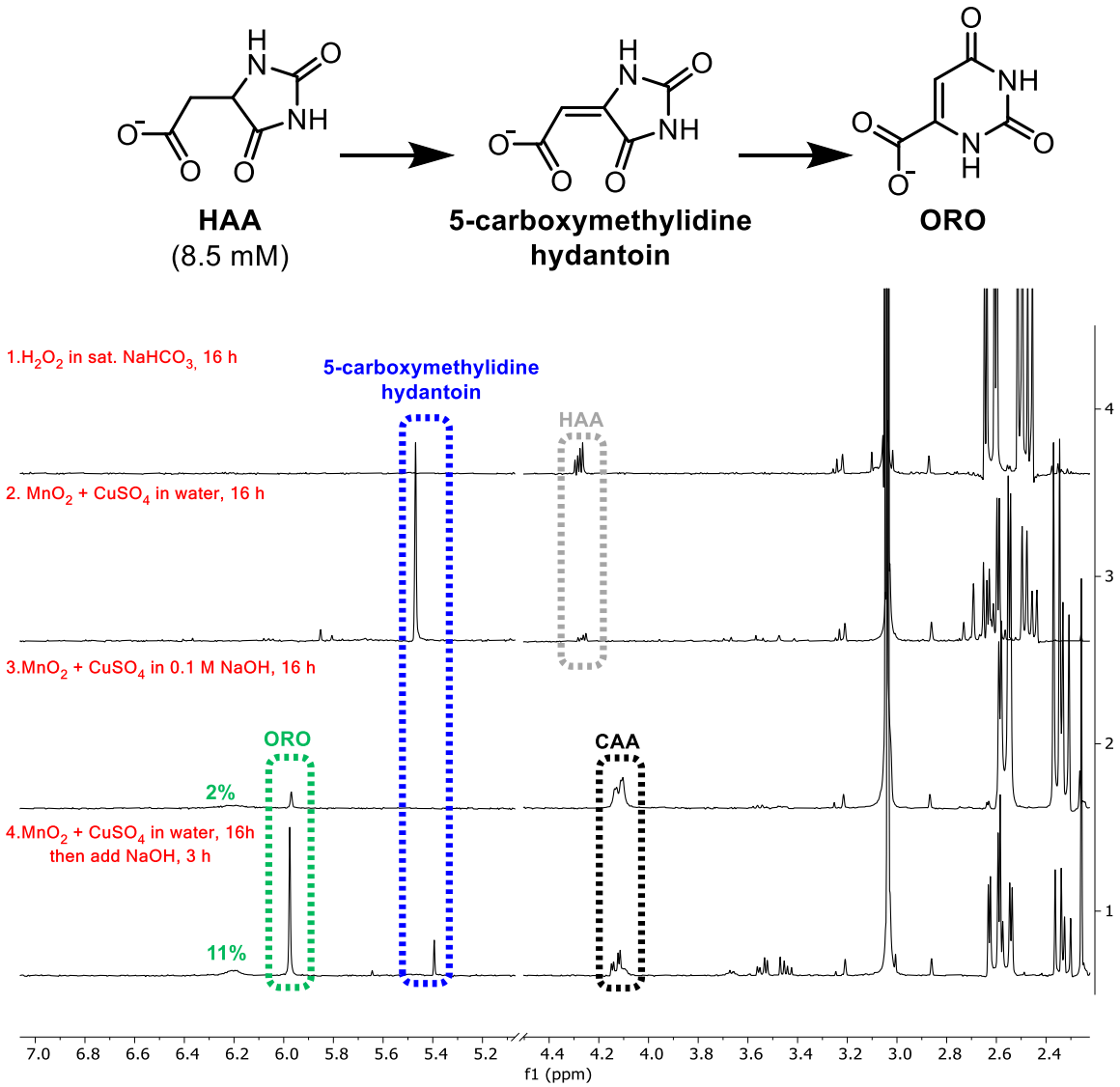


Figure 46. ^{1}H NMR result of HAA oxidation. Reaction conditions: 8.5 mM of HAA, 5 equiv. of MnO_2 (20 equiv. of H_2O_2), 0.4 equiv. of $CuSO_4$ (6 equiv. of NaOH in reaction 4) in 1 mL solvent at 60 ºC for 16 h. Yields were determined by ^{1}H NMR with water suppression techniques, integrating against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve.

We first exposed HAA to H_2O_2 but it did not result in any oxidation under the conditions we tested (Figure 46, reaction 1). In contrast, the MnO_2 and $CuSO_4$ couple indeed was able to oxidize HAA to 5-carboxymethylidine hydantoin (reaction 2). The use of 0.1 M NaOH as solvent resulted in 2% of ORO (reaction 3). The consumption of 5-carboxymethylidine hydantoin and increased formation of ORO (11%) was observed while NaOH was added to the reaction 2 after 16 h and was kept at 60 °C for another 3 h (reaction 4). This increasing yield might be explained by the fact that the hydrolysis of HAA back to CAA under alkaline pH occurs faster than its oxidation, but once 5-carboxymethylidine hydantoin is produced in water solvent, the addition of NaOH will hydrolyze only HAA and provoke the rearrangement of 5-carboxymethylidine hydantoin to ORO.

Thus, the adventitious conversion of CAA to HAA in its cyclization reaction is not necessarily a dead end, as it can can be funneled towards the same end-product, ORO.

3.4 Synthesis of pyrimidine nucleobases in one-pot

3.4.1 Cyclization and oxidation in one-pot

In order to assess the feasibility of a non-photochemical one-pot process mimicking pyrimidine nucleobase biosynthesis, we first attempted to combine the dehydrative cyclization and oxidation steps. Heating CAA in the presence of all reagents that were found to be optimal for cyclization and oxidation, 0.4 equiv. of Cu(II), 5 equiv. of MnO_2 and 12 equiv. of NaOH in 0.1 M HCl at 60 °C at t = 0 h did not result in any formation of ORO, nor 5-carboxymethylidine hydantoin (reaction 1 in Figure 47). This might be due to the pH neutralization by adding NaOH to the solvent, which is in agreement with the aforementioned influence of pH on the cyclization of CAA to DHO and HAA. Then we tried to add NaOH separately at t = 16 h, ORO was formed in 2% (reaction 2). The presence of MnO_2 in the beginning at t = 0 h could have triggered the decomposition of CAA to ASP, acetate and formate at pH 1. Lately, we isolated MnO_2 from the beginning and allowed the cyclization to complete before the destruction of CAA, so we added the oxidant separately at t = 16 h (reaction 3). Without the implication of NaOH during its oxidation step, the reaction towards 12% of ORO and 0.3% of URA could be achieved within the three-steps process. The intermediate 5-carboxymethylidine hydantoin was also obtained, indicating that the pathways proceeding through DHO and through HAA are both occurring.

The weak point of this one-pot synthesis is therefore human intervention. To our knowledge, there is no paper in the field that selectively produces numerous biological molecules using one-pot intervention-free prebiotic chemistry. Our work is thus only a first step toward our goal of an intervention-free process.

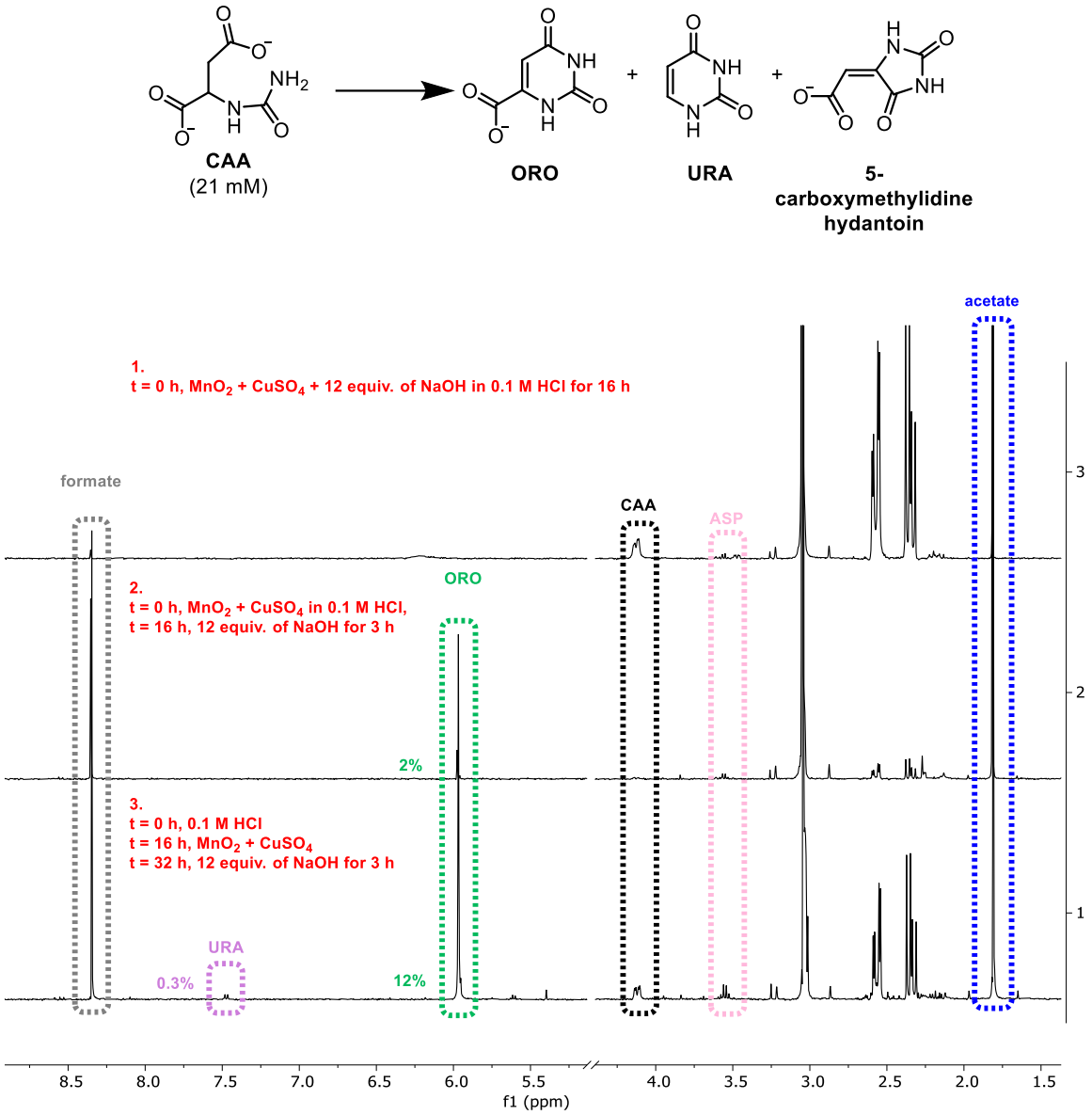


Figure 46. ^{1}H NMR result of one-pot synthesis of ORO starting from CAA. Reaction conditions: 21 mM of CAA, 5 equiv. of MnO_2, 0.4 equiv. of $CuSO_4$, 12 equiv. of NaOH in 1 mL of 0.1 M HCl at 60 ºC. Yields were determined by ^{1}H NMR with water suppression techniques, integrating against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve.

3.4.2 One-pot synthesis of ORO from ASP

In our next step, we tried to incorporate the carbamoylation reaction in our previously developed one-pot synthesis. As we showed previously in Figure 28, the addition of Cu(II) to the mixture of ASP and CAP can kill the carbamoylation. Thus, the Cu(II) must be added after the first step was finished. However, we did find some ORO traces (0.2%) when Cu(II) was added from the beginning. The addition of substantial acid (HCl) and base (NaOH) should also be avoided at t = 0 h because HCl acidifies the pH which was not suitable for the carbamoylation reaction, and NaOH hydrolyzed rapidly the DHO or HAA back to CAA. After testing different orders of addition, we observe ORO in 1% yield along with 14% CAA by heating ASP and CAP (1:2) at 60 ºC at pH 8, followed by the addition of Cu(II) and HCl at t = 16 h, MnO_2 at t = 32 h, and NaOH at t = 48 h after 54 h of total reaction time (Figure 47). Intermediate 5-carboxymethylidine hydantoin is notably observed by ^{1}H NMR in this process. Due to the low detection limit of ^{1}H NMR technique, we quantified our reaction alternatively by a more sensitive method, LC-QTOF-MS, which provides not only the UV absorbance of ORO, but also MS/MS pattern of the molecule. It was also possible to obtain ORO from ASP in one pot by adding all the reagents from the beginning apart from the basification of the reaction at t = 16 h with NaOH, however, the yield decreased to <1%.

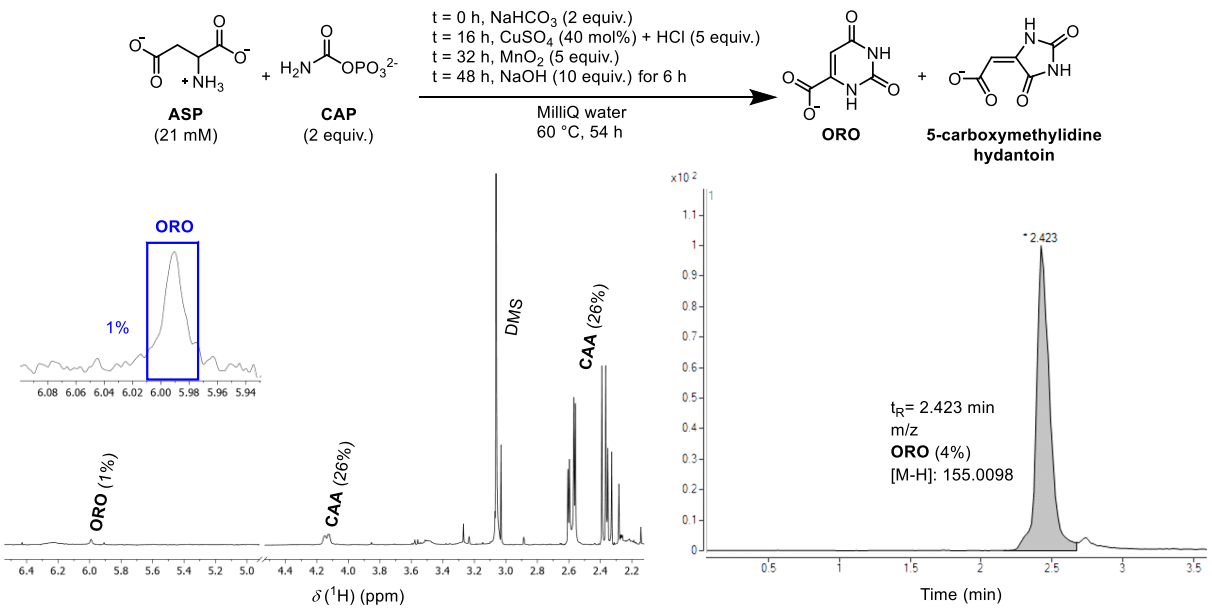


Figure 47. ^{1}H NMR and LC-QTOF-MS of the products of multi-step reaction of ASP and CAP to ORO. NMR yield was integrated against dimethylsulfone (DMS) as an internal standard and correlated with calibration curve. LC-QTOF-MS yield was calculated by comparing the peak area relative to the internal standard (maleic acid) against the calibration curve. [M-H] at m/z = 155.0098 was chosen as characteristic fragment for ORO. (Figure reproduced from *Yi, 2022*[149])

3.5 Conclusion and perspectives

Identifying a non-photochemical nonenzymatic analogue of pyrimidine nucleobase biosynthesis in which some steps are promoted by metals. Some of the steps are inhibited by metals or by unfavorable pH, which explains why the one-pot version of this chemistry necessitated timed addition of metals or timed changes in pH. Although pH gradients are present in some earth environment such as the hydrothermal vents, we feel it is premature to place our multi-step synthesis in a specific geological situation. Whether or not the particular set of conditions shown above is directly relevant to the origin of life, our findings provide experimental evidence to support the hypothesis that the key biological pathway to genetic molecules could have predated enzymes.

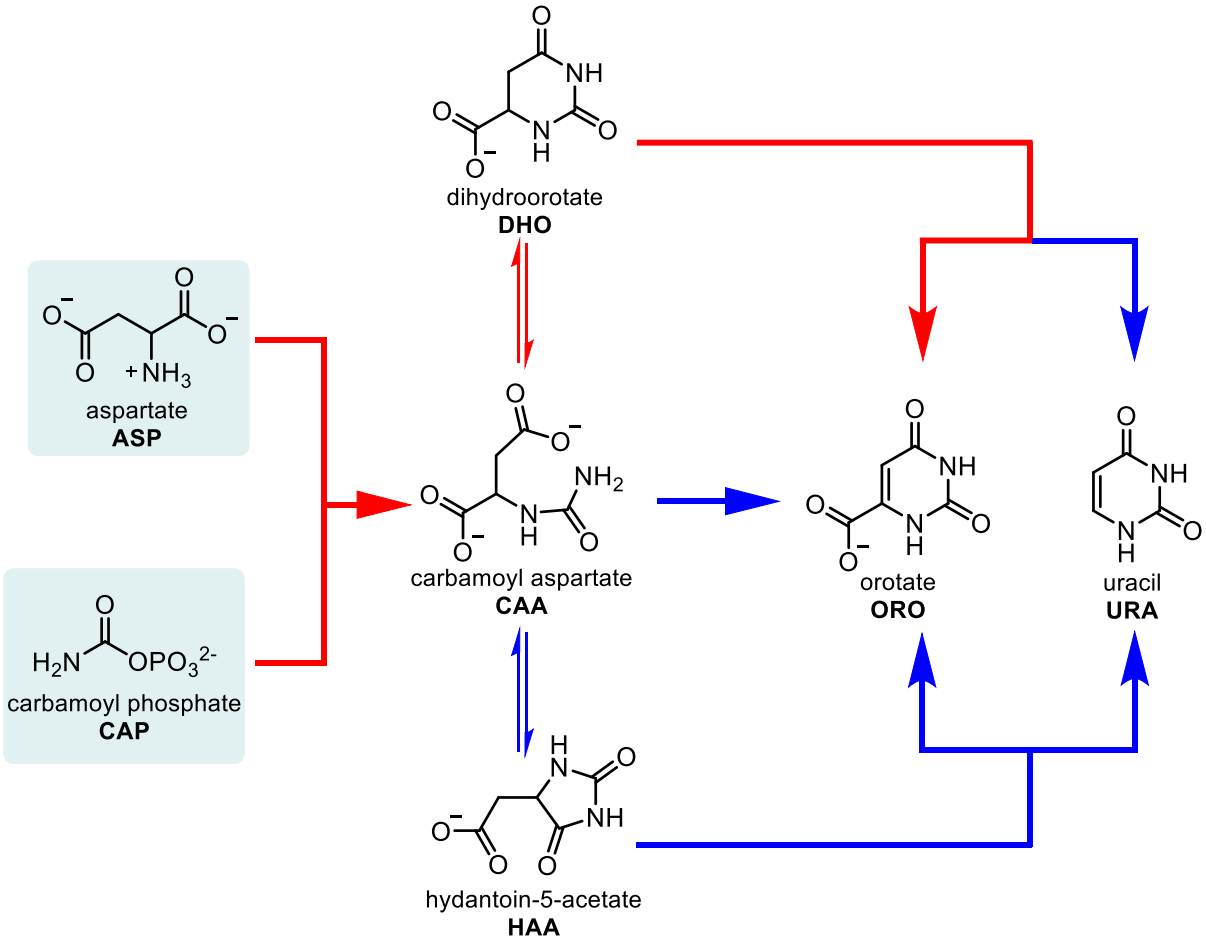


Figure 48. The first three steps of de novo pyrimidine ribonucleotide biosynthesis converting ASP and CAP to ORO, as shown in red arrows. Additional reactions observed in this work are described by blue arrows. (Figure reproduced from *Yi, 2022*[149])

In addition to the nonenzymatic analogue of the *de novo* path, parallel nonenzymatic reactions which do not belong to this specific biological pathway were also found to produce the same biological endpoint: orotate (Figure 48). The conversion of ASP to ORO does not necessarily need metals if potentially prebiotic alternative oxidants can be found. However, among all published results with themal conditions in aqueous solution, only the metal-mediated reaction sequence follows the *de novo* path via formation of DHO. This work indeed offered important evidence in line with the theoretical model that nonenzymatic reactions templated the evolution of core biological pathways.[26] Evolutionary refinement of the redundant branched nonenzymatic reaction network that are found in our work (both red and blue arrows) would logically result in a streamlined linear path like the one found in modern biochemistry (red arrows). In line with this idea, some enzymes still catalyze an equilibrium between CAA, HAA and DHO, even though HAA did not appear to have any biological acivities in the *de novo* pyrimidine biosynthesis. Our work can be considered as a first experimental step toward incorporating genetic molecules into a self-consistent metabolic framework for the origin of life. Future work should should focus on the intervention-free one-pot version of this pathway without timed changes to conditions.

4 Towards nonenzymatic synthesis of pyrimidine ribonucleotides

4.1 Preliminary experiments and perspectives

Even though numerous experiments highlighting the role of nonenzymatic thermal coupling of nucleobases and pentose sugars (see chapter 2.5.2.2), none of them ever mentioned the coupling between ORO and ribose-5-phosphate. However, the formation of orotidine phosphate is the key step in *de novo* pyrimidine ribonucleotide biosynthesis. To fill this gap, we envisaged to achieve the synthesis of OMP in the lab using nonenzymatic chemistry. We have first tried various procedures like thermal "dry-down" conditions or alkaline metal assisted coupling conditions using different kinds of phosphorylating agents, but ORO remained completely unreactive in the reaction (Figure 49).

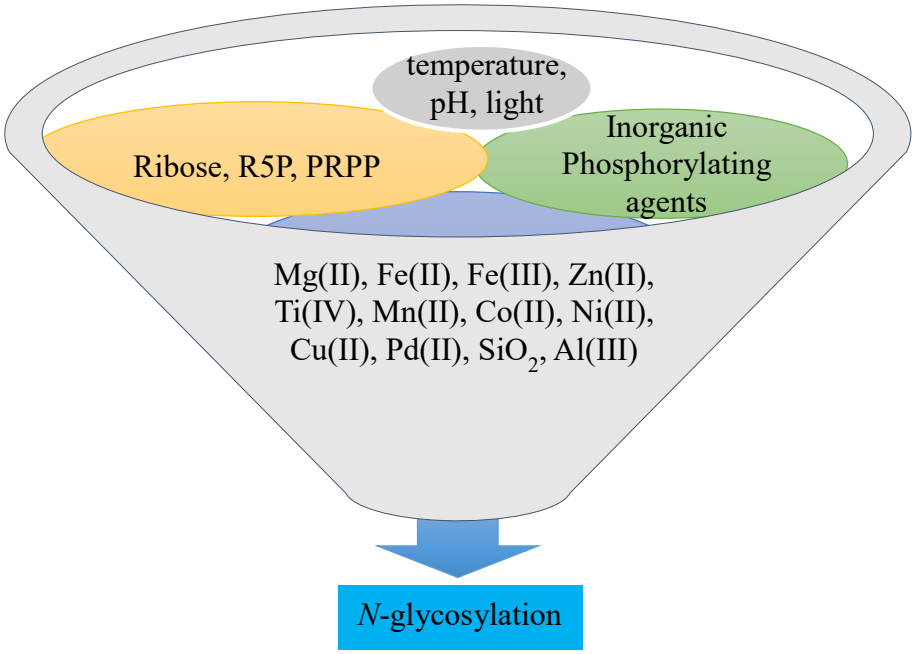


Figure 49. Key parameters for N-glycosylation step between ORO and pentose, each combination was screened under different conditions.

Chemical activating agents and directing groups were then engaged in our research. Based on some of Sutherland's work [147], we have tried CDI and isocyanomethane on the activation of carboxylic group on ORO (Figure 50a), with the expectation that the activation of the COOH group would prompt a pre-organized reactions system, via intermediates 1 and 2, which could bring close the pentose phosphate and the base and decrease the reaction barrier (Figure 50b).

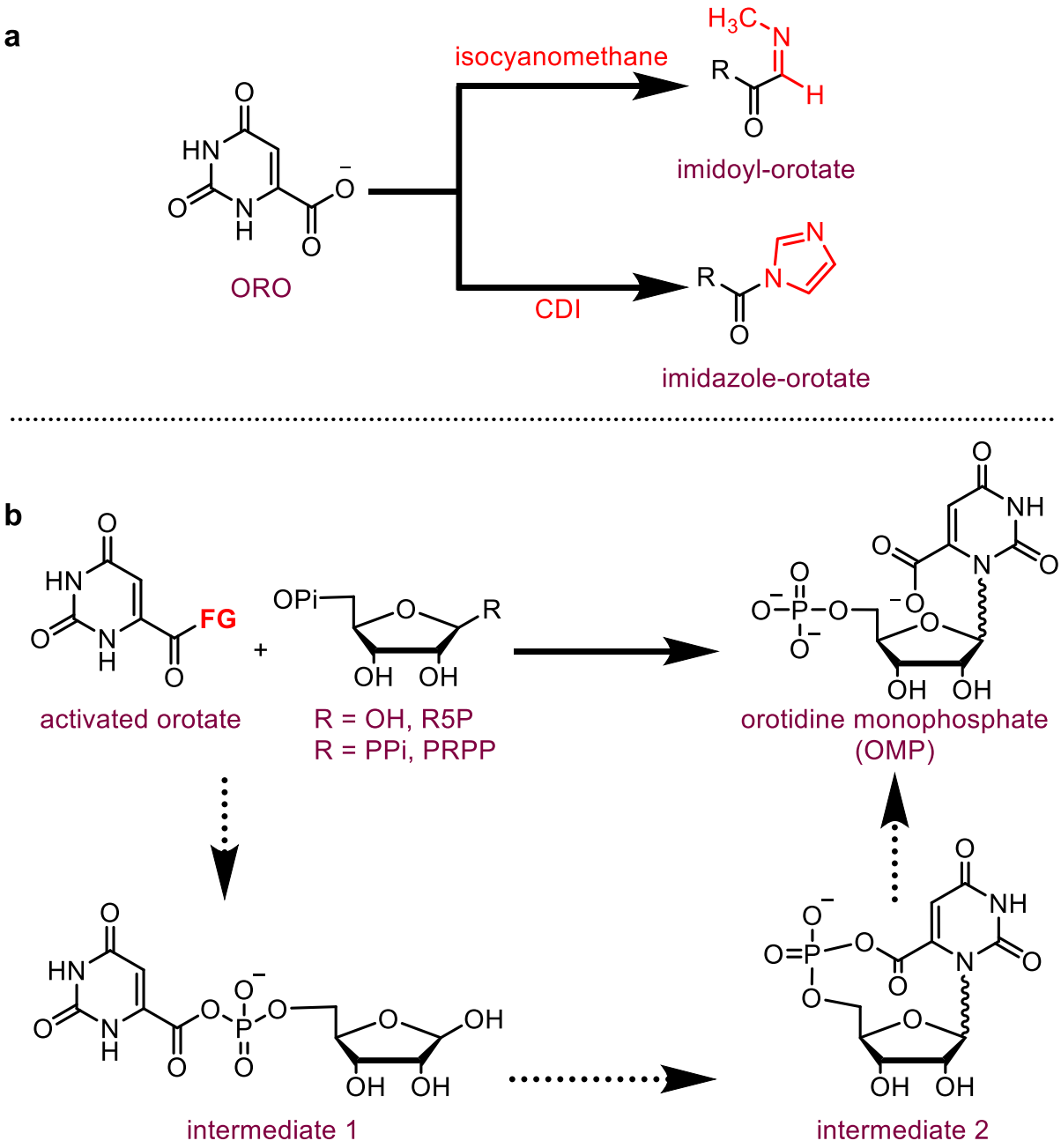


Figure 50. *a*. The isocyanomethane and CDI assisted COOH activation and their activated ORO. *b*. The proposed reaction scenario via intramolecular pre-organization of activated ORO to OMP.

To start investigating this reaction plausibility, we started with synthesizing imidazole-orotate by mixing and heating ORO and CDI in DMSO. The crystalized product was then subjected into water at room temperature for 16 h in the presence of PRPP and different metal salts. Interestingly, the reaction with the addition of $FeSO_4$ produced a new compound. The result was compared to the commercial OMP signal in LC-QTOF-MS analysis and it indicated different retention times. Thus, this newly formed molecule was suggested as intermediate 1 or as its structural isomer due to the identical m/z value found in analysis (Figure 51). The same signal at t_R = 3.7 min was obtained when R5P or ribose/TMP was used instead of PRPP.

However, the lack of NMR spectral information for this reaction due to analytical difficulties made the hypothetical formation of intermediate 1 debatable. The low concentration and low conversion of this reaction did not allow 2D NMR studies or ^{31}P NMR analysis. The only information we extracted from it was the MS analysis, which was not enough to confirm the structure of this new molecule. Therefore, new investigations are ongoing in our lab and focusing on pushing the reaction towards the formation of OMP. We have purchased an authentic sample of OMP from a commercial supplier and thus have enough information about the NMR spectra and MS patterns.

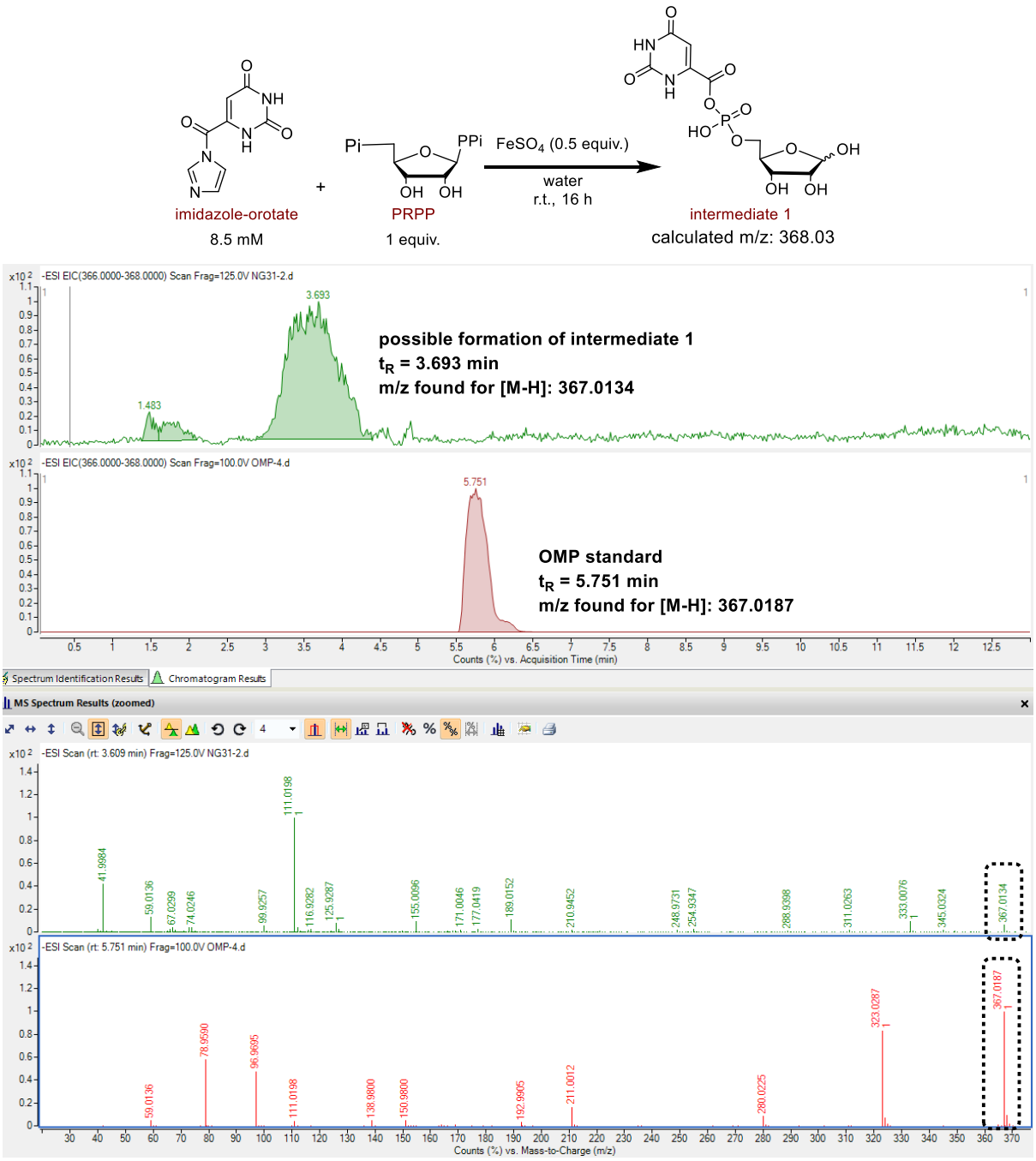


Figure 51. LC-QTOF-MS result of the coupling reaction between imidazole-orotate and PRPP. Reaction condition: 8.5 mM of imidazole-orotate, 1 equiv. of PRPP, 0.5 equiv. of $FeSO_4$ in 1 mL of water at room temperature for 16 h. The result is compared to commercial OMP (red signals).

PART III

5 Materials and methods

5.1 General information

^{1}H NMR spectra were recorded on a Bruker Avance Neo-500 (500 MHz) spectrometer at 0 °C or Bruker UltraShield Plus Avance III spectrometer (400 MHz) at ambient temperature (23 °C), in a H_2O:D_2O mixture (2:1) as solvent, with dimethyl sulfone as the internal standard (3.036 ppm at 0 °C and 3.056 ppm at ambient temperature). At 0 °C, water suppression was achieved using the Bruker ZGCPGPPR pulse program. The relaxation delay D1 was set to 13 s, with time domain size TD = 65536 and sweep width SWH = 10000.000 Hz. 32 scans were acquired for each sample. At ambient temperature, water suppression was achieved using the Bruker ZGESFPGP plulse program on 500 MHz spectrometer, the relaxation delay D1 = 1s, with TD = 65536, SWH = 10000.000 and 16 scans were required. On 400 MHz spectrometer, water suppression was achieved with the ZGESGPPE pulse program, with D1 = 40 s, TD = 32768, SWH = 8417.509 Hz and 16 scans were acquired for each sample. Integration was performed using *MestReNova v14.2.1* software.

U-LC-QTOF-MS analysis was performed on an Agilent 1290 Infinity II liquid chromatography (LC) system hyphenated with QTOF 6546 Accurate-Mass QTOF LC/MS system, equipped with a Dual Agilent Jet Stream spray source (ESI) and connected to an N_2 generator (Nitrocraft® NCP-032-R). All chromatographic separations were performed on an Agilent InfinityLab Poroshell 120 HILIC-Z column (2.1 x 150 mm, 1.9 µm), with a flow of 0.35 mL/min. Eluting buffers were A (10 mM CH_3COONH_4 in H_2O at pH 9) and B (10 mM CH_3COONH_4 in H_2O/MeCN (10/90) at pH 9), pH of both phases was adjusted by adding NH_4OH solution. The gradient was 0 → 2 min, 5% of A; 2 → 9 min, 23% of A; 9 → 11 min, 95 % of A; 11 → 17 min, 5% of A. ESI-QTOF-MS analysis was performed in a 2 GHz extended dynamic range, negative ionization mode, with fragmentor energy of 160 V, drying gas temp: 300°C, drying gas flow: 5 L/min, sheath gas temp: 350 °C, sheath gas flow: 11 L/min; nebulizer pressure: 50 psi, capillary V (+): 3 500 V, and skimmer: 65 V. The acquisition parameters were as follows: auto MS/MS mode; mass range 20–750 amu for MS and 20–750 amu for MS/MS experiment; and 1 spectrum/s acquisition. The collision-induced dissociation (CID) energy was optimized for 7 V. The identification of the compounds present in the samples was performed by comparison of MS/MS spectra with authentic samples (for compound **1**, the MS/MS spectra were compared with records from the SDBS database). Data analysis and integration were

performed using *Agilent MassHunter Workstation Quantitative Analysis for Q-TOF v.B.10.1* software.

5.2 Materials

Unless otherwise noted, all reagents and solvents were purchased from Sigma-Aldrich, Fluka, TCI, Acros organics or Carbosynth, and used without further purification. Water was obtained from a Milli-Q purification system (18.2 MΩcm). All reactions were carried out in 2 mL BRAND® microcentrifuge tubes with lid. Photo-oxdiations were carried out in Pyrex® quartz tubes sealed with cap. All stir bars were pre-washed with aqua regia, followed by distilled water and acetone, and oven dried to prevent any cross-contamination by metal salts. Further control experiments without the use of any metals were carried out in Thermo Scientific™ Digital Heating Shaking Drybath (1.5 mL tube block) without using any stir bars. The pH value was measured using VWR Dosatest® pH test strips (pH 0-14). The temperature of the sand bath was monitored by thermometer (Carl Roth®, DIN 12784, -10 °C to + 250 °C).

5.3 Analytical methods

5.3.1 Product identification by ^{1}H NMR

To facilitate ^{1}H NMR analysis of the reaction mixture, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove metallic ions in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. Then, the pH of each reaction was adjusted to 8 with $NaHCO_3$ powder if the reaction mixture was originally acidic, otherwise no pH adjustment was needed. ^{1}H NMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O).

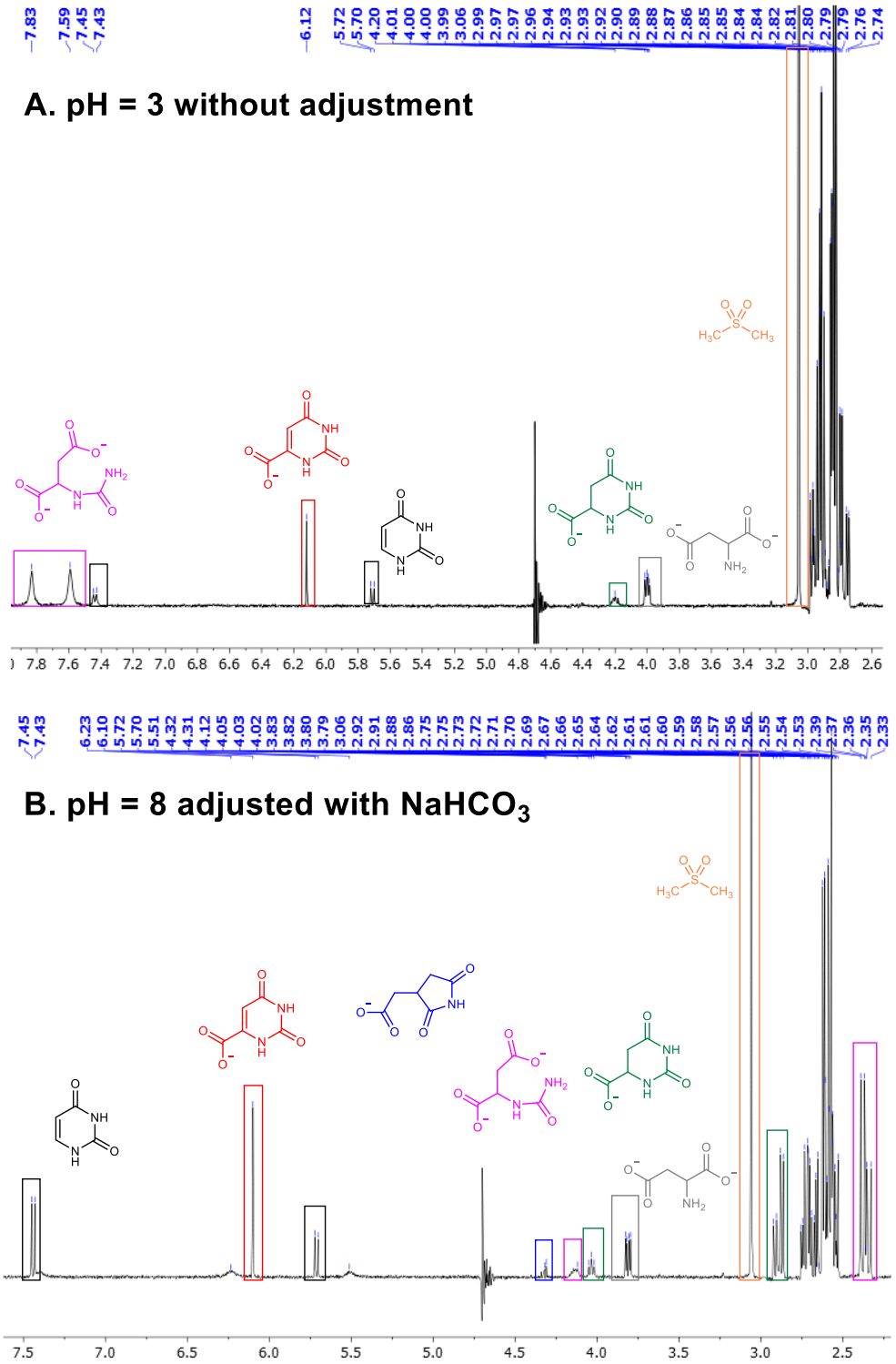


Figure S1. ^{1}H NMR (400 MHz, ns = 16, d1 = 40 s) of the mixture of all intermediates (authentic samples) detected in this study. **A**. Mixture of 6 compounds and internal standard dimethyl sulfone DMS in H_2O/D_2O (400/200) at pH 3 without pH adjustment. **B**. Mixture of 6 compounds and internal standard DMS in H_2O/D_2O (400/200) at pH 8 with the addition of $NaHCO_3$. The highlighted peaks are used for further quantitative characterization of each compound in the reaction mixtures.

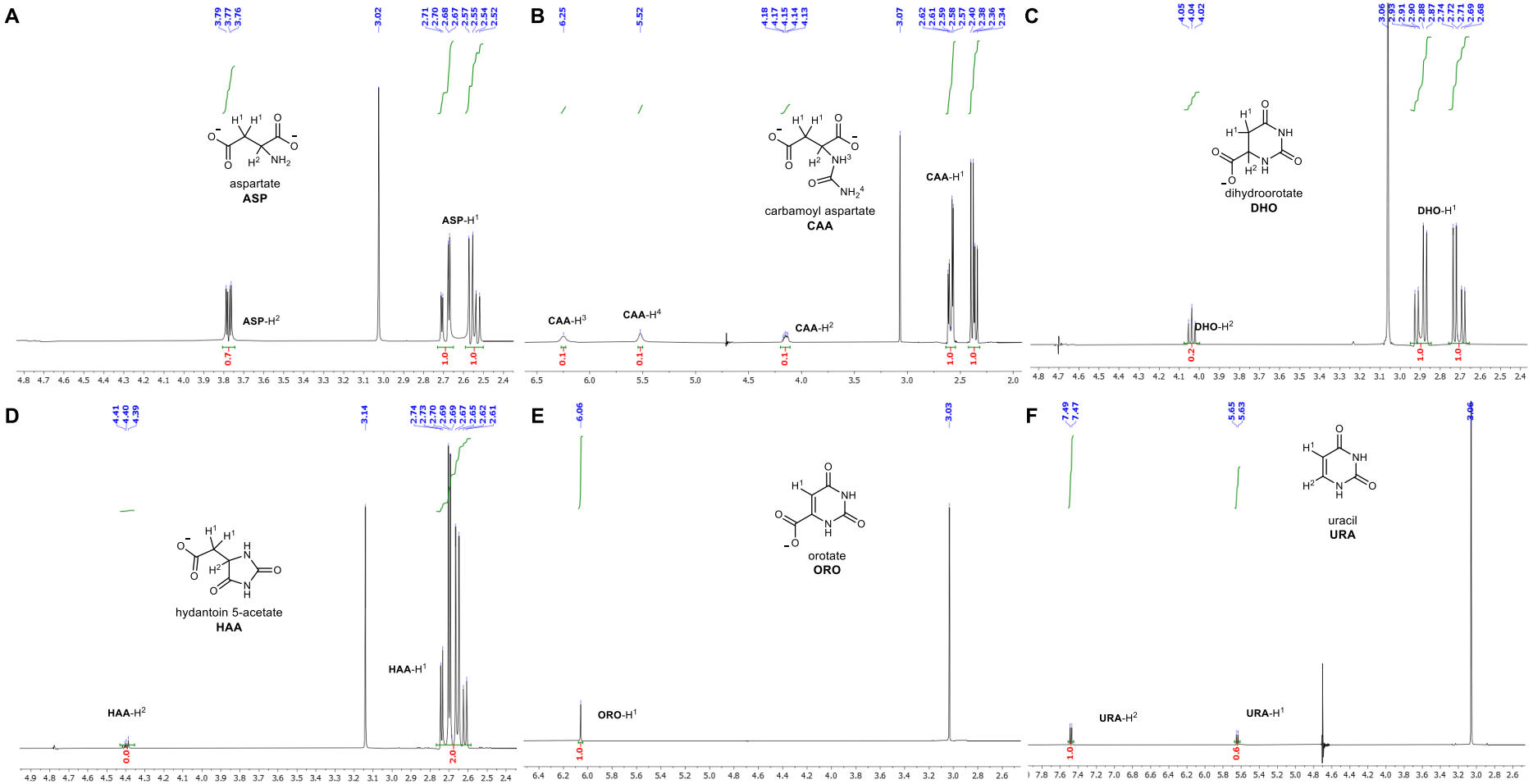
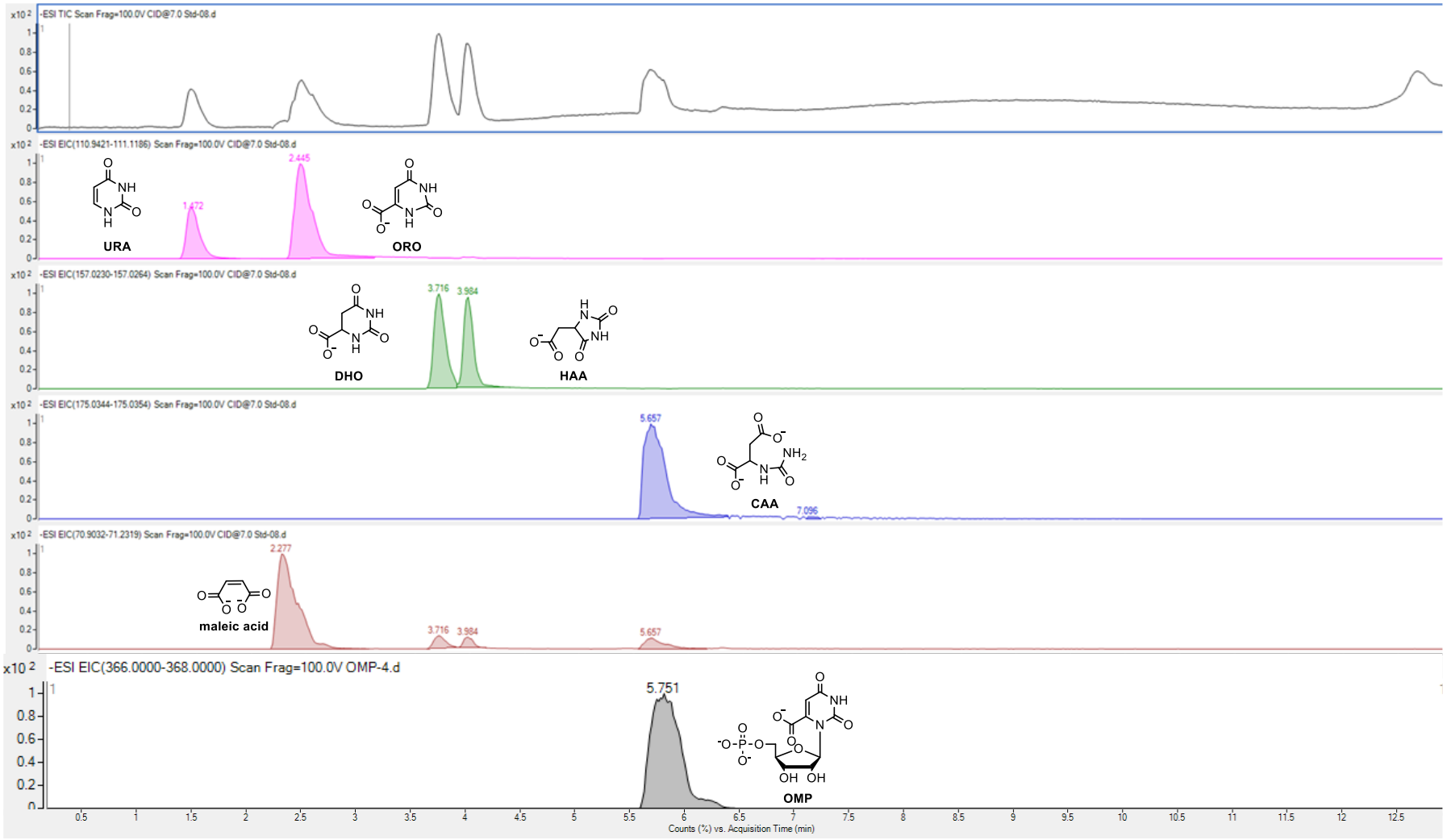


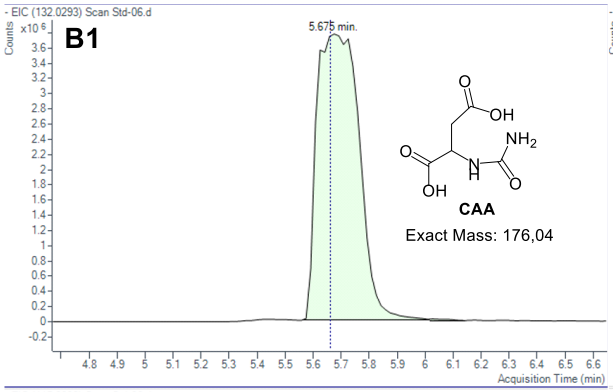
Figure S2. ^{1}H NMR (400 MHz, ambient temperature, ns = 16, d1 = 40 s) of an authentic sample (commercially available) of every intermediate detected in the study of pyrimidine nucleobase synthesis. *N.b.:* most water suppression methods attenuate other H signals around 4.7 ppm. For example, in the case of ASP, the H-1 proton signal is affected, and its integration is <1 (Fig. S2-A). [160]

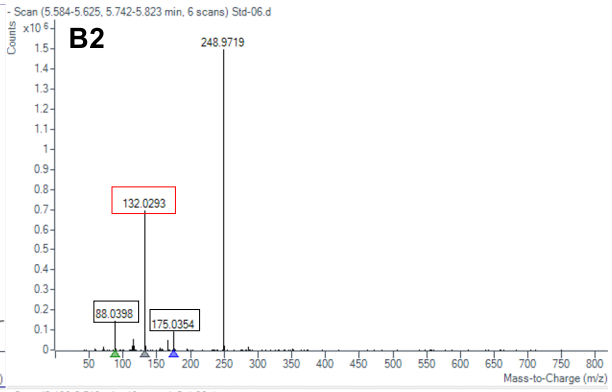
5.3.2 Product identification by LC-QTOF-MS/MS

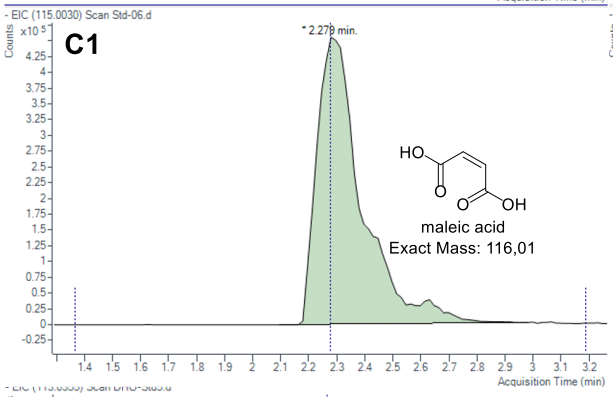
For LC-QTOF-MS/MS analysis, the same work up procedure as previously described in section A was applied. For LC-QTOF-MS/MS sample preparation, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes (with conical insert).

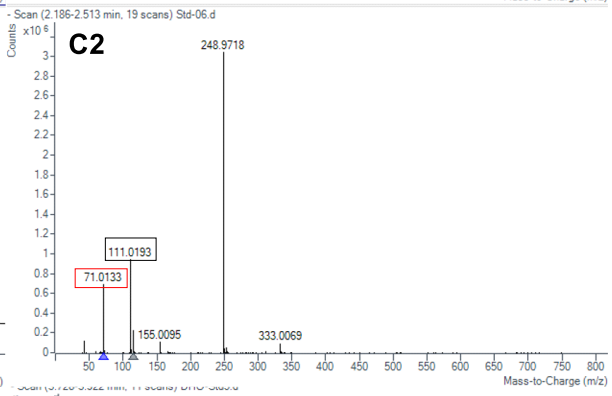
A. mixture of products

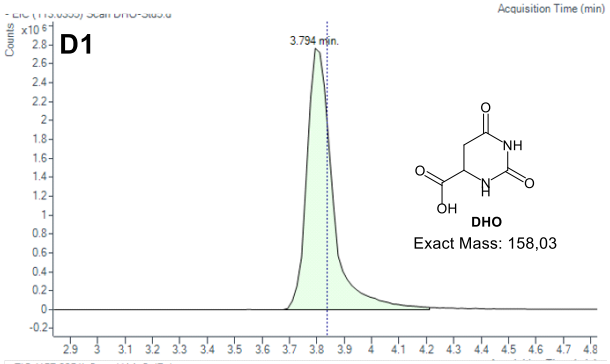


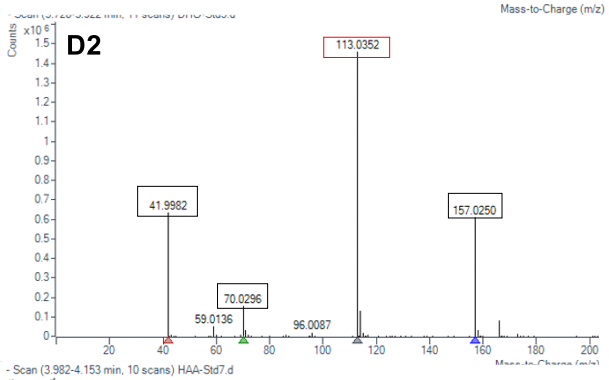

- EIC (132.0293) Scan Std-06.d
B1
5.675 min.
CAA
Exact Mass: 176,04
Acquisition Time (min)

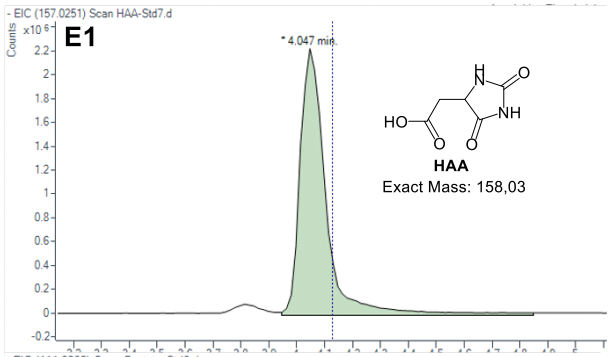

- Scan (5.584-5.625, 5.742-5.823 min, 6 scans) Std-06.d
B2
248.9719
132.0293
88.0398
175.0354
Mass-to-Charge (m/z)

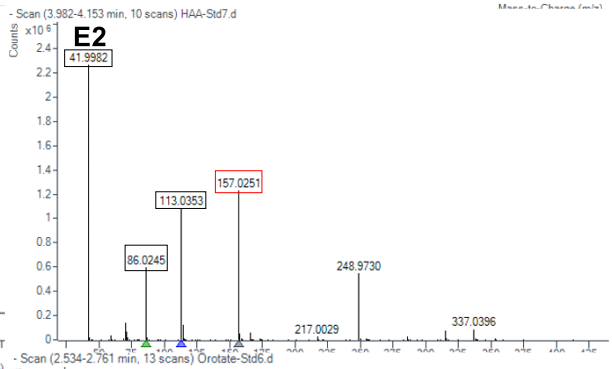

- EIC (115.0030) Scan Std-06.d
C1
2.279 min.
maleic acid
Exact Mass: 116,01
Acquisition Time (min)

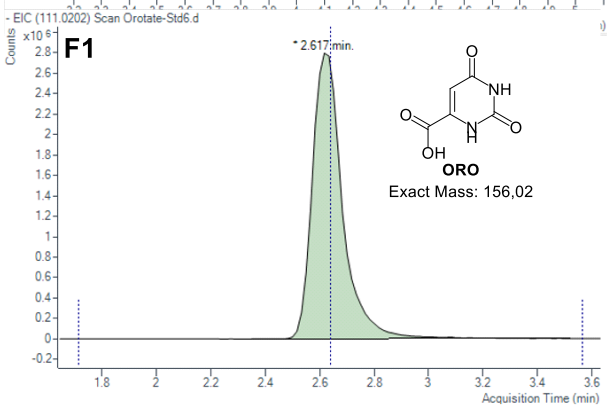

- Scan (2.186-2.513 min, 19 scans) Std-06.d
C2
248.9718
111.0193
71.0133
155.0095
333.0069
Mass-to-Charge (m/z)

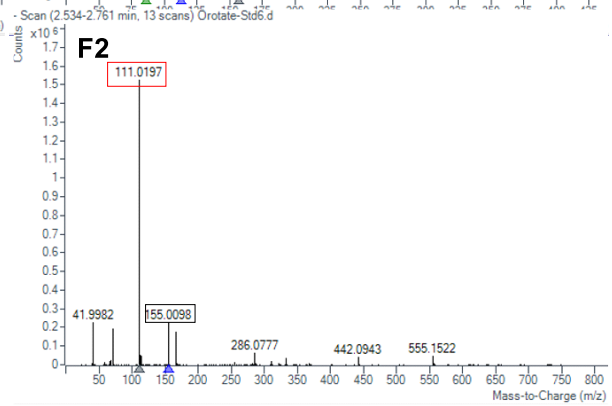

D1
3.794 min.
DHO
Exact Mass: 158,03


D2
113.0352
41.9982
157.0250
70.0296
59.0136
96.0087


- EIC (157.0251) Scan HAA-Std7.d
E1
4.047 min.
HAA
Exact Mass: 158,03


- Scan (3.982-4.153 min, 10 scans) HAA-Std7.d
E2
41.9982
157.0251
113.0353
86.0245
248.9730
217.0029
337.0396


- EIC (111.0202) Scan Orotate-Std6.d
F1
2.617 min.
ORO
Exact Mass: 156,02
Acquisition Time (min)


- Scan (2.534-2.761 min, 13 scans) Orotate-Std6.d
F2
111.0197
41.9982
155.0098
286.0777
442.0943
555.1522
Mass-to-Charge (m/z)

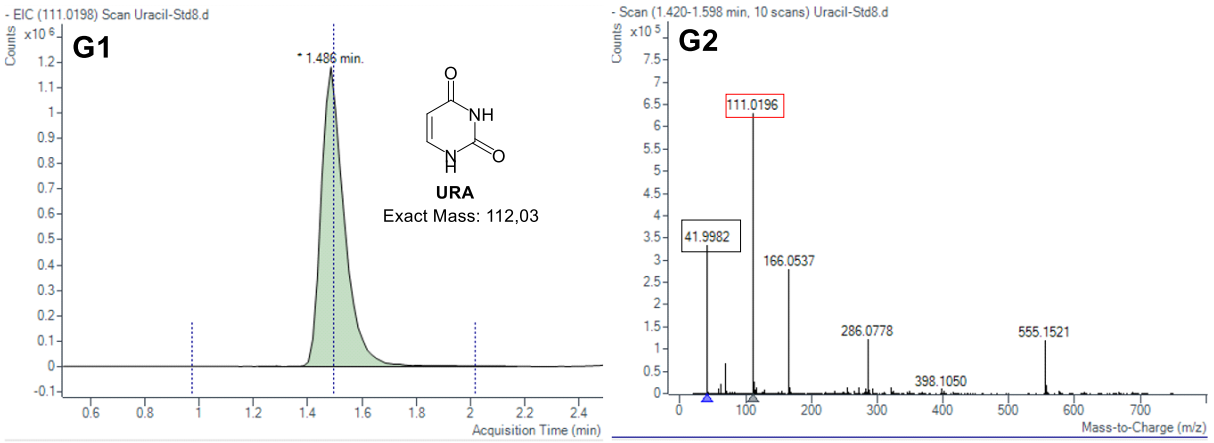


Figure S3. LC-QTOF-MS/MS profile of every intermediate (authentic samples) detected in this study. *A*: Total ion current (TIC) of the mixture of CAA/DHO/HAA/ORO/URA with internal standard: maleic acid, and extracted ion chromatogram (EIC) for each molecule from the TIC. *B1/B2*: LC profile and MS/MS profile of CAA. *C1/C2*: LC profile and MS/MS profile of internal standard maleic acid. *D1/D2*: LC profile and MS/MS profile of DHO. *E1/E2*: LC profile and MS/MS profile of HAA. *F1/F2*: LC profile and MS/MS profile of ORO. *G1/G2*: LC profile and MS/MS profile of URA. Highlighted m/z values with squares in the MS/MS profile correspond to the major fragment ions of the authentic sample. The most intense fragment ion is shown in a red square and it was selected as a compound qualifier for qualification and quantification of the corresponding molecule in LC-QTOF-MS/MS data analysis.

5.3.3 Quantification and error analysis by ^{1}H NMR

For ^{1}H NMR quantification, calibration curves were made to compensate for the small errors induced by the use of Chelex® Resin due to partial product adsorption.[106]

1000 µL aqueous solutions of D-carbamoyl-DL-aspartic acid CAA (0.08, 0.66, 1.31, 2.63, 5.25 and 21.01 mM) (Fig. S4-A), orotic acid ORO (0.06, 0.49, 0.99, 1.97, 2.37 and 4.74 mM) (Fig. S4-B), uracil URA (0.04, 0.07, 0.14, 0.29, 0.57 and 1.14 mM) (Fig. S4-C), L-dihydroorotic acid DHO (0.15, 1.19, 2.37, 4.74, 9.49 and 18.97 mM) (Fig. S4-D) and (0.29, 1.17, 4.66, 9.33, 18.66 and 37.32 mM) (Fig. S4-E), hydantoin-5-acetic acid HAA (0.79, 1.58, 3.16, 6.32, 12.65, and 25.30 mM) (Fig. S4-F) and (0.52, 1.04, 2.08, 4.16, 8.33 and 16.66 mM) (Fig. S4-G), *N*-carbamoyl -glycine (0.36, 0.72, 1.44, 2.88, 5.76 and 11.52 mM) (Fig. S4-H), D-carbamoyl-DL-glutamic acid (0.18, 0.37, 0.74, 1.47, 2.95 and 5.89 mM) (Fig. S4-I), were prepared by diluting their respective stock solutions (21.01 mM solution of CAA, 31.59 mM of ORO, 36.58 mM of URA, 37.95 mM of DHO, 25.30 mM of HAA, 11.52 mM of *N*-carbamoyl glycine, 5.89 mM of *N*-carbamoyl glutamic acid in MilliQ water) with MilliQ water.

During the preparation of the 1 mL calibration solutions from dilution, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was added to each vial and the suspension vortex-mixed for 1 h, followed by centrifugation at 6 000 rpm for 5 min. The pH of each stock solution was then adjusted to 8 using $NaHCO_3$ solid. ^{1}H NMR sample was prepared by taking 400 µL of the supernatant and adding 200 µL of 6.80 mM solution of internal standard (dimethyl sulfone in D_2O). Each of the standard calibration samples was subjected to ^{1}H NMR spectroscopy using the water suppression technique at ambient temperature described in the General information section above. The data from these measurements for every concentration allowed us to obtain six-point calibration plots for CAA, ORO, URA DHO, HAA, *N*-carbamoyl glycine and *N*-carbamoyl glutamic acid, correlating the substrate-to-standard ratios of peaks (2.35 ppm for CAA, 2.88 ppm for DHO, 4.31 ppm for HAA or 5.98 ppm for ORO, 7.49 ppm for URA, 3.58 ppm for *N*-carbamoyl glycine, 1.75 ppm for *N*-carbamoyl glutamic acid and 3.06 ppm for the DMS standard) with the product concentration (Figure S4). Yields were calculated by comparing the ^{1}H NMR integral against the calibration curve. Each reaction was performed at least twice, and reported yields are an average of those multiple runs.

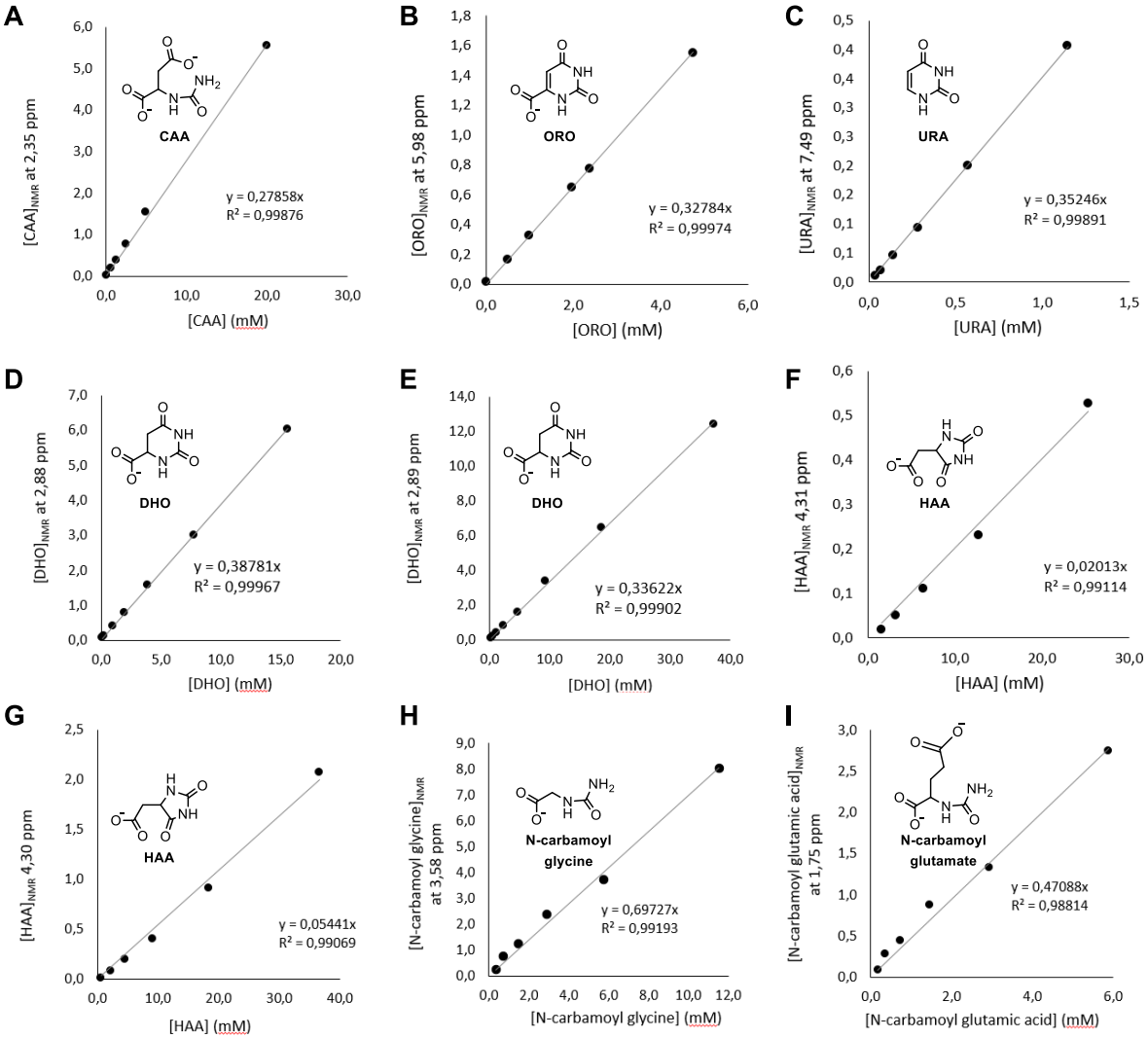


Figure S4. Correlation between the concentration of an aqueous solution of five intermediates CAA, DHO, HAA, ORO and URA and the measured ^{1}H NMR integral relative to DMS which was set to 6.0000. A/B/C/D/F: results were recorded on Bruker UltraShield Plus Avance III 400 (400 MHz), at ambient temperature with ns = 16 and d1 = 40 s. E/G/H/I: results were recorded on Bruker Ascend Spectroscope Avance Neo-500 (500 MHz), at ambient temperature with ns = 16 and d1 = 1 s.

For ^{1}H NMR quantification at 0 °C, 1000 µL aqueous solutions of ORO (0.52, 1.04, 2.08, 4.16, 8.33 and 16.66 mM), URA (0.28, 0.56, 1.12, 2.23, 4.46 and 8.92 mM) were prepared by diluting their respective stock solutions (33.31 mM of ORO, 17.84 mM of URA in MilliQ water) with MilliQ water. During the preparation of the stock solution, the same procedure was applied here as describe in the previous paragraph. Each of the standard calibration samples was subjected to NMR spectroscopy using water suppression technique at 0 °C described in General information section above. The data from these six measurements for every concentration allowed us to obtain six-point calibration plots for ORO and URA, correlating the substrate-to-standard ratios of peaks (5.77 ppm for ORO, 7.15 ppm for URA and 2.73 ppm for the two methyl peaks of the standard dimethyl sulfone DMS) with the product concentration (Figure S5). Yields were calculated by comparing the ^{1}H NMR integral against the calibration curve. Each reaction was performed at least twice, and reported yields are an average of those multiple runs.

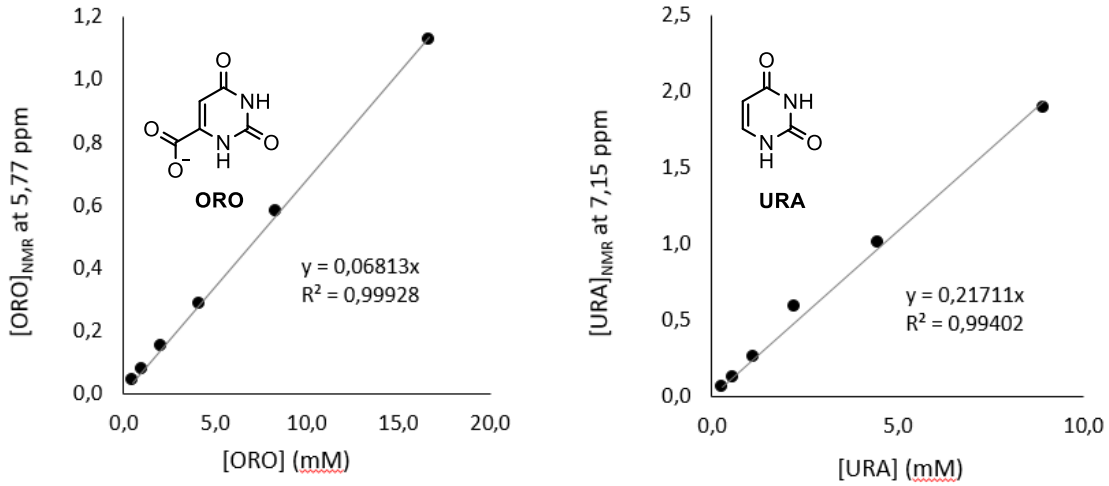


Figure S5. Correlation between the concentration of an aqueous solution of ORO and URA and the measured ^{1}H NMR integral relative to DMS which was set to 6.0000. Results were recorded on Bruker Ascend Spectrometer Avance Neo-500 (500 MHz), at 0 °C with ns = 32 and d1 = 13 s.

5.3.4 Quantification and error analysis by LC-QTOF-MS/MS

For LC-QTOF-MS/MS quantification, calibration curves were also made for the same reason.

For calibrations, 1000 µL aqueous solutions of DHO (0.06, 0.13, 0.26, 0.51, 1.03 and 2.06 mM), HAA (0.14, 0.28, 0.55, 1.11, 2.21 and 4.43 mM), ORO (0.12, 0.24, 0.48, 0.97, 1.93 and 3.86 mM) and URA (0.25, 0.50, 1.00, 2.01, 4.01 and 8.03 mM) were prepared by diluting with their respective stock solutions (8.22 mM of DHO, 8.85 mM of HAA and 8.03 mM of URA in MilliQ water) with MilliQ water.

During the preparation of their stock solution, the same procedure was applied as described for the previous ^{1}H NMR analysis. The LC-QTOF-MS/MS sample was prepared by taking 150 µL of the supernatant and adding 50 µL of 8.6 mM solution of the internal standard (maleic acid in MilliQ water). Each of the standard calibration samples was subjected to LC chromatography using gradient and fragmentation technique described in the General information section above. The data from these measurements for every concentration allowed us to obtain six-point calibration plots for DHO, HAA, ORO and URA correlating the substrate-to-standard ratios of peaks (t_R = 3.794 min and m/z [M-COOH] = 113.0352 for DHO, 4.047 min and m/z [M-H] = 157.0251 for HAA, t_R = 2.617 min and [M-COOH] = 111.0197 for ORO, t_R = 1.472 min and [M-H] = 111.0196 for URA, 2.223 min and m/z [M-COOH] = 71.0135 for maleic acid) with the product concentration (Figure S6). Yields were calculated by comparing the peak area relative to the internal standard against the calibration curve. Each reaction was performed at least twice, and reported yields are an average of those multiple runs.

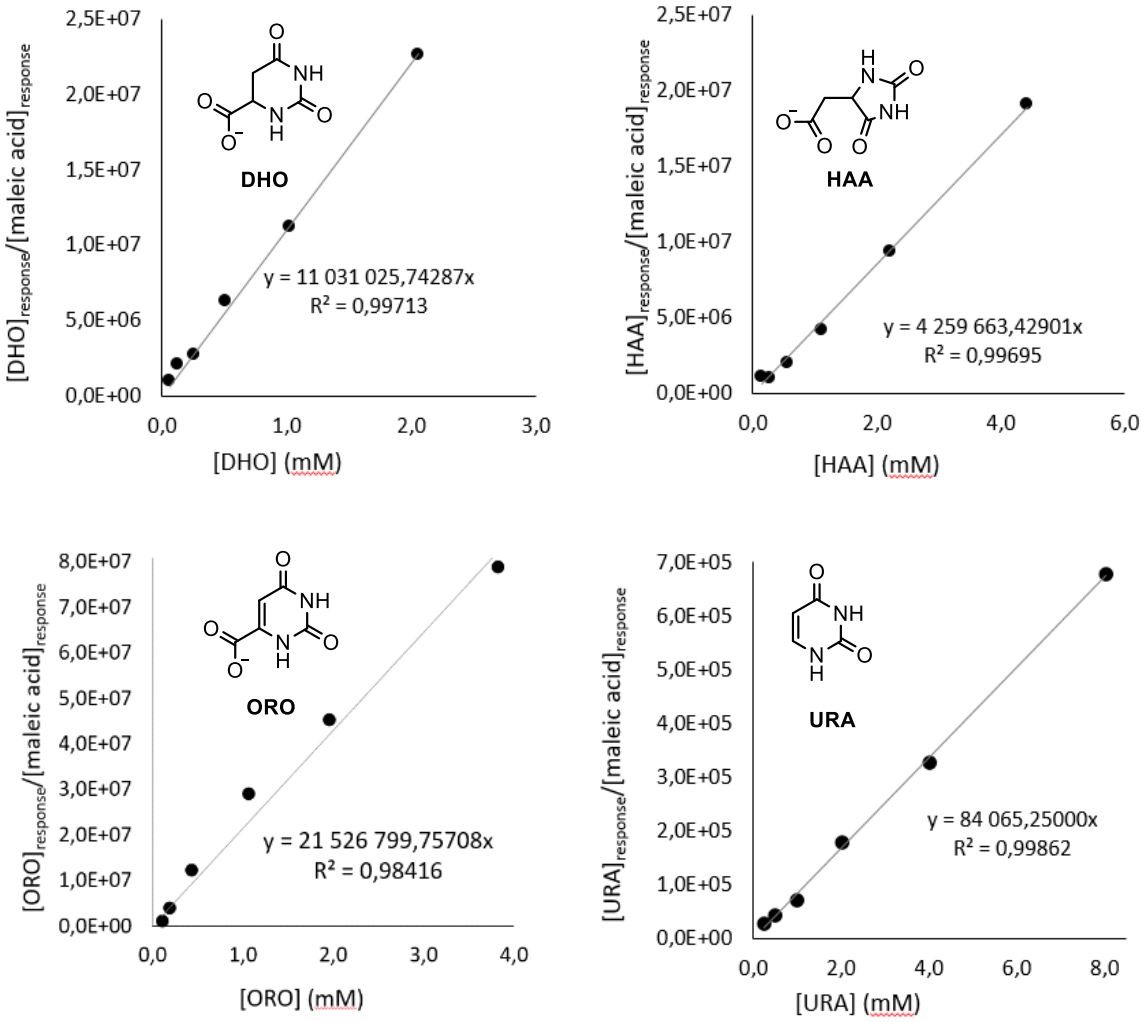


Figure S6. Correlation between the concentration of an aqueous solution of DHO, HAA, ORO and URA the measured relative response to maleic acid in LC-QTOF-MS/MS.

5.4 Synthetic procedures

5.4.1 Synthesis of starting materials

5.4.1.1 DAP synthesis

Diamido phosphate was synthesized according to a previously reported procedure.[161] A mixture of phenyl phosphorodiamidate (4.9915 g, 29.0 mmol), NaOH (2.3998 g, 60 mmol) and H_2O (15 mL) was stirred at 110 °C for 10 min. The solvent was evaporated under vacuum and 70 mL of ethanol was added at 0 °C in ice bath. The grey precipitate was filtered and dissolved in 50 mL of cold water. The resulting solution was washed two times with DCM (30 mL × 2) and multiple times with EtOAc until no phenol or phenoxide was present in the organic phase (checked by TLC). The aqueous phase was concentrated and was added dropwise to 50 mL of cold ethanol under vigorous stirring to afford a white precipitate. The precipitate was filtered, washed with cold ethanol and dried under vacuum to afford DAP as a white powder.

5.4.1.2 aTMP synthesis

Amido triphosphate was synthesized according to modified procedure from literature.[162] A mixture of sodium trimetaphosphate (90 g, 0.29 mol) and 30% (w/w) aqueous ammonia solution (300 mL, 4.7 mmo) was added to 600 mL of water. The reaction was stirred at room temp. for 48 h. The solvent was evaporated under vacuum until pH reached about 8-9. The synthesized compound was liquid and stored at 0 °C for further use.

5.4.1.3 Imidazole-orotate ester synthesis

To a solution of orotic acid monohydrate (1 equiv., 5.7 mmol, 1 g) in 10 mL of DMSO, was added carbonyldiimidazole (2.2 equiv., 12.5 equiv., 2 g). The mixture was stirred at 60 °C for 2 h. Then the brownish solution was kept stirring at room temp. overnight. Solvent was partially evaporated under vacuum. DCM was added to wash off the excess DMSO and imidazole. The obtained dark oil was then poured into Et_2O, brown precipitate was obtained and washed several time with DCM with sonication. The obtained brown solid was used for further reactions. ^{1}H NMR (500 MHz, DMSO-d6) δ 11.06 (s, 1H), 8.46 (s, 1H), 7.41 (s, 2H), 5.82 (s, 1H).

5.4.2 General procedure of the carbamoylation reaction

5.4.2.1 Carbamoylation with KOCN

To a 2 mL BRAND® microcentrifuge tube were added 1 mL of D,L-aspartic acid (1 equiv., 8.5 μmol, 1.1 mg), ammonium bicarbonate (10 equiv., 85 μmol, 6.7 mg) and potassium cyanate (1 equiv., 8.5 μmol, 0.7 mg) from stock solutions. The tube was then sealed, and the reaction mixture was stirred at 60 °C (in sand bath) for 16 h. Then, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was added to the reaction tube to correlate the error with the calibration curve. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. ^{1}H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.2.2 Carbamoylation with aTMP and CDI

To a 2 mL BRAND® microcentrifuge tube were added 1 mL of D,L-aspartic acid (1 equiv., 8.5 μmol, 1.1 mg), ammonium bicarbonate (10 equiv., 85 μmol, 6.7 mg) from stock solutions and sodium amido triphosphate (1 equiv., 8.5 μmol, 2.6 mg) together with carbonyl diimidazole (1 equiv., 8.5 μmol, 1.4 mg). The tube was then sealed, and the reaction mixture was stirred at 60 °C (in sand bath) for 16 h. Then, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was added to the reaction tube to correlate the error with the calibration curve. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. ^{1}H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.2.3 Carbamoylation with CAP

To a 2 mL BRAND® microcentrifuge tube were added 1 mL of D,L-aspartic acid (1 equiv., 8.5 μmol, 1.1 mg) and lithium carbamoylphosphate dibasic hydrate (1 equiv., 8.5 μmol, 1.5 mg) from stock solutions, as well as metal salt, with or without base (see Table S1). The tube was then sealed, and the reaction mixture was stirred at 60 °C (in sand bath) for 16 h. Then, pH was measured and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was added to the

reaction tube to remove transition metal ions in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 using $NaHCO_3$ solid if needed. ^1H qNMR sample was prepared by taking 400 µL of the supernatant and adding 200 µL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 µL of the supernatant and 50 µL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.2.4 Competitive carbamoylation between four amino acids

To a 2 mL BRAND® microcentrifuge tube were added 1 mL of glycine (1 equiv., 8.5 µmol, 0.6 mg), D,L-alanine (1 equiv., 8.5 µmol, 0.8 mg), D,L-aspartic acid (1 equiv., 8.5 µmol, 1.1 mg), D,L-glutamic acid monohydrate (1 equiv., 8.5 µmol, 1.4 mg) and lithium carbamoylphosphate dibasic hydrate (1 equiv., 8.5 µmol, 1.5 mg) from stock solutions, and $NaHCO_3$ (2 equiv., 17 µmol, 1.4 mg). The tube was then sealed, and the reaction mixture was stirred at 60 °C (in sand bath) for 16 h. Then, pH was measured (pH = 8) and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was added to the reaction tube to correlate the error related to the calibration curves. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. ^1H NMR sample was prepared by taking 400 µL of the supernatant and adding 200 µL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O).

5.4.3 General procedure of the dehydrative cyclization reaction

5.4.3.1 Metal-promoted cyclization

To a 2 mL BRAND® microcentrifuge tube was added 1 mL of D-carbamoyl-DL-aspartic acid (1 equiv., 8.5 µmol, 1.5 mg) from stock solution and metal salt, with or without base (see Table S2). The tube was then sealed, and the reaction mixture was stirred at 60 °C (temperature of the sand bath) for 16 h. Then, pH was measured and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid if needed. ^1H qNMR sample was prepared by taking 400 µL of the supernatant and adding 200 µL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 µL of the supernatant

and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.3.2 Phosphorylating agent promoted cyclization

To a 2 mL BRAND® microcentrifuge tube was added 1 mL of D-carbamoyl-DL-aspartic acid (1 equiv., 21 μmol, 3.7 mg) from stock solution in water and metal salt, metals salt (5 equiv., 105 μmol) and $POCl_3$ (5 equiv., 105 μmol, 10 μL). The tube was then sealed, and the reaction mixture was stirred at 60 °C (temperature of the sand bath) for 16 h. Then, pH was measured and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using KOH or $NaHCO_3$ solid if needed. ^1H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.4 General procedure of the oxidation reaction

5.4.4.1 Photo-oxidation

To a 10 mL a Pyrex® quartz tube with cap was added 1 mL of L-dihydroorotic acid (1 equiv., 8.5 μmol, 1.3 mg) from stock solution and sulfur salt (1 equiv., 8.5 μmol), with or without $NaHCO_3$. The mixture was first flushed with argon balloon for 10 min with continuous sonication, then was sealed. The reaction was stirred at ambient temperature (20 °C) in a photobox (UV-A + UV-B + UV-C) for 16 h. Then, pH was measured and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid while needed. ^1H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.4.2 Thermal oxidation of DHO

To a 2 mL BRAND® microcentrifuge tube was added 1 mL of L-dihydroorotic acid (1 equiv., 8.5 μmol, 1.3 mg) from stock solution and metal salt, with or without base (see **Table S3**). The tube was first flushed with an argon balloon for 10 min with continuous sonication, then sealed, and the reaction mixture was stirred at 60 °C (in a sand bath) for 16 h. Then, pH was measured and 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid while needed. ^1H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.4.3 Thermal oxidation of HAA

To a 2 mL BRAND® microcentrifuge tube were added 1 mL of hydantoin-5-acetic acid (1 equiv., 8.5 μmol, 1.3 mg) from a stock solution and metal (see **Table S4**). For **entry 3**, 5.9 equiv. (100 μL of 1 M NaOH solution) of NaOH was added from the beginning at t = 0 h. For **entry 4** and **5**, NaOH solution was added at t = 16 h and the reaction was maintained stirring at 60 °C for another 3 h. The reaction tube was first flushed with argon balloon for 10 min with continuous sonication, then sealed, and the reaction mixture was stirred at 60 °C (in a sand bath) for 16 h. Then, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid if needed. ^1H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethyl sulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.5 General procedure of the one-pot reaction

5.4.5.1 From CAA to ORO

entry 1 of Table S5: To a 2 mL BRAND® microcentrifuge tube were added 1 mL of *N*-carbamoyl-DL-aspartic acid (1 equiv., 21 μmol, 3.3 mg) from stock solution. 5 equiv. of $FeCl_3$ and 10 equiv. of NH_4HCO_3 were added at t = 0 h. The tube was first flushed with an argon balloon for 10 min with continuous sonication, then sealed, and the reaction mixture was stirred at 60 °C (in a sand bath) for 16 h. Then 0.4 equiv. of $CuSO_4$ and 5 equiv. of MnO_2 were added at t = 16 h, the reaction was kept stirring at 60 °C for under Ar atmosphere for another 16 h (t_{total} = 32 h).

entry 3-4 of **Table S5**: Same as above, but at t = 0 h, 0.4 equiv. of $CuSO_4$ was added to 1 mL of solution containing CAA. The tube was first flushed with an argon balloon for 10 min with continuous sonication, then sealed, and the reaction mixture was stirred at 60 °C (in a sand bath) for 16 h. Then, 5 equiv. of MnO_2 was added at t = 16 h and the mixture was maintained stirring at 60 °C for an extra 16 h. At t = 32 h, 9.5 equiv. of NaOH (200 μL of 1 M NaOH solution) was added and the reaction was kept stirring at 60 °C for 3 h (t_{total} = 35 h). After 35 h, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid if needed. ^{1}H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethylsulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.5.2 From ASP to ORO

entry 5 of Table S5 and Table S6**:** To a 2 mL BRAND® microcentrifuge tube were added 1 mL of D,L-aspartic acid (1 equiv., 21 μmol, 2.8 mg) and lithium carbamoylphosphate dibasic hydrate (2 equiv., 42 μmol, 6.4 mg) and 2 equiv. of $NaHCO_3$ from stock solutions. The tube was first flushed with argon balloon for 10 min with continuous sonication, then sealed, and the reaction mixture was stirred at 60 °C (in a sand bath) for 16 h. Then, at t = 16 h, 4.8 equiv. of HCl (100 μL of 1 M HCl solution) and 0.4 equiv. of $CuSO_4$ was added and the mixture was maintained under stirring at 60 °C for an extra 16 h. At t = 32 h, 5 equiv. of MnO_2 was added

and the mixture was maintained under stirring at 60 °C for an extra 16 h. At t = 48 h, 9.5 equiv. of NaOH (200 μL of 1 M NaOH solution) was added and the reaction was kept stirring at 60 °C for 6 h (t_{total} = 54 h). After 54 h, 150 mg of Chelex® 100 Resin (50-100 mesh, sodium form) was used to remove transition metals in the reaction mixture. After vortex mixing for 1 h, the resulting suspension was centrifuged at 6 000 rpm for 5 min. The pH of each reaction was adjusted to 8 at the end using $NaHCO_3$ solid if needed. A ^1H qNMR sample was prepared by taking 400 μL of the supernatant and adding 200 μL of 6.8 mM solution of internal standard (dimethylsulfone in D_2O). For LC-QTOF-MS/MS analysis, 150 μL of the supernatant and 50 μL of 8.6 mM solution of internal standard (maleic acid in MilliQ water) were added to glass microtubes.

5.4.6 General procedure of *N*-glycosylation

To a 2 mL BRAND® microcentrifuge tube were added imidazole-orotate ester (1 equiv., 8.5 μmol) and pentose phosphate (D-ribose-5-phosphate disodium salt hydrate, or 5-phospho-D-ribose 1-diphosphate pentasodium salt, 1 equiv., 8.5 μmol) or D-ribose together with trisodium trimetaphosphate (1 equiv. of each), metal salt (0.5 equiv., 4.3 μmol) in 1 mL of water. The reaction was stirred on Shaking Drybath for 16 h. Then, the resulting suspension was centrifuged. For LC-QTOF-MS/MS analysis, 150 μL of the supernatant were added to glass microtubes.

5.5 Experimental data

Table S1: Reaction condition optimization for the aspartate ASP substitution with product quantification by ^{1}H qNMR

Entry	Substrate	[substrate] mM	solvent	pH	conditions				base		product quantification			
					MgO (eq.)	$MgCl_2$ (eq.)	$CuSO_4$ (eq.)	$FeCl_3$ (eq.)	NH_4HCO_3 (eq.)	$NaHCO_3$ (eq.)	CH at 2.35 ppm (1H) integral relative to DMS (6H)[e]		[CAA] (mM)	yield of CAA (%)[c]
1	ASP+CAP	8.26	water	9	-	-	-	-	-	10		0.6790	2.44	29.49
												0.7466	2.68	32.39
												0.8425	3.02	36.59
											Average	**0.7398**	**2.71**	**32.82 ± 3.57**
2	ASP+CAP	8.26	water	8	-	-	-	-	-	2		1.0050	3.61	43.62
												0.9904	3.56	43.02
												0.9845	3.53	42.76
											Average	**0.9917**	**3.57**	**43.13 ± 0. 44**
3	ASP+CAP	8.26	water	8	5	-	-	-	-	2		0.7609	2.73	33.02
												0.9497	3.41	41.25
												1.0200	3.66	44.30
											Average	**0.9434**	**3.27**	**39.52 ± 5.83**
4	ASP+CAP	8.26	water	8	2	-	-	-	-	2		0.7796	2.80	33.86
												0.7060	2.54	30.71
												0.7911	2.84	34.36
											Average	**0.7498**	**2.73**	**32.98 ± 1.98**
5	ASP+CAP	8.26	water	7	-	5	-	-	-	2		0.7055	2.53	30.64
												0.8419	3.02	36.57
												0.7784	2.79	33.78
											Average	**0.7741**	**2.78**	**33.66 ± 2.97**
6	ASP+CAP	8.26	water	8	-	-	0.4	-	-	2		0.0818	0.29	3.55
												0.0520	0.18	2.20
												0.1363	0.49	5.92
											Average	**0.0970**	**0.32**	**3.89 ± 1.88**
7	ASP+CAP	8.26	water	3	-	-	-	5	-	2		-	-	-
												-	-	-
												-	-	-

											Average	**-**	**-**	**-**
8	ASP+CAP	8.26	water	8	5	-	-	-	2	-		0.6233	2.24	27.07
												0.5853	2.10	25.37
												0.4861	1.74	21.11
											Average	**0.5745**	**2.03**	**24.52 ± 3.07**
9	ASP+CAP	8.26	water	8	-	-	-	-	2	-		0.3209	1.16	13.97
												0.2329	0.84	10.12
												0.2347	0.84	10.19
											Average	**0.2549**	**0.95**	**11.43 ± 2.20**
10	ASP+CAP	8.26	1 M pH 8 phosphate buffer	8	-	-	-	-	-	-		0.2358	0.85	10.24
												0.2081	0.75	9.04
												0.2752	0.99	11.95
											Average	**0.2358**	**0.86**	**10.41 ± 1.46**
11	ASP+CAP	8.26	water	8	5	-	-	-	-	-		0.6872	2.46	29.81
												0.5746	2.06	24.96
												0.6166	2.21	26.78
											Average	**0.6233**	**2.24**	**27.18 ± 2.45**
12	ASP+CAP	8.26	water	7	-	-	-	-	-	-		0.2435	0.87	10.58
												0.2241	0.80	9.73
												0.2197	0.79	9.54
											Average	**0.2329**	**0.82**	**9.95 ± 0.55**
13 [a]	ASP+CAP	8.26	water	8	-	-	-	-	-	2		1.4795	5.31	64.21
												1.5450	5.55	67.11
												1.3766	4.94	59.79
											Average	**1.4637**	**5.27**	**63.70 ± 3.69**
14 [a]	ASP+CAP	8.26	water	8	5	-	-	-	-	2		1.5430	5.54	67.02
												1.5693	5.63	68.11
												1.5450	5.55	67.11
											Average	**1.5455**	**5.57**	**67.41 ± 0.61**
15 [a,b]	ASP+CAP	20.96	water	8	-	-	-	-	-	2		0.1262	0.45	5.48
												0.1602	0.58	6.96
												0.1575	0.57	6.84
											Average	**0.1462**	**0.53**	**6.43 ± 0.82**
16 [a]	ASP+CAP	20.96	water	8	-	-	-	-	-	2		4.3883	15.75	75.15

												4.0788	14.64	69.85
												4.9865	17.90	85.40
											Average	**4.3010**	**16.10**	**76.80 ± 7.91**
												-	-	-
17 [c]	ASP+CAP	20.96	water	8	-	-	-	-	-	2		-	-	-
												-	-	-
											Average	**-**	**-**	**-**
												-	-	-
18 [d]	ASP+CAP	8.26	water	3	-	-	-	-	-	-		-	-	-
												-	-	-
											Average	**-**	**-**	**-**

[a] 2 equiv. of **CAP** was used. [b] The reaction was carried out at ambient temperature (20 °C). [c] The reaction was carried out at 0 °C. [d] The pH of the reaction was adjusted to 3 with diluted HCl solution. [e] Reactions were quantified on Bruker 400 MHz.

Table S2: Reaction condition optimization for the cyclization of CAA with product quantification by ^{1}H qNMR or LC-QTOF-MS/MS

Entry[a]	Substrate	[CAA] mM	solvent	pH	conditions					base		product quantification					
												DHO			HAA		
					MgO (eq.)	$FeCl_3$ (eq.)	$CuSO_4$ (eq.)	$CuCl_2$ (eq.)	$ZnSO_4$ (eq.)	NH_4HCO_3 (eq.)	$NaHCO_3$ (eq.)	[CAA] response relative to internal standard[a]	[DHO] (mM)	yield of DHO (%)	HAA response relative to internal standard	[HAA] (mM)	yield of HAA (%)
1	CAA	21.01	0.1 M HCl	1	-	-	-	-	5	-	-	0.3974	1.03	4.89	0.6258	11.50	54.75
												0.4256	1.10	5.22	0.6661	12.24	58.27
												0.4062	1.05	4.99	0.5707	10.49	49.92
												Average	**1.06**	**5.03 ± 0.17**	**0.6209**	**11.41**	**54.31 ± 4.19**
2	CAA	21.01	0.1 M HCl	1	5	-	-	-	-	-	-	-	-	-	0.0298	0.55	2.61
												-	-	-	0.0219	0.40	1.90
												-	-	-	0.0320	0.59	2.80
												Average	**-**	**-**	**0.0279**	**0.51**	**2.44 ± 0.47**
3	CAA	21.01	0.1 M HCl	1	-	-	5	-	-	-	-	0.3300	0.85	4.05	0.4767	8.76	41.70
												0.3754	0.97	4.61	0.4511	8.29	39.46
												0.3383	0.87	4.13	0.4942	9.08	43.22
												Average	**0.90**	**4.26 ± 0.30**	**0.4740**	**8.71**	**41.46 ± 1.89**
4	CAA	21.01	0.1 M HCl	1	-	-	0.4	-	-	-	-	0.3238	0.84	3.99	0.4577	8.41	40.04
												0.2588	0.67	3.18	0.4511	8.29	39.46
												0.4218	1.09	5.18	0.3811	7.00	33.34
												Average	**0.87**	**4.12 ± 1.01**	**0.4300**	**7.90**	**37.61 ± 3.71**
5	CAA	21.01	0.1 M HCl	1	-	-	-	0.4	-	-	-	0.4317	1.11	5.29	0.4550	8.36	39.80
												0.3694	0.95	4.53	0.4689	8.62	41.02
												0.4843	1.25	5.94	0.3863	7.10	33.79
												Average	**1.10**	**5.25 ± 0.71**	**0.4367**	**8.03**	**38.20 ± 3.87**
6[b]	CAA	21.01	0.1 M HCl	1	-	-	-	-	5	-	-	-	-	-	-	-	-
												-	-	-	-	-	-
												-	-	-	-	-	-
												Average	**-**	**-**	**-**	**-**	**-**
7[c]	CAA	8.52	0.1 M HCl	1	-	-	0.4	-	-	-	-	0.1252	0.32	3.79	0.0585	2.91	34.12
												0.0725	0.19	2.20	0.0298	1.48	17.38
												0.1089	0.29	3.30	0.0492	2.44	28.69
												Average	**0.27**	**3.10 ± 0.81**	**0.0458**	**2.28**	**26.73 ± 8.54**
8	CAA	8.52	0.1 M HCl	1	-	-	-	-	-	-	-	-	-	-	15281353.3215	3.59	42.12
												-	-	-	10206934.1432	2.40	28.14

													-	-	-	14051214.2229	3.30	38.73
												Average	-	-	-	**13179833.8959** [d]	**3.09**	**36.33 ± 7.30**
9	CAA	8.52	water	2	-	-	-	-	5	-	-		2606941.2406	0.24	2.77	6777855.4870	1.59	18.68
													2369528.0900	0.21	2.52	4188179.5308	0.98	11.54
													3749865.6778	0.34	3.99	4777830.1401	1.12	13.17
												Average	**2908778.3362** [d]	**0.26**	**3.10 ± 0.79**	**5247955.0526** [d]	**1.23**	**14.47 ± 3.74**
10	CAA	8.52	water	2	-	5	-	-	-	-	-		2195131.0456	0.20	2.34	15144583.3962	3.56	41.75
													3264010.2972	0.30	3.47	19064064.0628	4.48	52.55
													3286456.0526	0.30	3.50	23277042.4948	5.46	64.16
												Average	**2915199.1318** [d]	**0.26**	**3.10 ± 0.66**	**19161896.6513** [d]	**4.50**	**52.82 ± 11.21**
11	CAA	8.52	water	4	-	-	5	-	-	-	-		264759.6438	0.02	0.28	323944.5442	0.08	0.89
													278083.4436	0.03	0.30	234090.2463	0.05	0.65
													223143.0944	0.02	0.24	310741.8893	0.07	0.86
												Average	**255328.7273** [d]	**0.02**	**0.27 ± 0.03**	**289592.2266** [d]	**0.07**	**0.80 ± 0.133**
12	CAA	8.52	water	4	-	-	-	-	-	-	-		-	-	-	-	-	-
													-	-	-	-	-	-
													-	-	-	-	-	-
												Average	-	-	-	-	-	-
13	CAA	8.52	water	8	-	-	0.4	-	-	-	2		769067.4726	0.07	0.82	-	-	-
													996736.0773	0.09	1.06	-	-	-
													796171.3416	0.07	0.85	-	-	-
												Average	**853991.6305** [d]	**0.08**	**0.91 ± 0.13**	-	-	-
14[c]	CAA	8.52	water	8	-	-	0.4	-	-	2	-		0.1076	0.28	3.26	-	-	-
													0.1114	0.29	3.37	-	-	-
													0.0916	0.24	2.77	-	-	-
												Average	**0.1035**	**0.27**	**3.13 ± 0.32**	-	-	-
15	CAA	8.52	water	4	-	-	0.4	-	-	-	-		395629.8766	0.04	0.42	432166.9720	0.10	1.19
													429625.9123	0.04	0.46	481748.1303	0.11	1.33
													141316.0672	0.01	0.15	303880.2932	0.07	0.84
												Average	**322190.6187** [d]	**0.03**	**0.34**	**405931.7985** [d]	**0.10**	**1.12 ± 0.25**

16[c]	CAA	8.52	water	5	-	5	-	-	-	10	-		0.1053	0.28	3.20	0.0583	2.90	34.00
													0.0725	0.19	2.2	0.0298	1.48	17.38
													0.133	0.34	4.03	0.0553	2.75	32.26
												Average	**0.1036**	**0.27**	**3.14 ± 0.92**	**0.0478**	**2.38**	**27.88 ± 9.14**
17[c]	CAA	8.52	water	3	-	5	-	-	-	2	-		0.0491	0.13	1.49	0.0263	1.31	15.34
													0.0869	0.22	2.63	0.0401	1.99	23.39
													0.0350	0.09	1.06	0.0318	1.58	18.55
												Average	**0.0570**	**0.15**	**1.73 ± 0.81**	**0.0327**	**1.63**	**19.09 ± 4.05**

[a] Unless otherwise specified, yields were quantified on Bruker 500 MHz at 20 °C. [b] Reaction was carried out at 20 °C. [c] Yields were quantified on Bruker 400 MHz. [d] Yields were quantified by LC-QTOF-MS/MS.

Table S3: Reaction condition optimization for the oxidation of DHO with product quantification by ^{1}H qNMR

Entry	Substrate	[substrate] mM	solvent	pH	conditions											base	products quantification						
																	ORO				URA		
					$CuSO_4$ (eq.)	$CuCl_2$ (eq.)	$CoCl_2$ (eq.)	$NiSO_4$ (eq.)	$FeSO_4$ (eq.)	MnO_2 (eq.)	$Na_2S_2O_4$ (eq.)	H_2O_2 (eq.)	NH_4HCO_3 (eq.)	$NaHCO_3$ (eq.)	NaOH (eq.)		CH at 6.080 ppm (1H) integral relative to DMS 6H	[ORO] (mM)	yield of **ORO** (%)[c]	CH at 7.445 ppm (1H) integral relative to DMS 6H	[URA] (mM)	yield of **URA** (%)
1[a]	DHO	37.95	0.1 M pH 8 phosphate buffer	8	0.4	-	-	-	-	5	-	-	-	-	-		0.4293	6.30	16.60	-	-	-
																	0.2888	4.24	11.17	-	-	-
																	0.4936	7.24	19.09	-	-	-
																Average	**0.4039**	**5.93**	**15.62 ± 4.05**	**-**	**-**	**-**
2[a]	DHO	37.95	0.1 M pH 10 phosphate buffer	10	0.4	-	-	-	-	5	-	-	-	-	-		0.6097	8.95	23.58	-	-	-
																	0.6926	10.16	26.79	-	-	-
																	0.6571	9.64	25.42	-	-	-
																Average	**0.6531**	**9.58**	**25.26 ± 1.61**	**-**	**-**	**-**
3[a]	DHO	9.49	saturated $NaHCO_3$	9	0.4	-	-	-	-	5	-	-	-	-	-		0.2863	4.20	44.29	-	-	-
																	0.234	3.43	36.2	-	-	-
																	0.351	5.15	54.31	-	-	-
																Average	**0.2904**	**4.26**	**44.93 ± 9.07**	**-**	**-**	**-**
4[a]	DHO	8.22	water	13	0.4	-	-	-	-	5	-	-	-	-	5.9		-	-	-	-	-	-
																	-	-	-	-	-	-
																	-	-	-	-	-	-
																Average	-	-	-	-	-	-
5[a,d]	DHO	42.38	saturated $NaHCO_3$	9	-	-	-	-	-	-	2	-	-	-	-		1.0306	15.13	35.70	0.0562	0.16	0.38
																	0.8207	12.05	28.43	0.0904	0.26	0.68
																	1.2375	18.16	42.86	0.0430	0.12	0.22
																Average	**1.0296**	**15.11**	**35.66± 7.22**	**0.0632**	**0.18**	**0.42 ± 0.23**
6[a]	DHO	37.95	saturated $NaHCO_3$	9	-	-	-	-	-	-	-	20	-	-	-		0.9047	13.28	35.00	-	-	-
																	0.9777	14.35	37.82	-	-	-
																	0.7737	11.36	29.93	-	-	-
																Average	**0.8854**	**13.00**	**34.25 ± 4.00**	**-**	**-**	**-**
7	DHO	37.95	saturated $NaHCO_3$	9	0.4	-	-	-	-	-	-	20	-	-	-		0.0900	1.32	3.48	0.7338	2.09	5.49
																	0.1083	1.59	4.19	0.5450	1.55	4.08
																	0.0421	0.62	1.63	0.5070	1.44	3.79
																Average	**0.0801**	**1.18**	**3.10 ± 1.32**	**0.5953**	**1.69**	**4.45 ± 0.91**
8	DHO	8.22	water	9	0.4	-	-	-	-	5	-	-	10	-	-		0.4408	1.34	15.21	0.0296	0.08	1.02

																	0.6097	1.73	21.04	0.0099	0.03	0.34
																	0.5053	1.43	17.44	0.0161	0.05	0.55
																Average	**0.5186**	**1.50**	**17.90 ± 2.94**	**0.0185**	**0.05**	**0.64 ± 0.35**
9	DHO	8.22	water	9	0.4	-	-	-	-	5	-	-	-	10	-		0.1422	0.43	4.91	-	-	-
																	0.1800	0.51	6.21	-	-	-
																	0.1400	0.40	4.83	-	-	-
																Average	**0.1541**	**0.45**	**5.32 ± 0.77**	**-**	**-**	**-**
10	DHO	8.22	water	4	0.4	-	-	-	-	5	-	-	-	-	-		1.3860	4.23	51.42	0.0432	0.02	0.36
																	1.3372	4.07	49.61	0.0410	0.02	0.31
																	1.3488	4.11	50.04	0.0719	0.03	0.54
																Average	**1.3573**	**4.14**	**50.36 ± 0.95**	**0.0520**	**0.02**	**0.40 ± 0.12**
11 [b]	DHO	8.22	water	4	0.4	-	-	-	-	5	-	-	-	-	-		0.1430	0.44	4.94	-	-	-
																	0.1462	0.41	5.05	-	-	-
																	0.1357	0.38	4.69	-	-	-
																Average	**0.1416**	**0.41**	**4.89 ± 0.19**	**-**	**-**	**-**
12	DHO	8.22	water	4	-	0.4	-	-	-	5	-	-	-	-	-		1.2734	3.89	47.32	0.0427	0.12	1.47
																	1.0444	3.19	38.66	0.0402	0.13	1.38
																	1.0023	3.06	37.10	0.0442	0.14	1.52
																Average	**1.1067**	**3.38**	**41.03 ± 5.51**	**0.0424**	**0.13**	**1.46 ± 0.07**
13	DHO	8.22	water	4	-	-	0.4	-	-	5	-	-	-	-	-		0.9071	2.77	33.63	0.1004	0.37	4.44
																	0.8706	2.66	32.30	0.1078	0.39	4.70
																	0.8273	2.53	30.69	0.0862	0.31	3.76
																Average	**0.8683**	**2.65**	**32.21 ± 1.47**	**0.0981**	**0.36**	**4.30 ± 0.49**
14	DHO	8.22	water	4	-	-	-	0.4	-	5	-	-	-	-	-		0.8060	2.46	29.90	0.0541	0.15	1.87
																	0.7794	2.37	28.91	0.0517	0.14	1.63
																	0.9103	2.77	33.77	0.0511	0.14	1.61
																Average	**0.8319**	**2.53**	**30.86 ± 2.57**	**0.0523**	**0.14**	**1.70 ± 0.14**
15	DHO	8.22	water	4	-	-	-	-	0.4	5	-	-	-	-	-		0.7459	2.27	27.67	0.0408	0.13	1.59
																	0.6529	1.99	24.22	0.0373	0.12	1.45
																	0.8639	2.63	32.01	0.0346	0.08	0.85
																Average	**0.7542**	**2.30**	**27.97 ± 3.90**	**0.0376**	**0.11**	**1.30 ± 0.39**
16	DHO	8.22	water	4	0.4	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
																	-	-	-	-	-	-
																	-	-	-	-	-	-

																Average	-	-	-	-	-	-
17	DHO	8.22	water	4	-	-	-	-	-	5	-	-	-	-	-		0.3444	1.05	12.78	0.0221	0.06	0.68
																	0.5874	1.79	21.79	0.0612	0.17	1.88
																	0.5421	1.67	20.11	0.0680	0.18	2.79
																Average	**0.4913**	**1.50**	**18.23 ± 4.79**	**0.0504**	**0.14**	**1.78 ± 1.06**
18	DHO	20.87	water	4	0.4	-	-	-	-	5	-	-	-	-	-		2.2872	6.98	33.43	0.1430	0.49	2.33
																	2.1693	6.62	31.70	0.1850	0.60	2.90
																	1.9674	6.00	28.75	0.2733	0.89	4.28
																Average	**2.1413**	**6.53**	**31.29 ± 2.37**	**0.2004**	**0.66**	**3.17 ± 1.00**

a Yields were quantified by ^{1}H qNMR on Bruker 500 MHz at 0 °C. ***b*** Reaction was carried out at 20 °C. ***c*** Unless otherwise mentioned, reactions were quantified by^{1}H qNMR on Bruker 400 MHz at at 20 °C. ***d*** Reaction was done under UV irradiation at 15 °C.

Table S4: Reaction condition optimization for the oxidation of HAA with product quantification by LC-QTOF-MS/MS and ^{1}H qNMR

Entry	Substrate	[substrate] mM	solvent	pH	conditions			base		products quantification						
										product 1	ORO			URA		
					$CuSO_4$ (eq.)	MnO_2 (eq.)	H_2O_2 (eq.)	NH_4HCO_3 (eq.)	NaOH (eq.)	CH at 5.455 ppm (1H) integral relative to DMS 6H	CH at 6.080 ppm (1H) integral relative to DMS 6H	[ORO] (mM)	yield of ORO (%)[a]	LC-QTOF Response relative to maleic acid	[URA] (mM)	Yield of URA (%)
1	HAA	8.22	saturated $NaHCO_3$	9	-	-	20	-	-	-	-	-	-	-	-	-
										-	-	-	-	-	-	-
										-	-	-	-	-	-	-
										Average -	-	-	-	-	-	-
2	HAA	8.22	water	4	0.4	5	-	-	-	0.3827	-	-	-	44049.8823	0.39	4.78
										0.3603	-	-	-	26970.1742	0.24	2.93
										0.4019	-	-	-	70906.5050	0.63	7.69
										Average **0.3816 ± 0.0209**	-	-	-	**47308.8538**[b]	**0.42**	**5.13 ± 2.40**
3[c]	HAA	8.22	water	13	0.4	5	-	-	5.9	-	0.0557	0.17	2.07	-	-	-
										-	0.0540	0.16	2.00	-	-	-
										-	0.0514	0.15	1.90	-	-	-
										Average -	**0.0537**	**0.16**	**1.99 ± 0.08**	-	-	-
4[d]	HAA	8.22	water	13	0.4	5	-	-	5.9	0.0550	0.2934	0.89	10.89	-	-	-
										0.0614	0.2794	0.85	10.37	-	-	-
										0.0611	0.3137	0.96	11.64	-	-	-
										Average **0.0592 ± 0.0006**	**0.2955**	**0.90**	**10.97 ± 0.64**	-	-	-
5[d]	HAA	20.87	water	13	0.4	5	-	-	5.9	0.2783	0.5839	1.78	8.53	-	-	-
										0.2900	0.6265	1.91	9.16	-	-	-
										0.3400	0.6527	1.99	9.54	-	-	-
										Average **0.3028 ± 0.0339**	**0.6210**	**1.89**	**9.08 ± 0.51**	-	-	-

[a] Reactions were quantified by ^{1}H qNMR on Bruker 400 MHz at ambient temperature. [b] Reactions were quantified by LC-QTOF. [c] All reagents were added in pot from t = 0 h. [d] Reagents were added separately, NaOH was added at t = 16 h.

Table S5: Reaction condition optimization for the one-pot synthesis of ORO with product quantification by LC-QTOF-MS/MS and ^{1}H qNMR

Entry	Substrate	[substrate] mM	solvent	conditions			base			products detected in the mixture post-reaction (%)						
										ORO				URA		product 1
				$FeCl_3$ (eq.)	$CuSO_4$ (eq.)	MnO_2 (eq)	NH_4HCO_3 (eq)	$NaHCO_3$ (eq.)	NaOH (eq.)	[ORO] response relative to internal standard [a]	[ORO] (mM)	yield of ORO (%) [b]	CH at 7.445 ppm (1H) integral relative to DMS 6H	[URA] (mM)	yield of URA (%) [b]	CH at 5.455 ppm (1H) integral relative to DMS 6H
1	CAA	8.52	water	5	0.4	5	10	-	-	0.0076	0.02	0.23	0.0200	0.06	0.56	-
										0.0082	0.03	0.24	0.0057	0.02	0.17	-
										0.0125	0.05	0.37	0.0125	0.04	0.37	-
										Average	**0.0094**	**0.03**	**0.28 ± 0.08**	**0.0127**	**0.04**	**0.37 ± 0.20**	**-**
2 [c]	CAA	8.52	water	-	0.4	5	-	-	5.9	-	-	-	-	-	-	-
										-	-	-	-	-	-	-
										-	-	-	-	-	-	-
										Average	-	-	-	-	-	-	-
3	CAA	8.52	water	-	0.4	5	-	-	5.9	0.0398	0.12	1.19	-	-	-	0.0061
										0.0504	0.16	2.30	-	-	-	0.0118
										0.0555	0.18	2.53	-	-	-	0.0047
										Average	**0.0486**	**0.15**	**2.01 ± 0.72**	-	-	-	**0.0075**
4	CAA	8.52	0.1 M HCl	-	0.4	5	-	-	9.5	0.0532	0.16	1.59	-	-	-	-
										0.0502	0.15	1.50	-	-	-	-
										0.0745	0.23	3.22	-	-	-	-
										Average	**0.0593**	**0.18**	**2.10 ± 0.97**	-	-	-	-
5	CAA	21.01	0.1 M HCl	-	0.4	5	-	-	9.5	0.8350	2.54	12.12	0.0239	0.07	0.33	0.0134
										0.8999	2.74	13.06	0.011	0.03	0.15	0.0164
										0.8418	2.58	12.23	0.0301	0.08	0.42	0.0022
										Average	**0.8589**	**2.62**	**12.47 ± 0.51**	**0.0217**	**0.06**	**0.30 ± 0.14**	**0.0107**
										91058359.7700	4.23	20.14	2601.6435	0.04	0.36	-
										118827930.4800	5.52	26.29	2086.5798	0.03	0.29	-
										114953106.6600	5.34	25.45	1983.2408	0.03	0.28	-
										Average	**108279798.9700** [d]	**5.03**	**23.96 ± 3.33**	**2223. 8214** [d]	**0.03**	**0.31 ± 0.05**	-
6	ASP+CAP	21.01	Water /0.1 M HCl	-	0.4	5	-	2	9.5	0.0970	0.28	1.31	-	-	-	-
										0.0759	0.25	1.24	-	-	-	-
										0.0990	0.33	1.62	-	-	-	-
										Average	**0.0906**	**0.29**	**1.39 ± 0.20**	-	-	-	-
										14638223.3200	0.68	3.24	-	-	-	-
										18943583.1200	0.88	4.19	-	-	-	-

		18082511.1600	0.84	4.00	-	-	-	-
	Average	**17221439.2000** [d]	**0.80**	**3.81 ± 0.50**	-	-	-	-

[a] Unless otherwise specified, reactions were quantified by ^{1}H qNMR on Bruker 400 MHz at ambient temperature. [b] Reported yields are an average of at least 3 runs. [c] All reagents are added together from the beginning at t = 0 h. [d] Yields were quantified by LC-QTOF-MS/MS.

Table S6: Reaction condition optimization for the one-pot synthesis of ORO with product quantification by LC-QTOF-MS/MS and ^{1}H qNMR

Entry	time (h)	ASP (mM)	CAP (eq.)	$NaHCO_3$ (eq.)	$CuSO_4$ (eq.)	MnO_2 (eq.)	HCl (eq.)	NaOH (eq.)		LC-QTOF-MS/MS response relative to maleic acid [a]	[ORO] (mM)	yield of ORO (%)	CH at 2.35 ppm (1H) integral relative to DMS (6H) [b]	[CAA] (mM)	yield of CAA (%)
1	t=0	21	2	2	0.4	5	-	4.8		-	-	-	1.1504	3.90	18.59
	t=16	-	-	-	-	-	-	-		-	-	-	0.9607	3.26	15.52
	t=32	-	-	-	-	-	-	-		-	-	-	1.1100	3.87	13.57
	t=48	-	-	-	-	-	-	-	**Average**	**-**	**-**	**-**	**1.0737**	**3.68**	**15.89 ± 2.53**
2	t=0	21	2	2	0.4	5	-	-		1086539.657	0.05	0.24	0.9607	3.26	15.52
	t=16	-	-	-	-	-	-	4.8		961794.9809	0.04	0.21	0.9434	3.20	15.24
	t=32	-	-	-	-	-	-	-		785507.7047	0.04	0.17	1.0029	3.50	12.31
	t=48	-	-	-	-	-	-	-	**Average**	**944614.1141**	**0.04**	**0.21 ± 0.03**	**0.9690**	**3.32**	**14.36 ± 1.78**
3	t=0	21	2	2	-	-	-	-		-	-	-	1.0809	3.67	17.47
	t=16	-	-	-	0.4	5	-	4.8		-	-	-	1.0737	3.64	17.35
	t=32	-	-	-	-	-	-	-		-	-	-	1.2813	4.45	16.01
	t=48	-	-	-	-	-	-	-	**Average**	**-**	**-**	**-**	**1.1453**	**3.92**	**16.94 ± 0.81**
4	t=0	21	2	2	-	-	-	-		1086539.657	0.05	0.24	3.0034	10.19	48.53
	t=16	-	-	-	0.4	-	-	-		1682971.998	0.08	0.37	2.6154	8.87	42.26
	t=32	-	-	-	-	5	-	4.8		2682362.7529	0.12	0.59	3.8183	13.06	48.05
	t=48	-	-	-	-	-	-	-	**Average**	**1817291.4690**	**0.08**	**0.40 ± 0.18**	**3.1457**	**10.71**	**46.28 ± 3.49**
5	t=0	21	2	2	-	-	-	-		3599344.5821	0.17	0.80	0.5988	2.03	9.68
	t=16	-	-	-	0.4	-	-	-		4224594.3847	0.20	0.93	0.7261	2.46	11.73
	t=32	-	-	-	-	5	-	-		3749275.1772	0.18	0.82	0.7190	2.55	12.04
	t=48	-	-	-	-	-	-	4.8	**Average**	**3857738.0480**	**0.18**	**0.85 ± 0.07**	**0.6813**	**2.35**	**11.15 ± 1.29**
6	t=0	21	2	2	-	-	-	-		4685025.7517	0.22	1.04	3.6659	12.44	59.24
	t=16				-	-	4.8	-		5200546.5560	0.24	1.15	3.3978	11.53	54.90
	t=32	-	-	-	-	5	-	-		5289091.1471	0.26	1.23	4.1611	14.22	67.42
	t=48	-	-	-	-	-	-	4.8	**Average**	**5058221.1516**	**0.24**	**1.14 ± 0.09**	**3.7416**	**12.73**	**60.52 ± 6.36**
7	t=0	21	2	2	-	-	-	-		14638223.3200	0.68	3.24	1.4229	4.83	17.78
	t=16				0.4	-	4.8	-		18943583.1200	0.88	4.19	1.4331	4.86	17.91
	t=32	-	-	-	-	5	-	-		18082511.1600	0.84	4.00	1.9974	6.87	24.91
	t=48	-	-	-	-	-	-	9.5	**Average**	**17221439.2000**	**0.80**	**3.81 ± 0.50**	**1.6178**	**5.52**	**20.20 ± 4.08**

[a] Reactions were quantified by LC-QTOF-MS/MS. [b] Reactions were quantified by ^{1}H qNMR on Bruker 400 MHz at ambient temperatu.

APPENDIX

A. Representative ^{1}H NMR spectrum and LC-QTOF-MS chromatograms

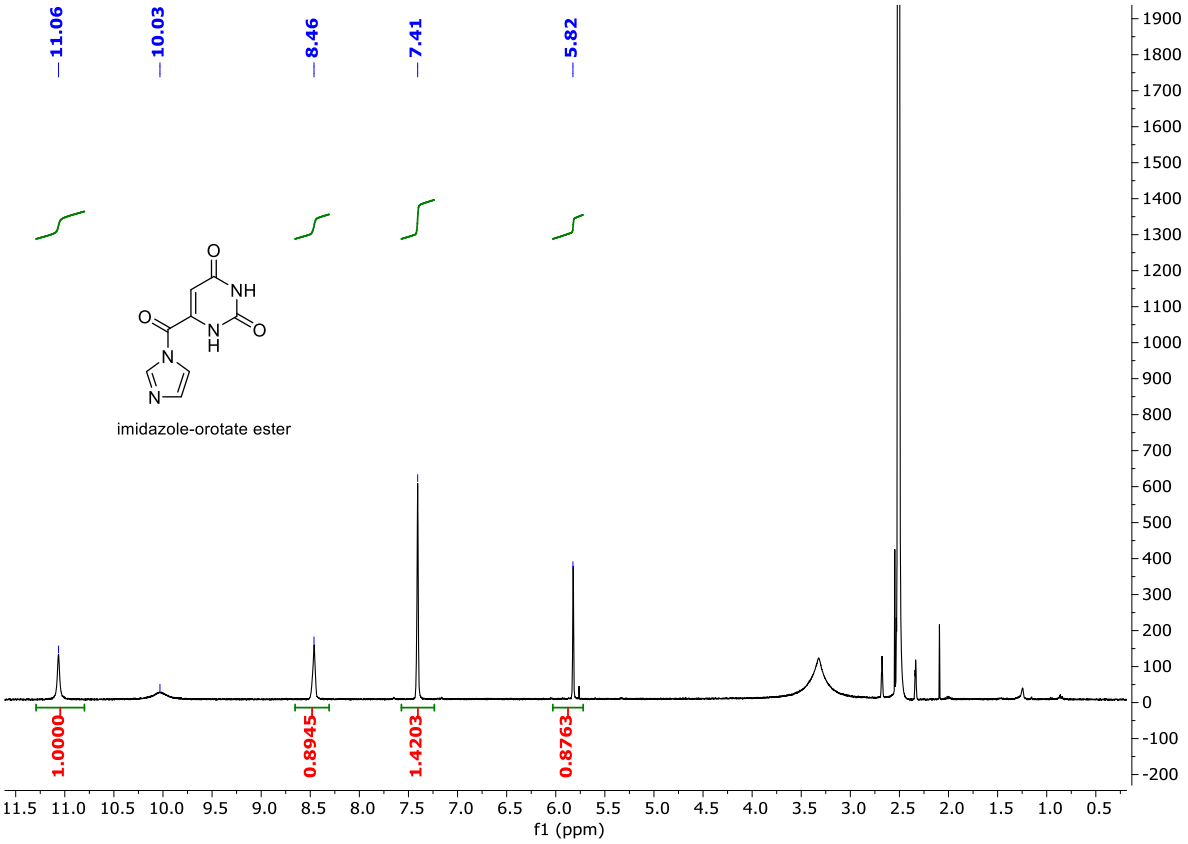
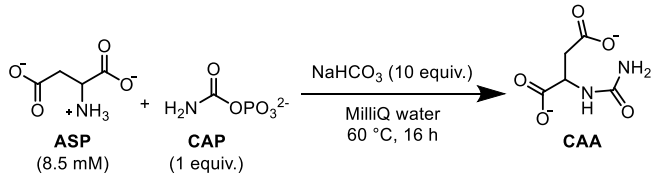


Table S1 Entry 1



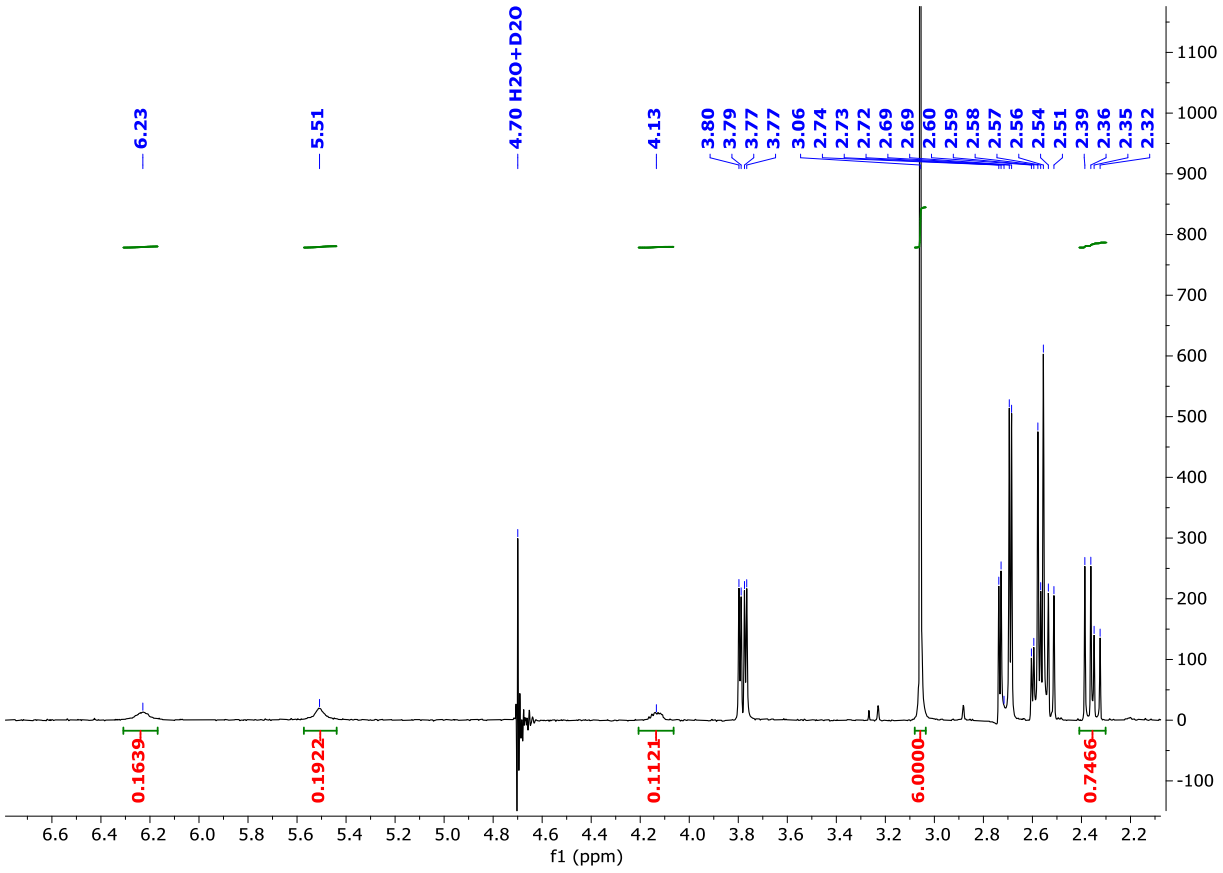
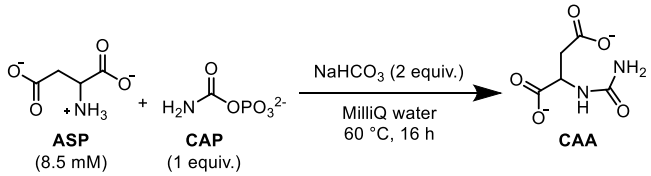


Table S1 Entry 2



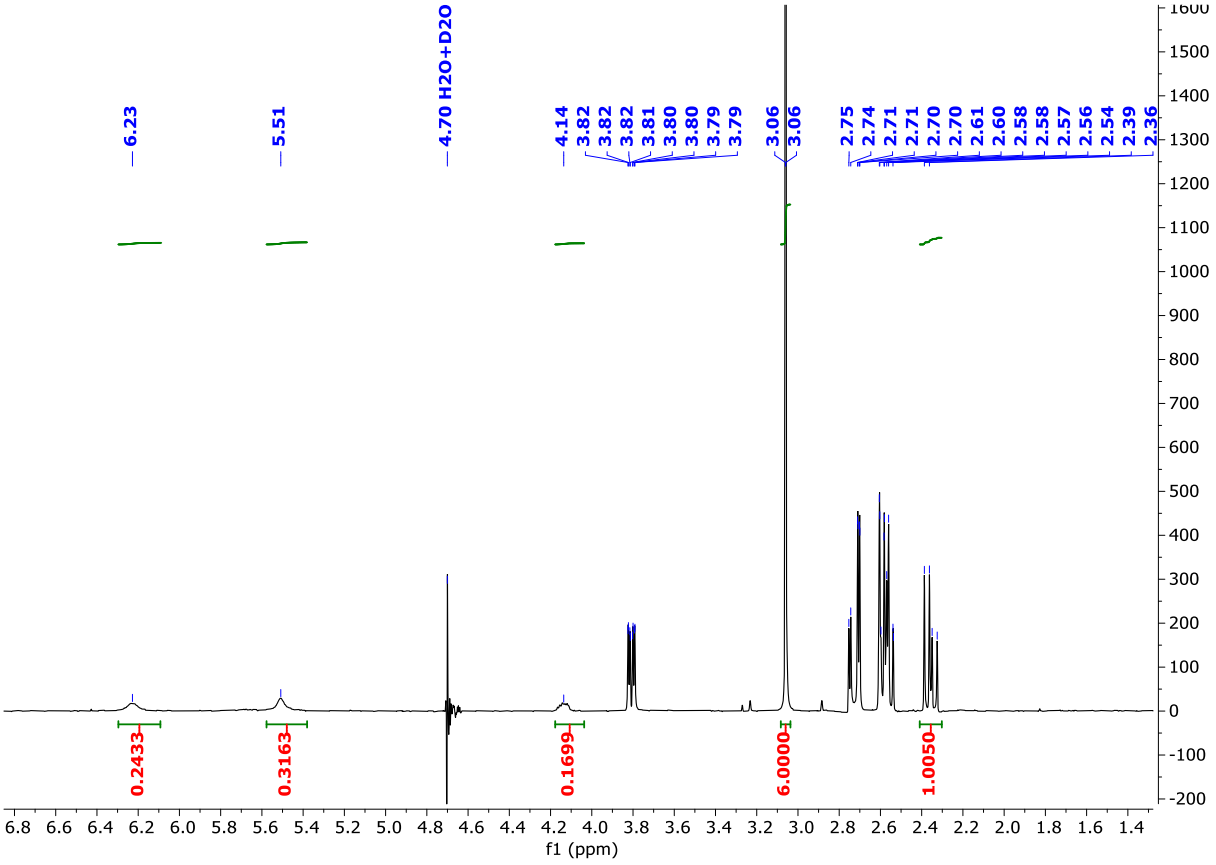
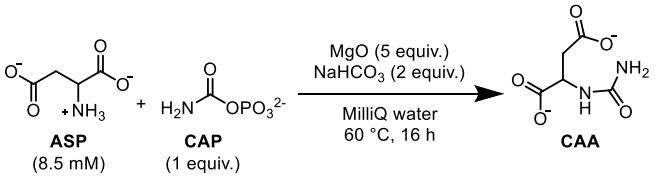


Table S1 Entry 3



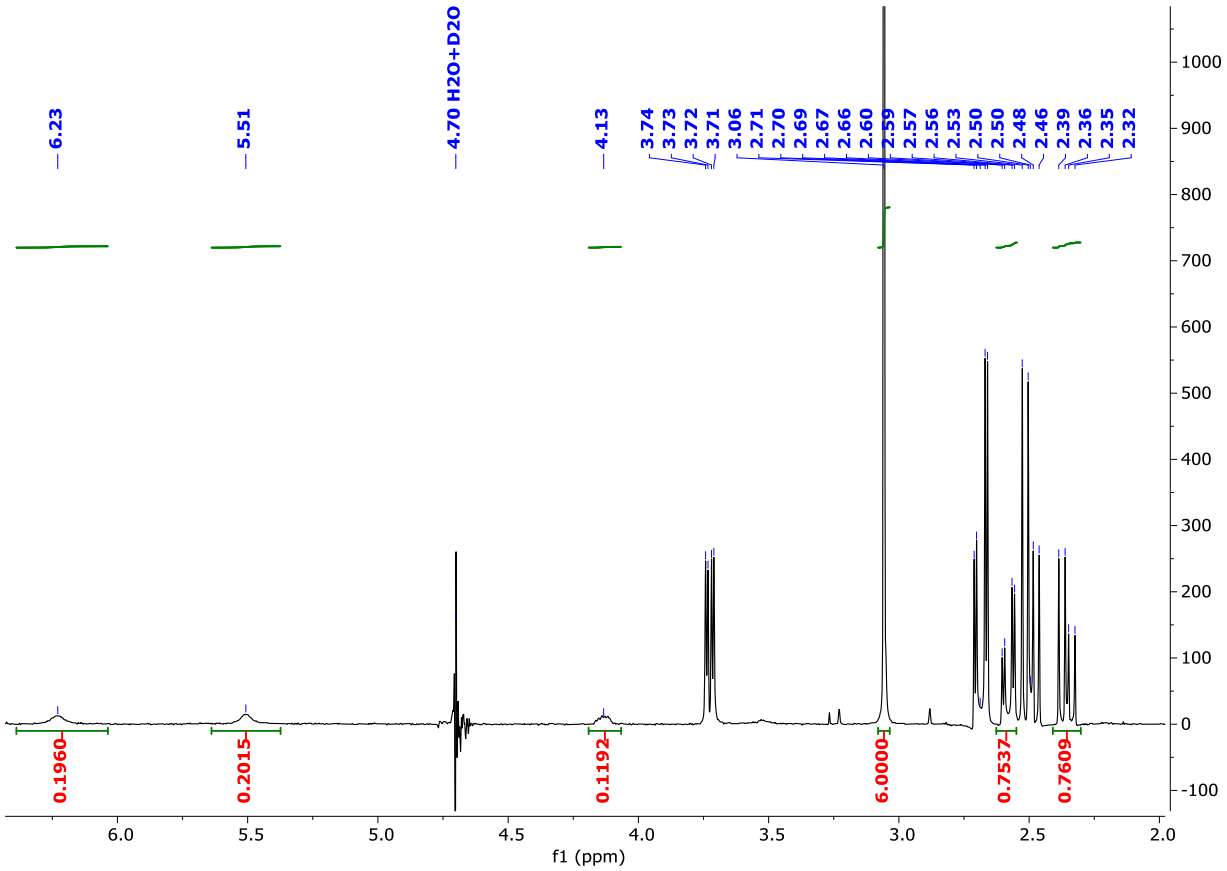
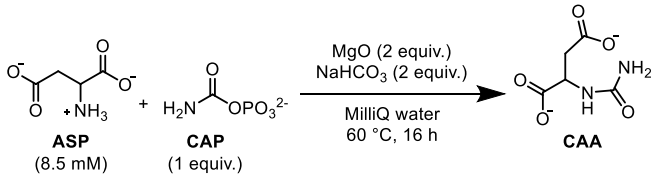


Table S1 Entry 4



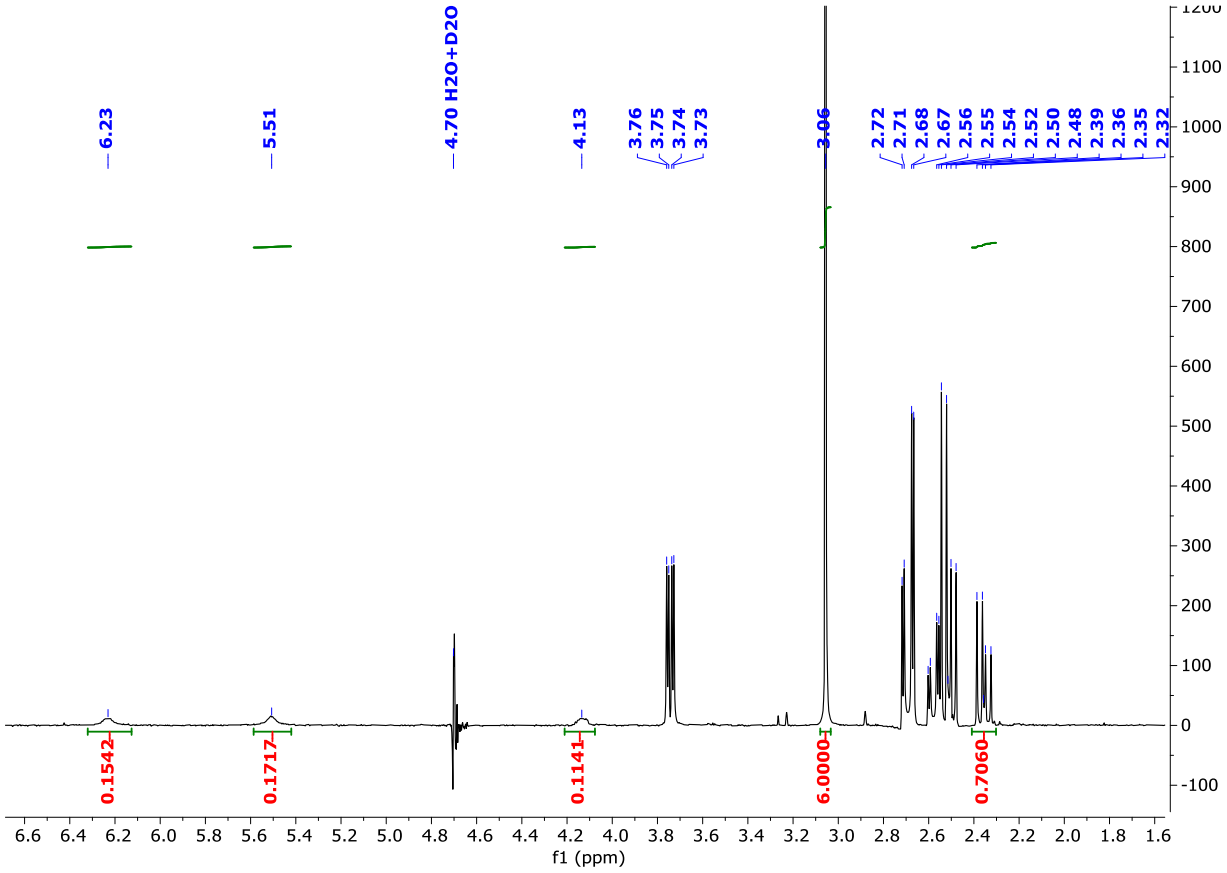
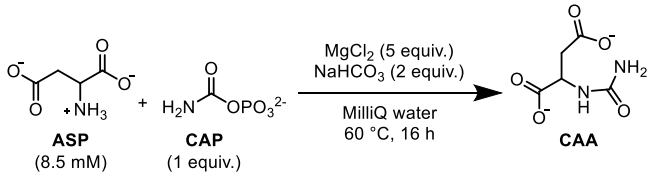


Table S1 Entry 5



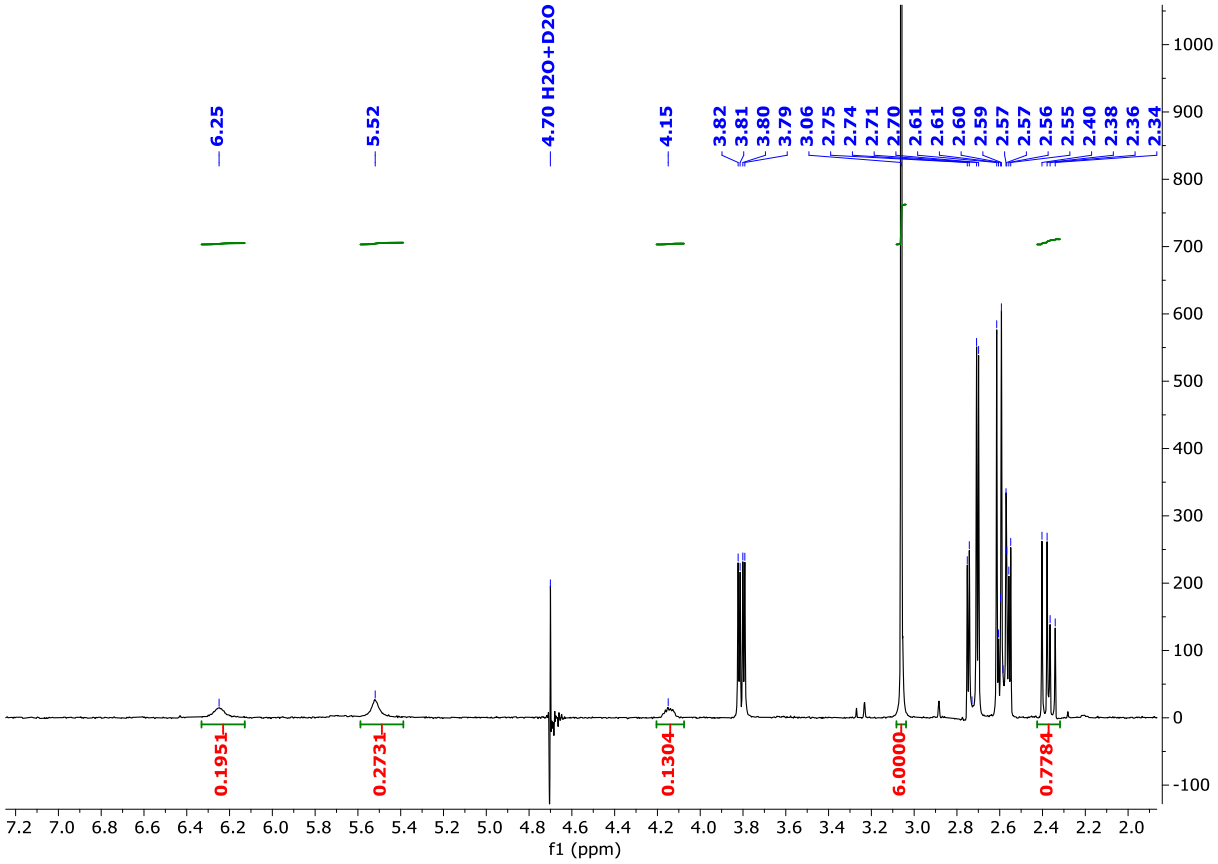


Table S1 Entry 6

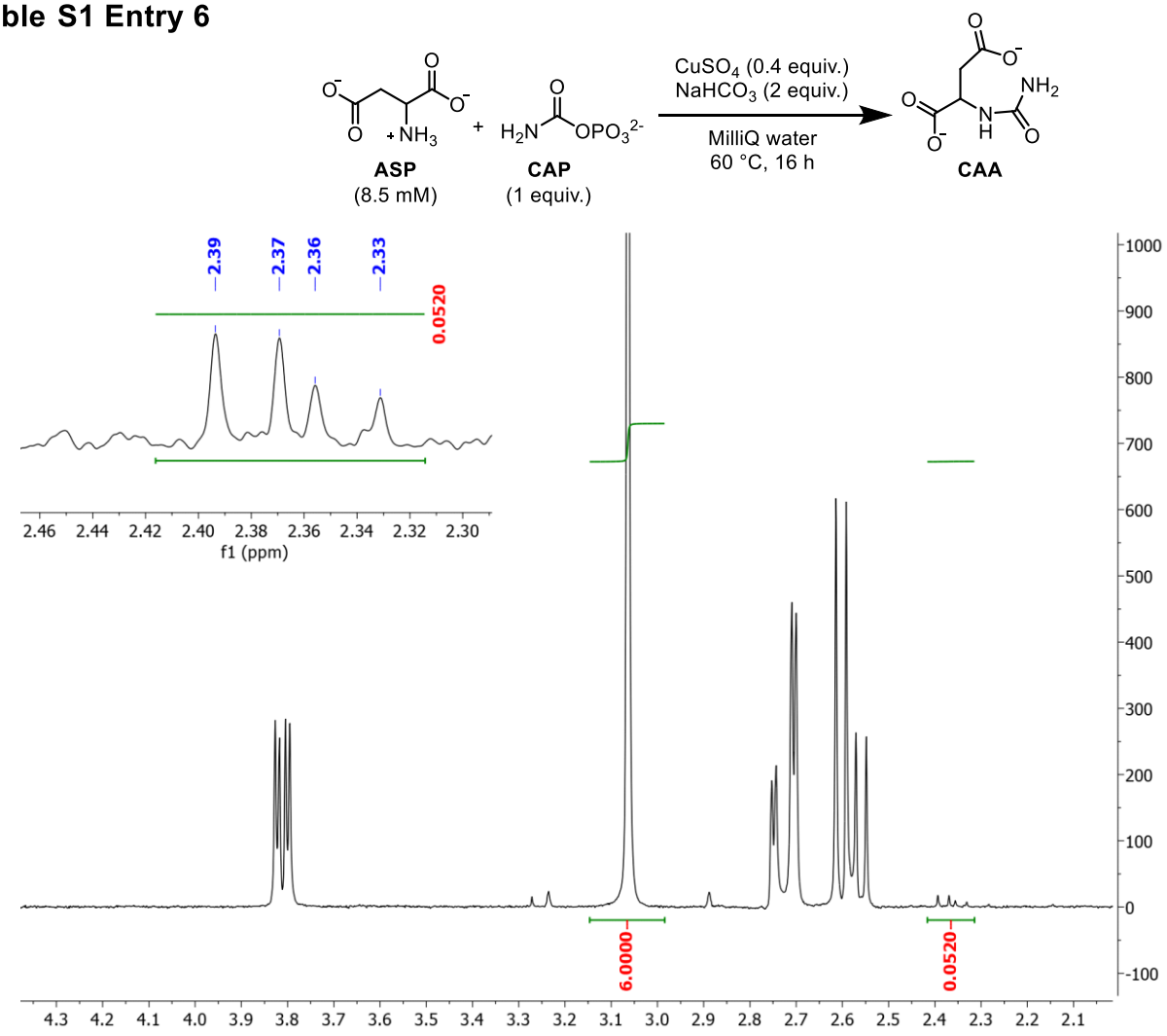
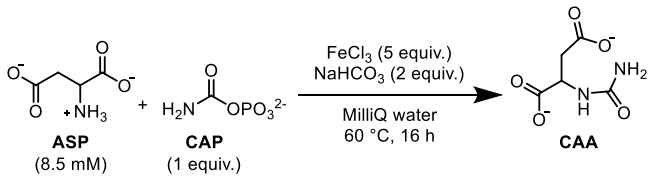


Table S1 Entry 7



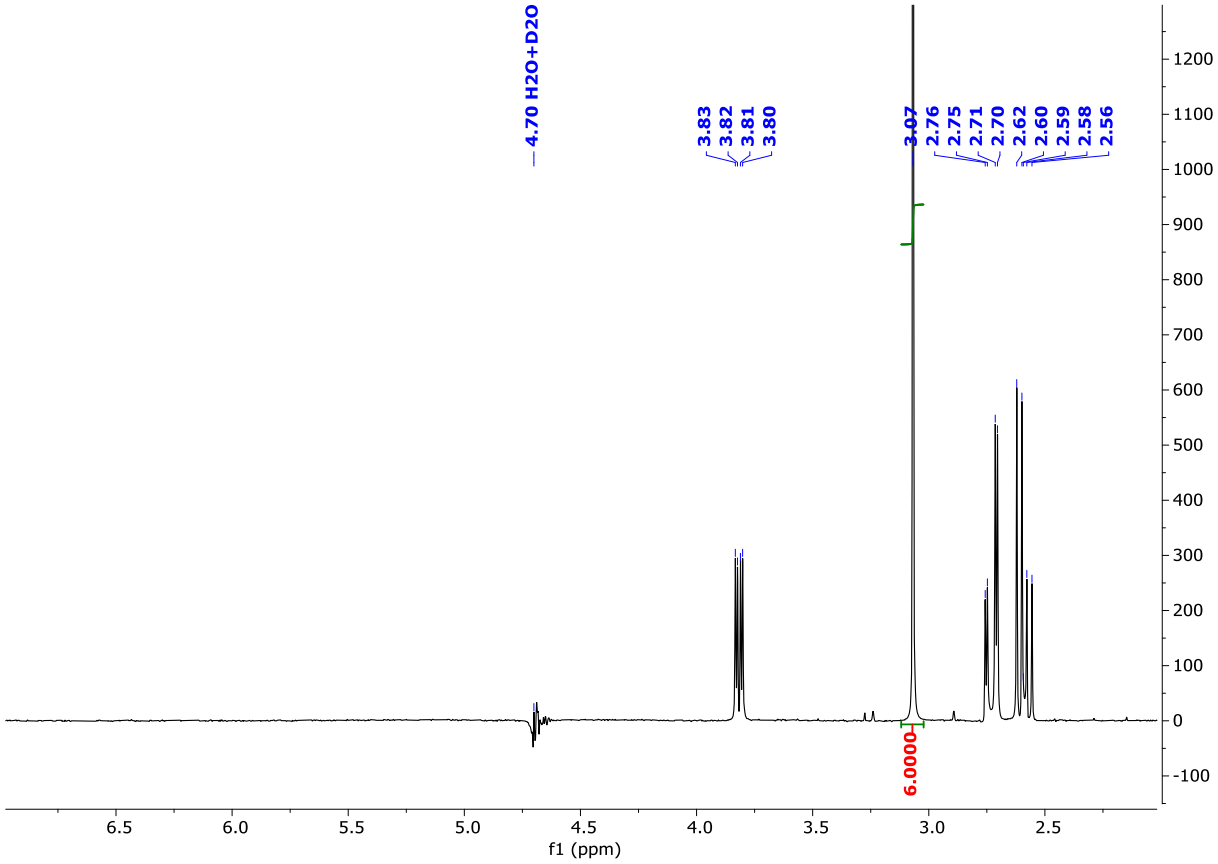


Table S1 Entry 8

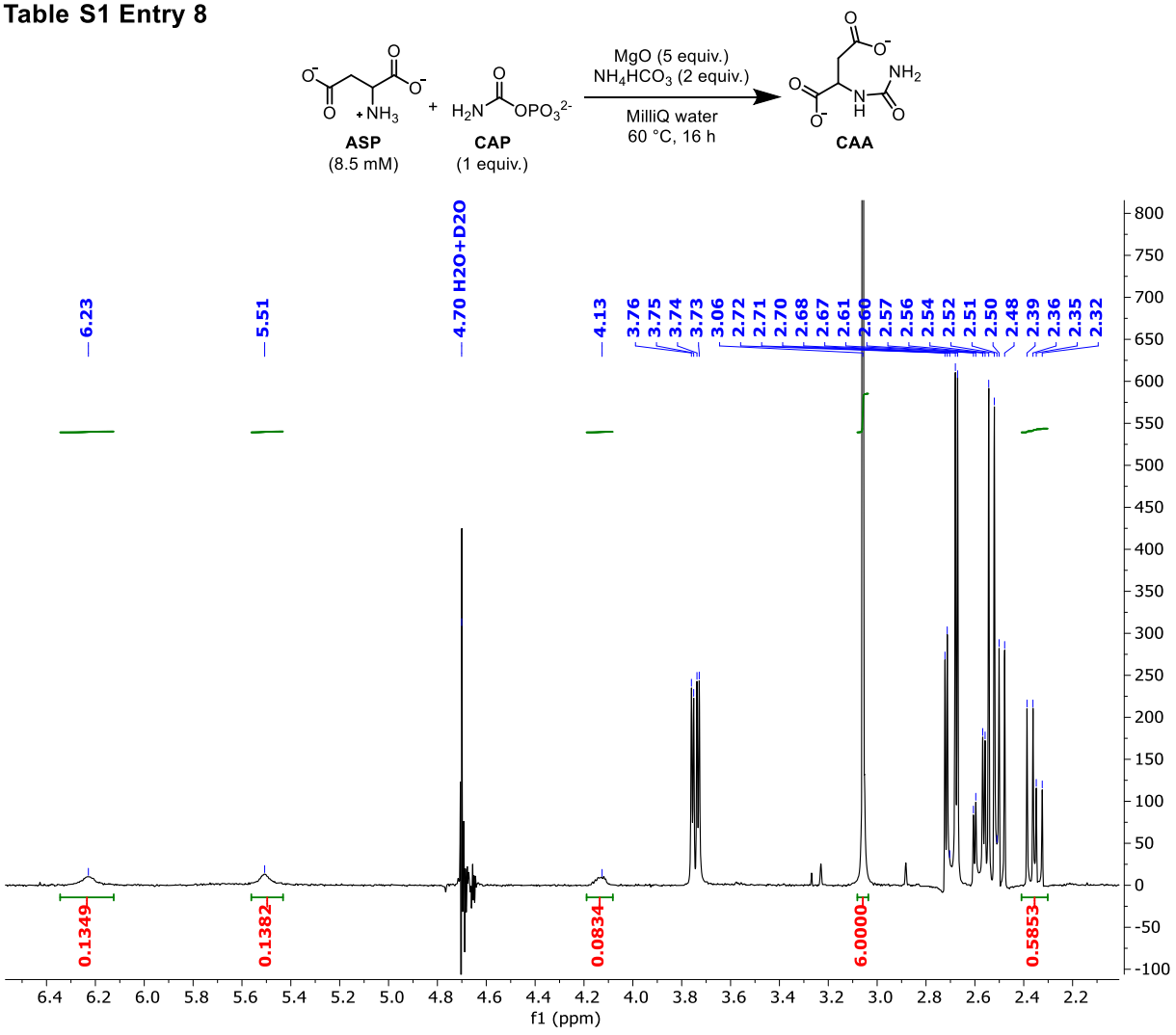
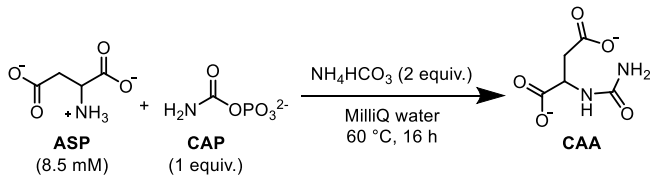


Table S1 Entry 9



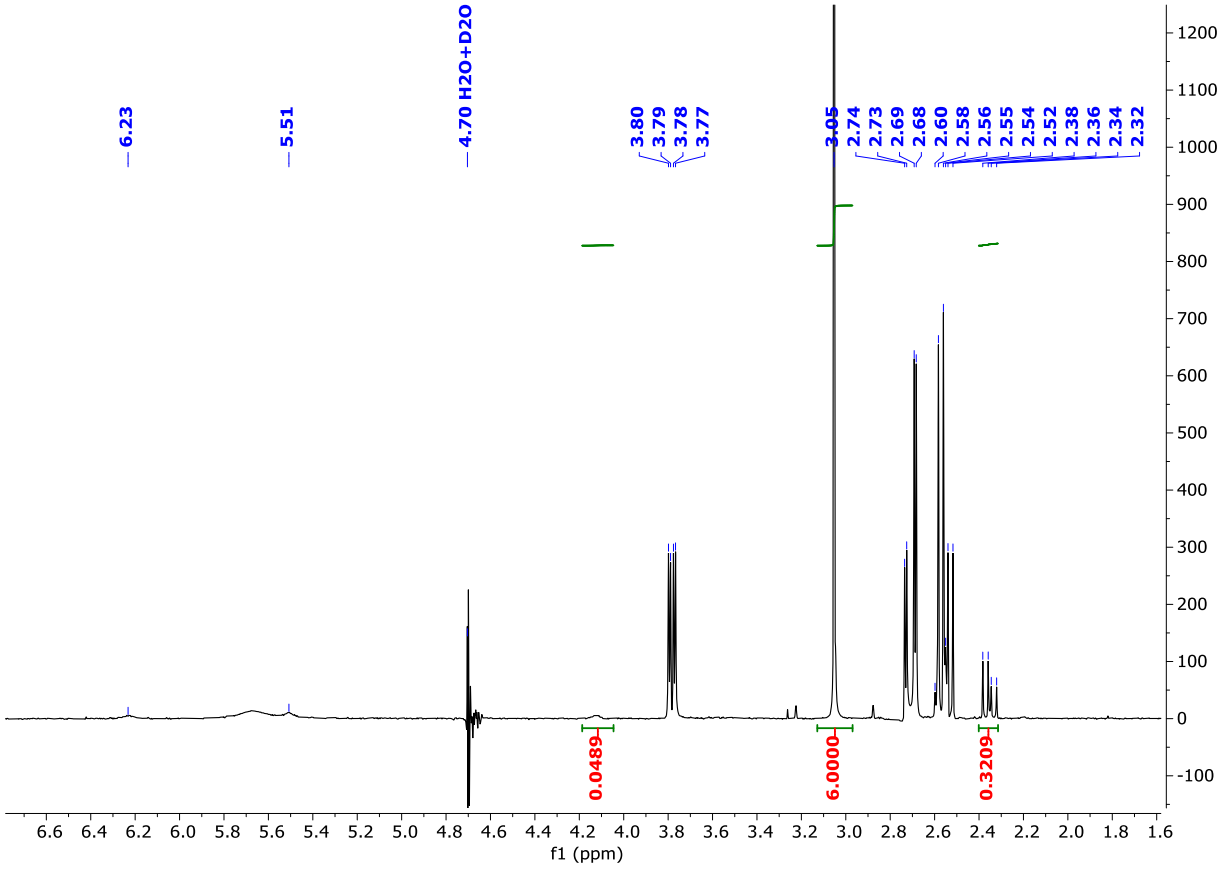
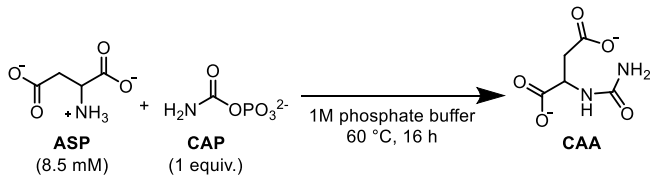


Table S1 Entry 10



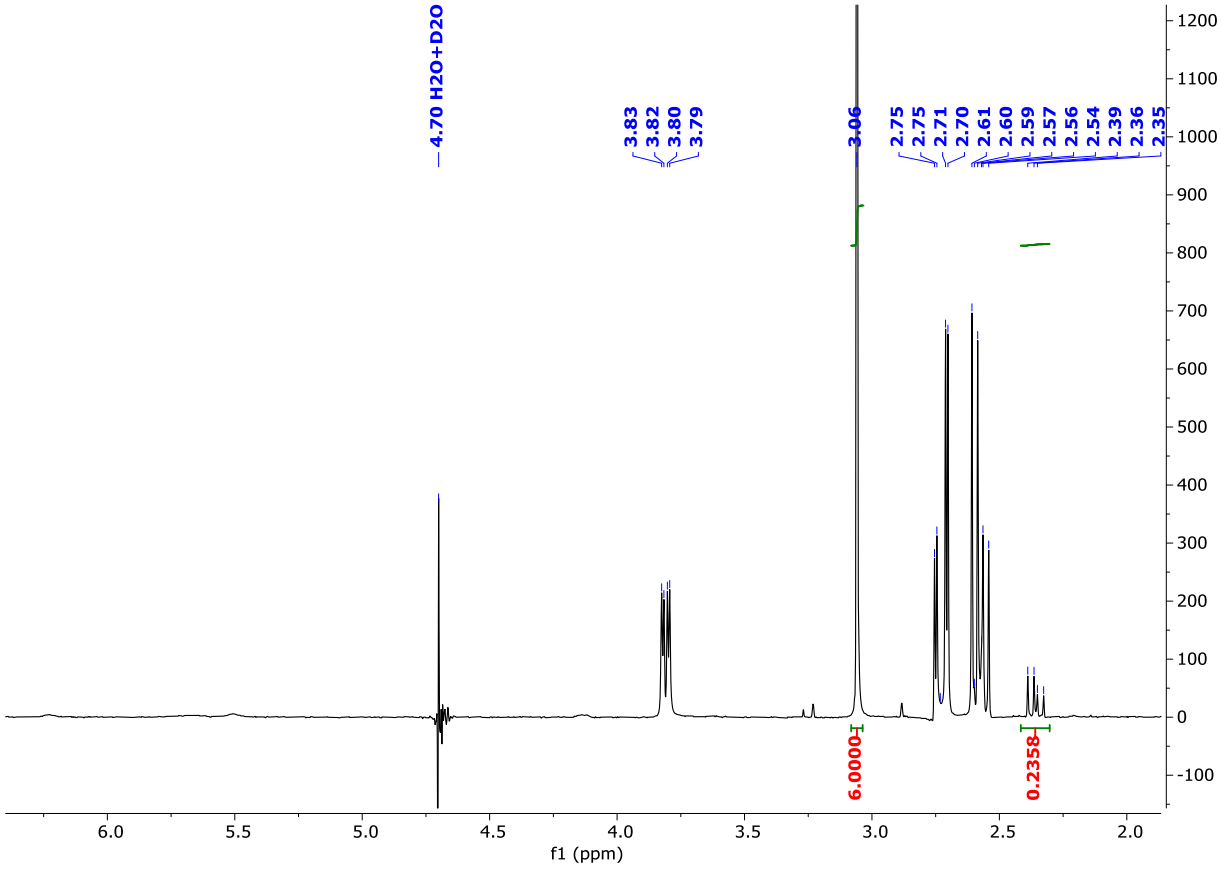


Table S1 Entry 11

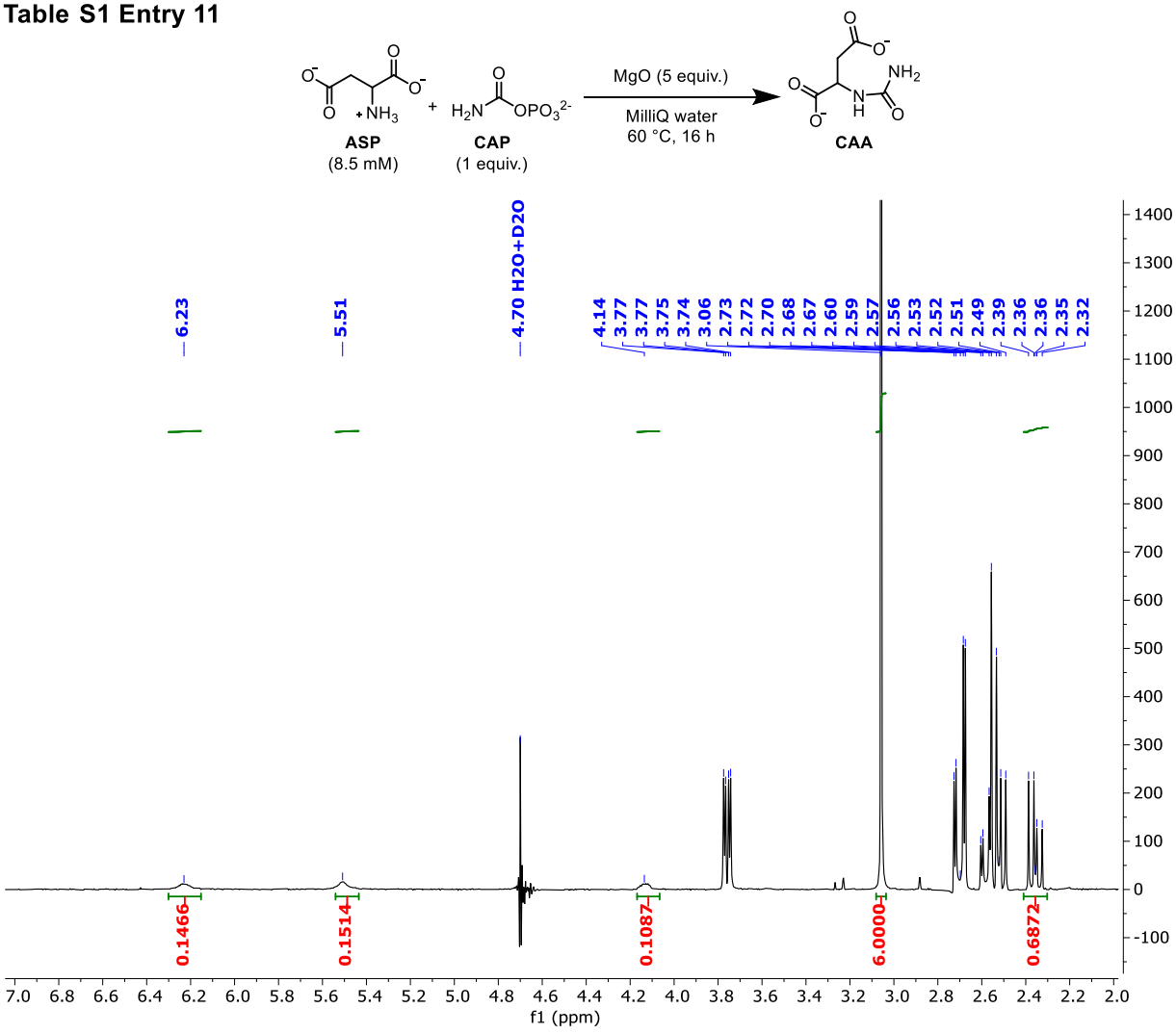
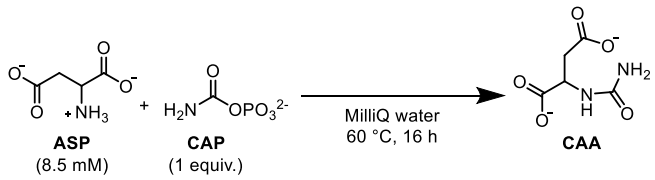


Table S1 Entry 12



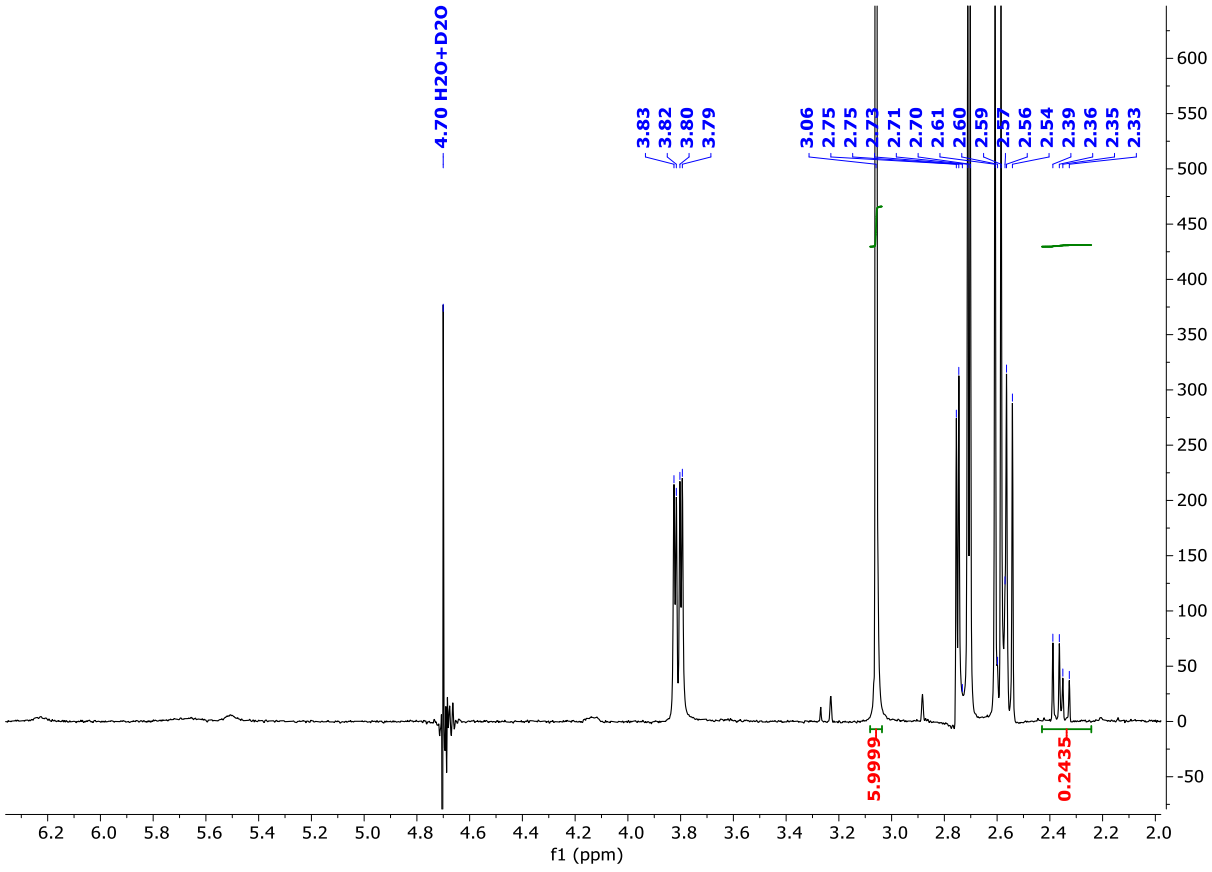
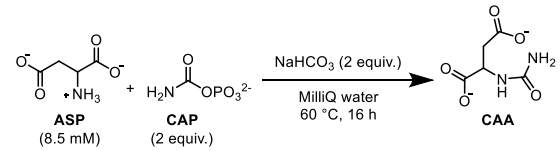
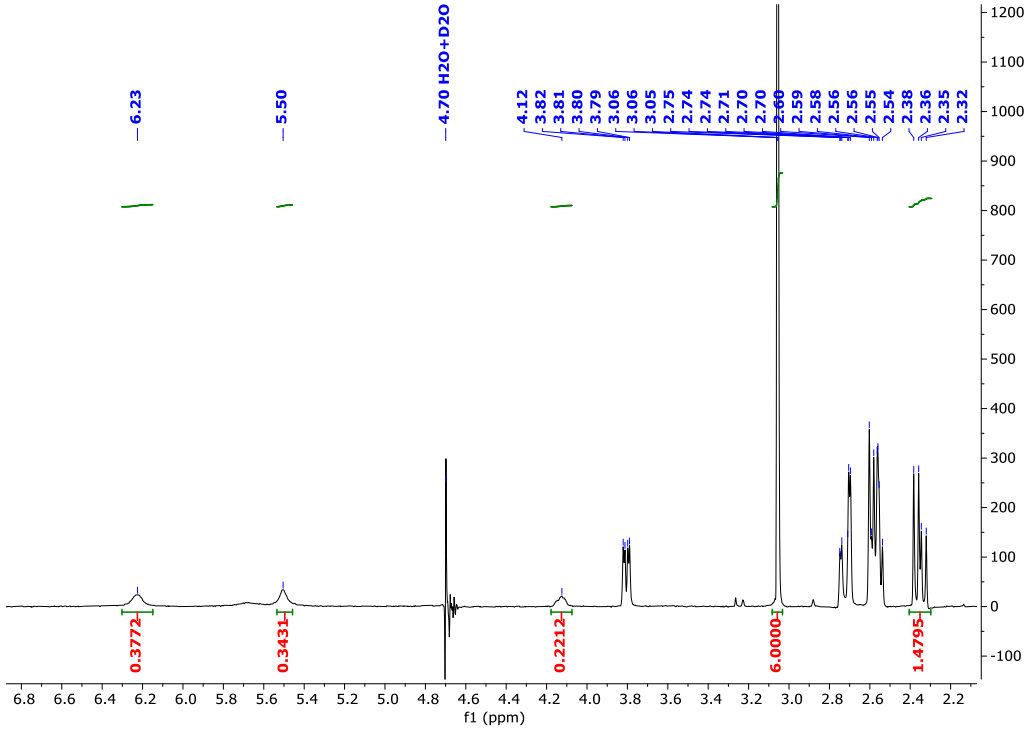
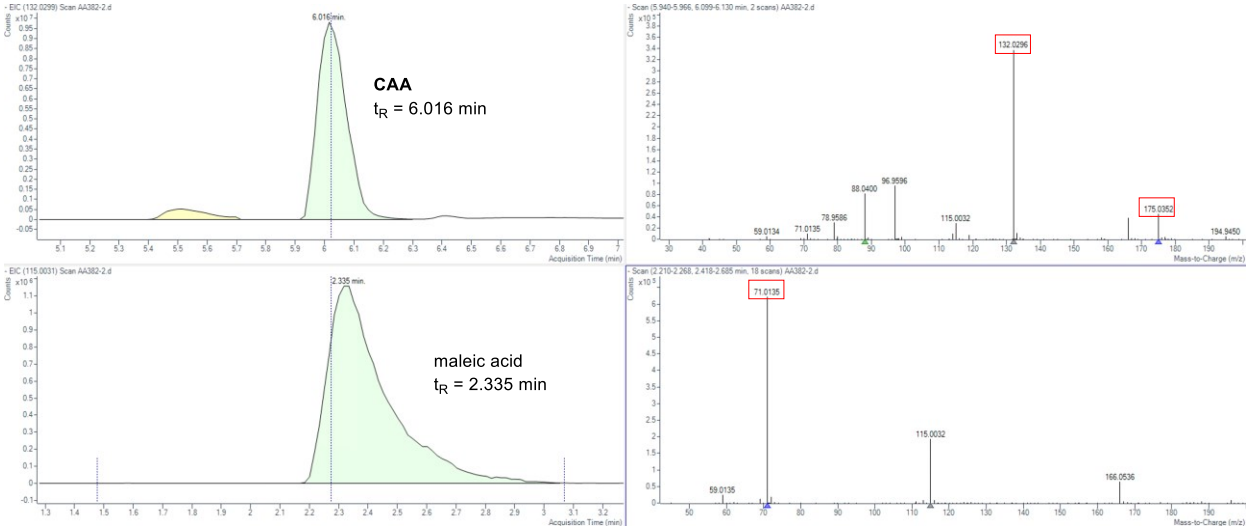


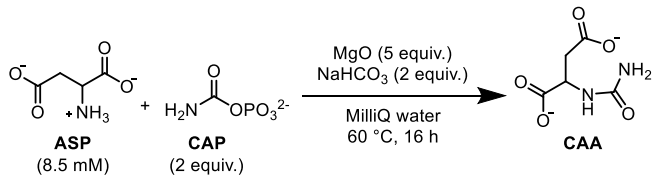
Table S1 Entry 13







S1 Entry 14


MgO (5 equiv.)
NaHCO3 (2 equiv.)
MilliQ water
60 °C, 16 h
ASP
(8.5 mM)
CAP
(2 equiv.)
CAA

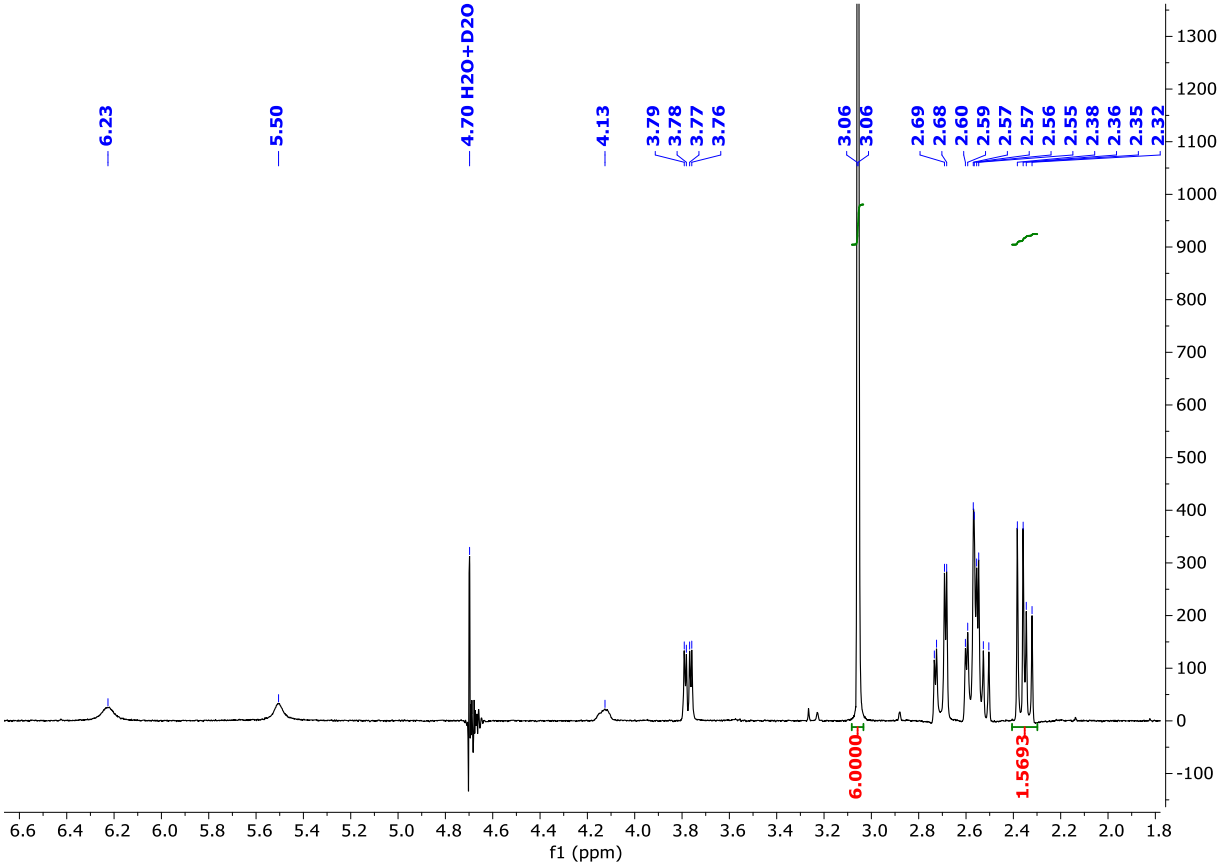
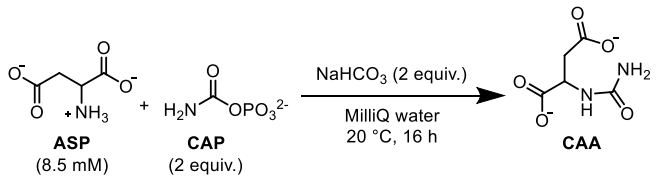

6.23
5.50
4.70 H2O+D2O
4.13
3.79
3.78
3.77
3.76
3.06
3.06
2.69
2.68
2.60
2.59
2.57
2.57
2.56
2.55
2.38
2.36
2.35
2.32
6.0000
1.5693
f1 (ppm)

Table S1 Entry 15



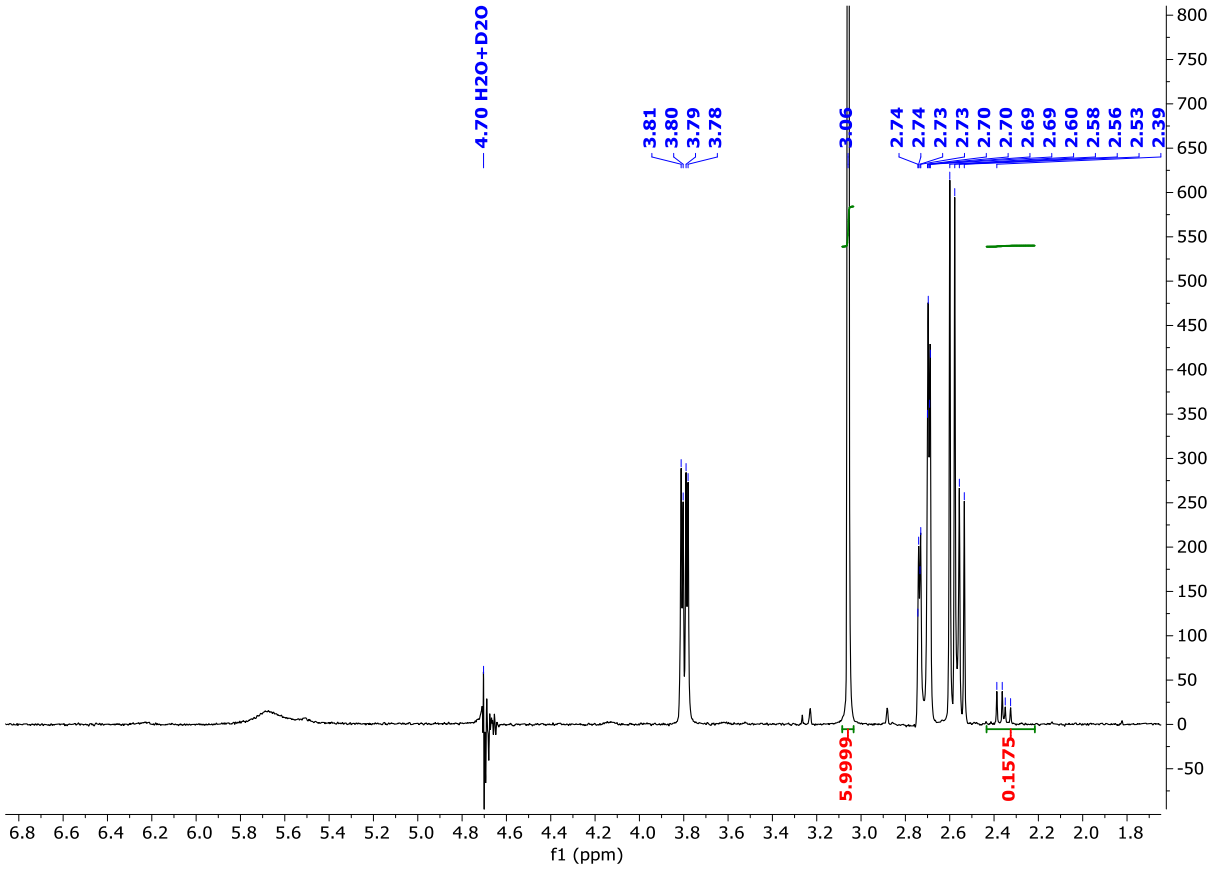
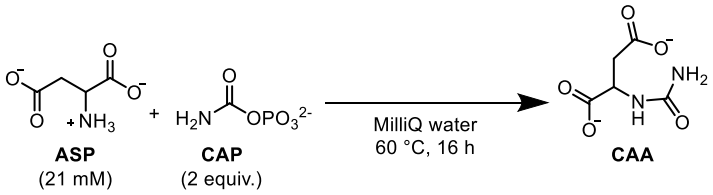


Table S1 Entry 16



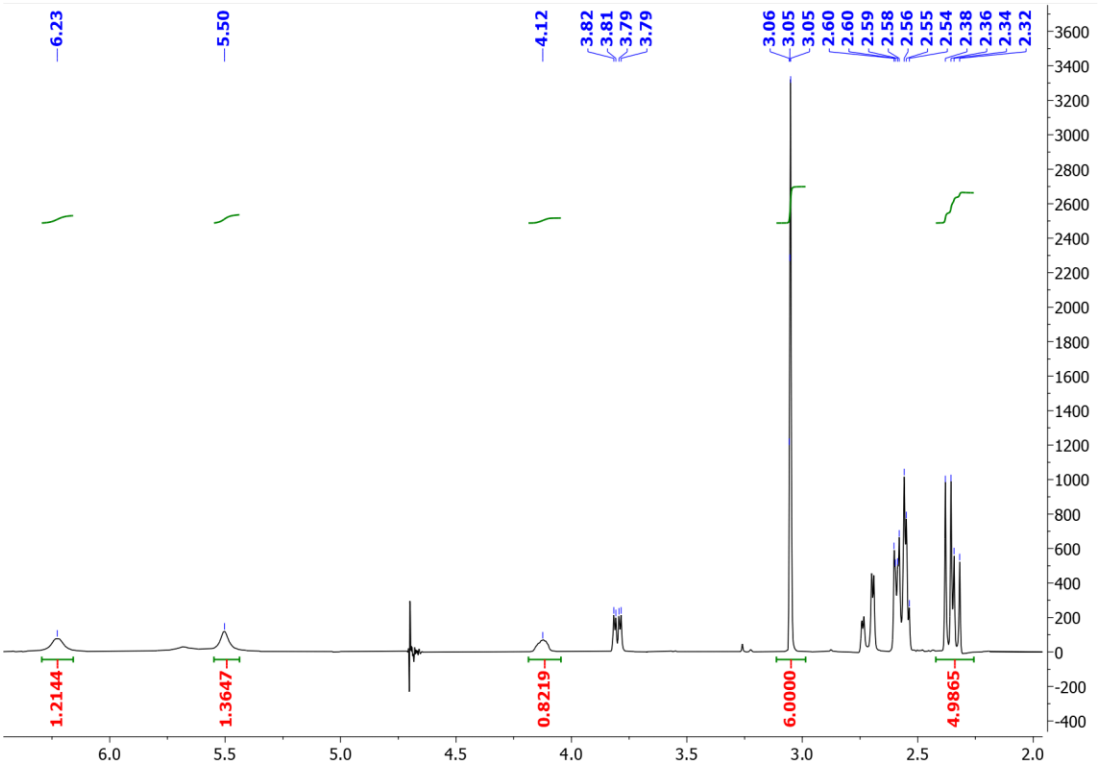
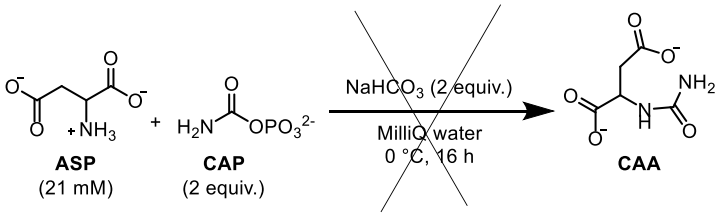


Table S1 Entry 17



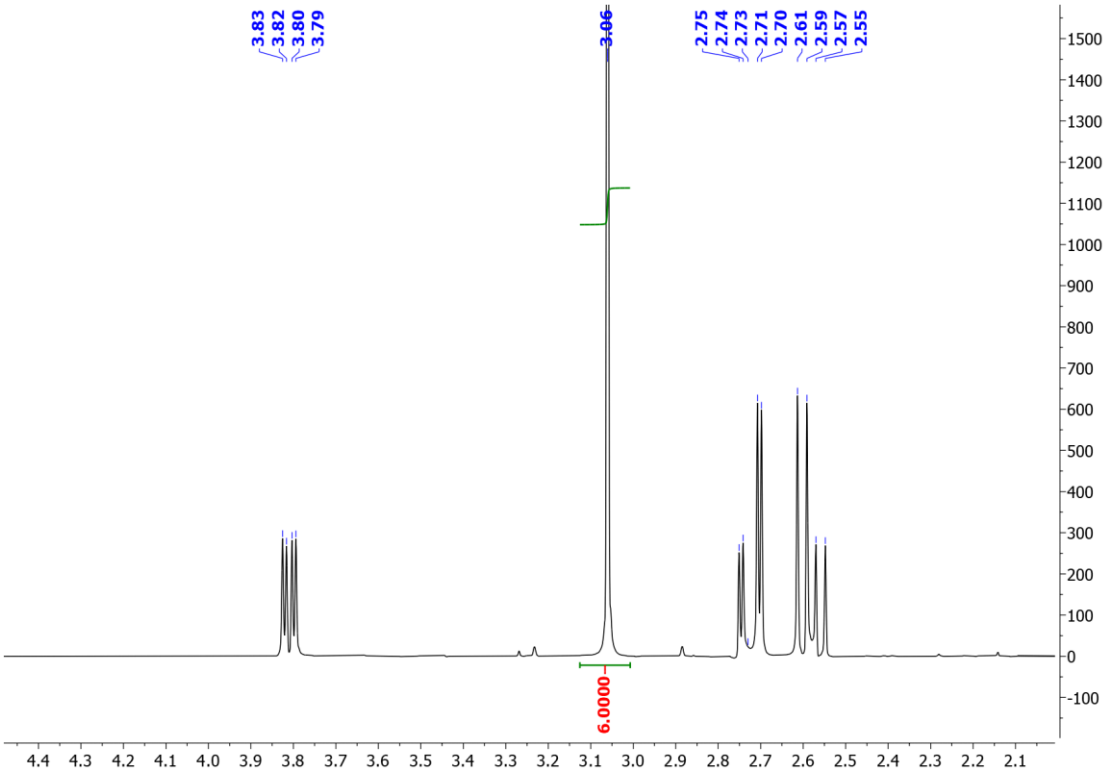
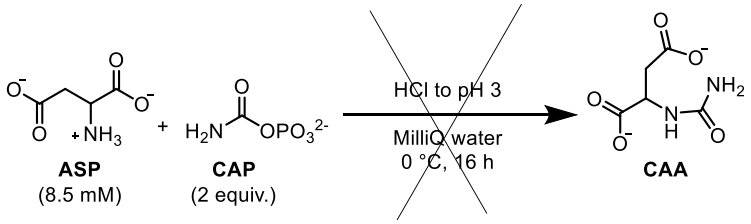


Table S1 Entry 18



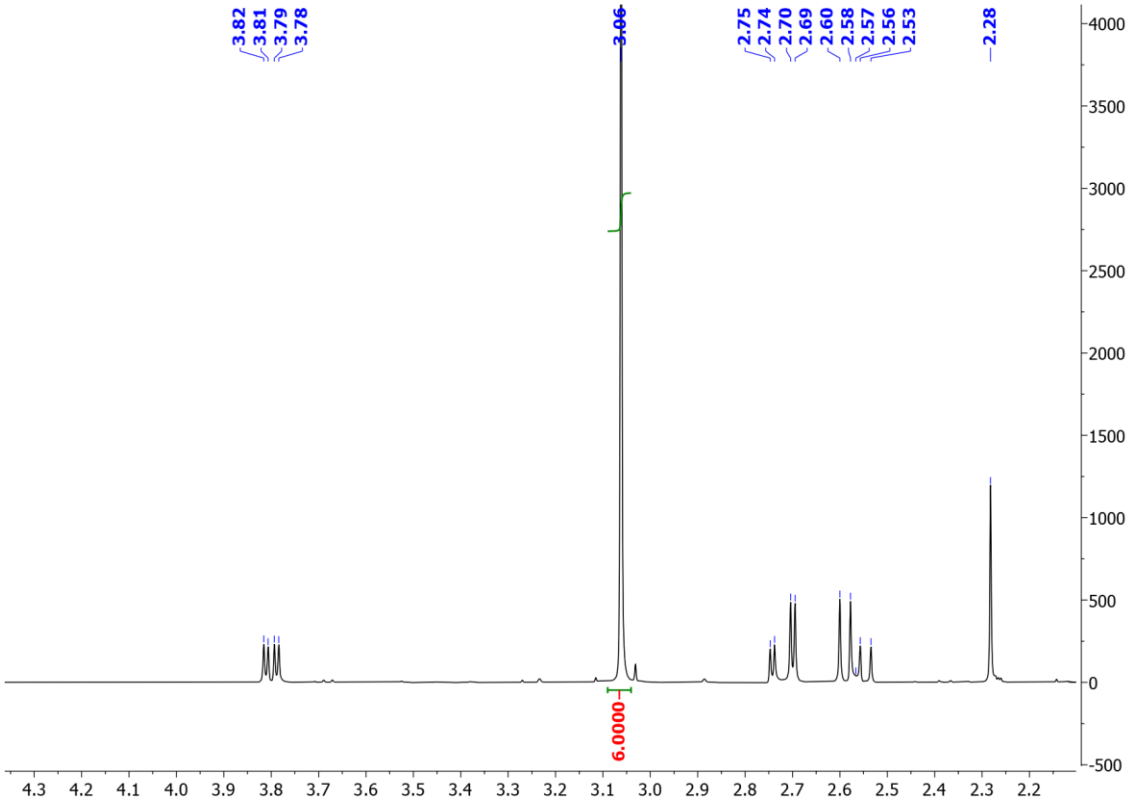
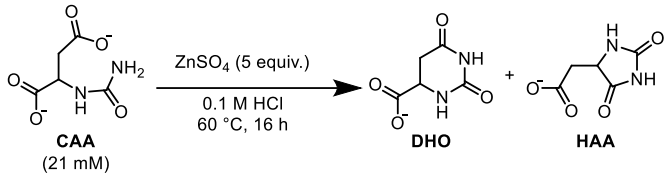


Table S2 Entry 1



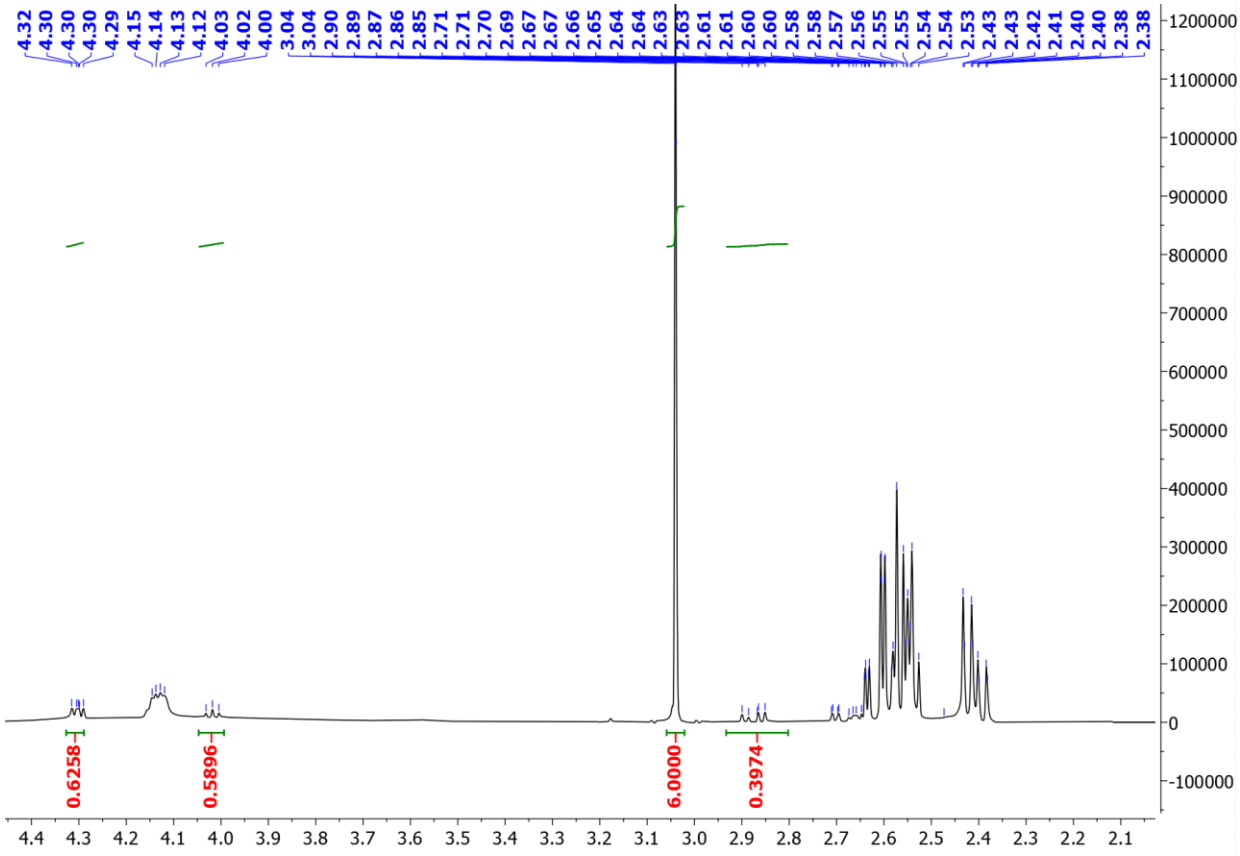
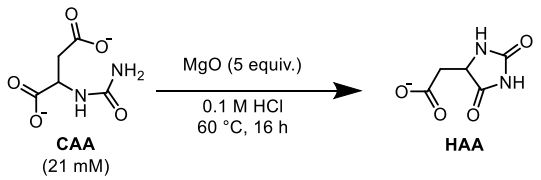


Table S2 Entry 2



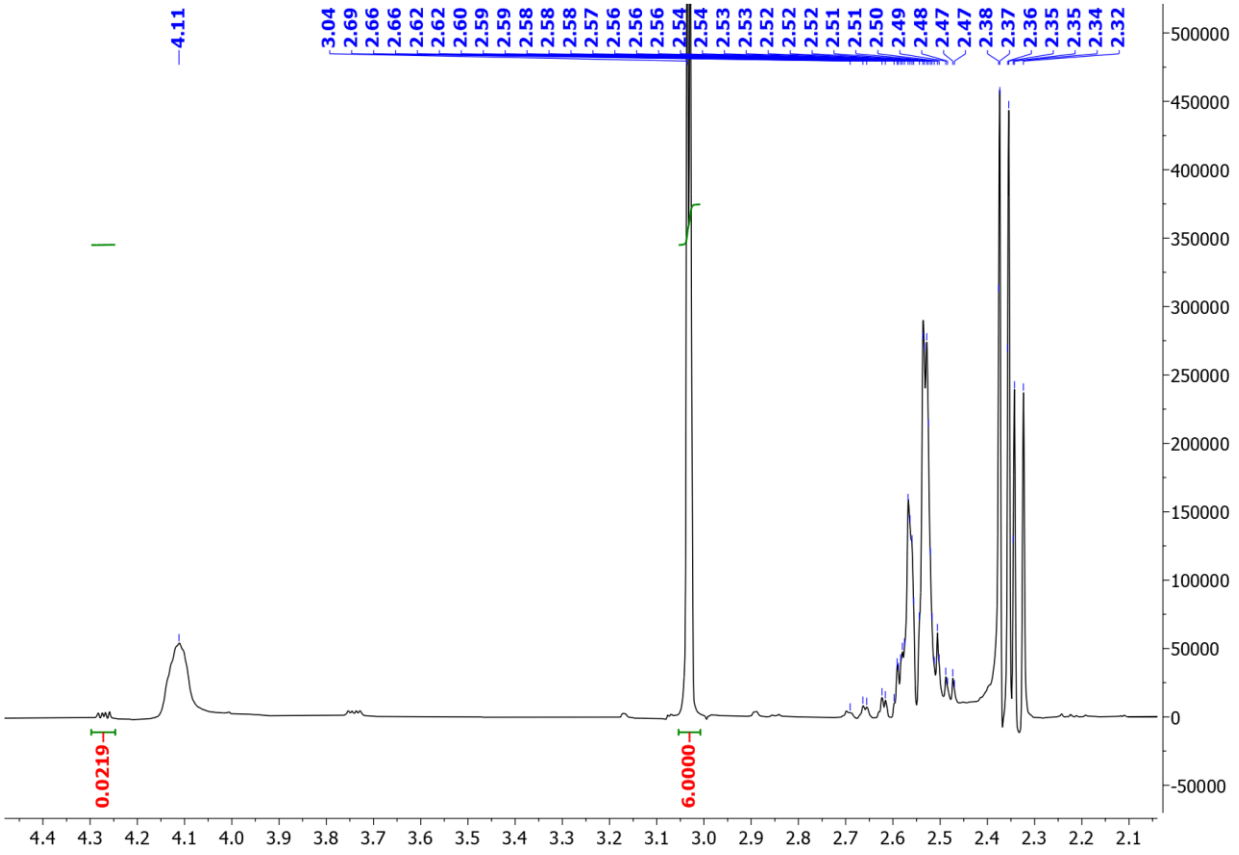
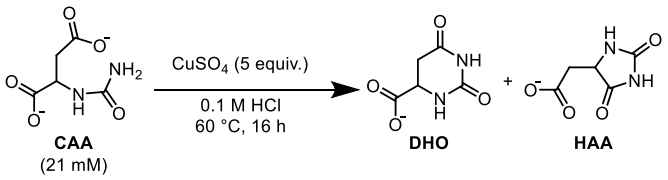


Table S2 Entry 3



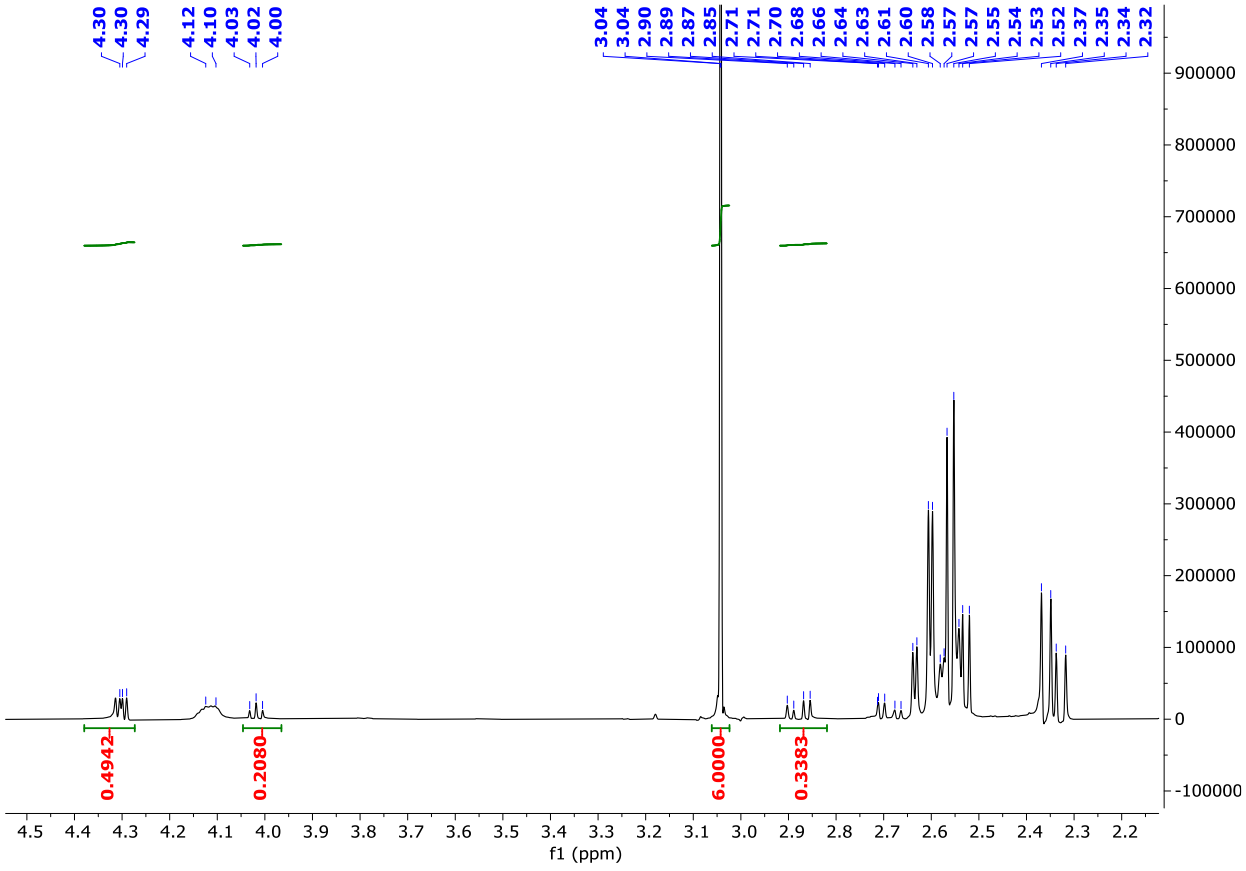
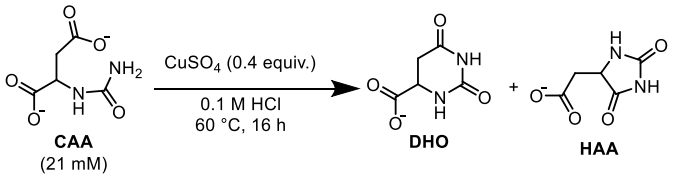


Table S2 Entry 4



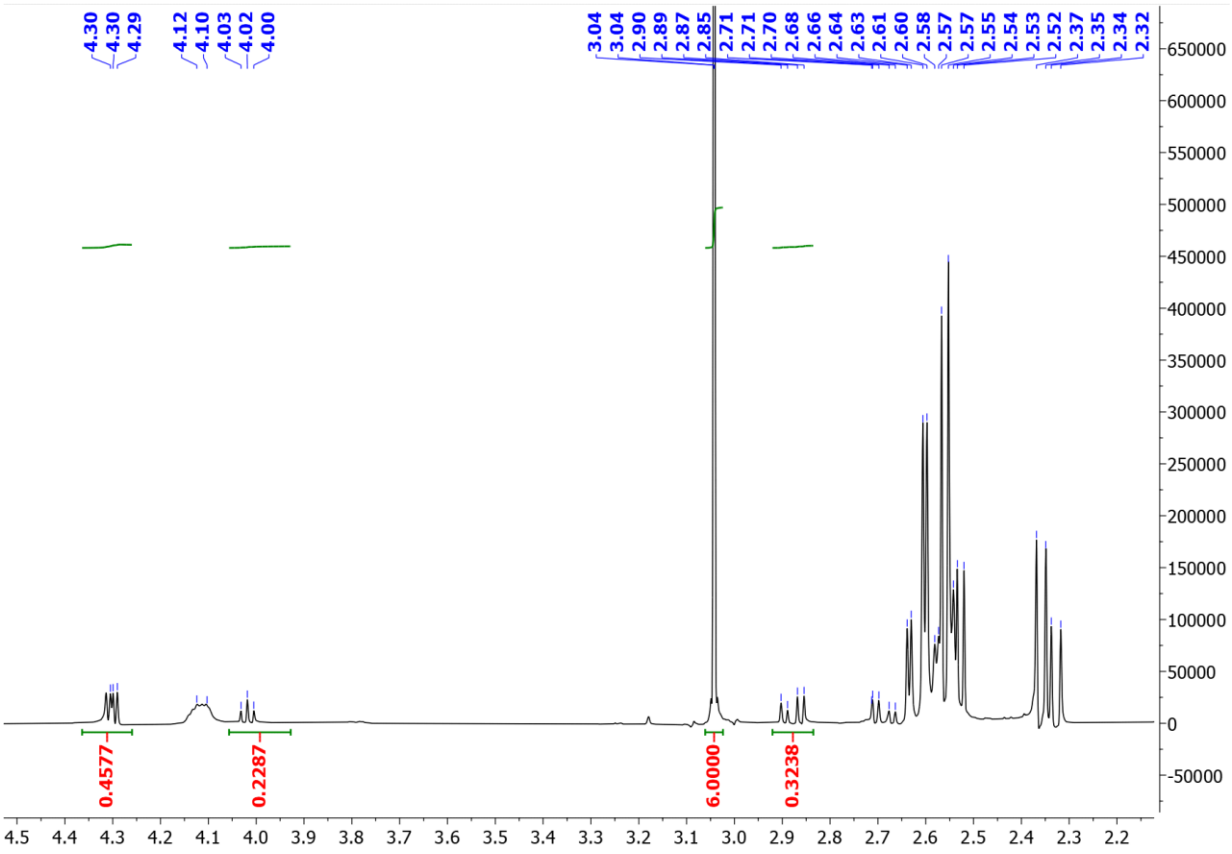
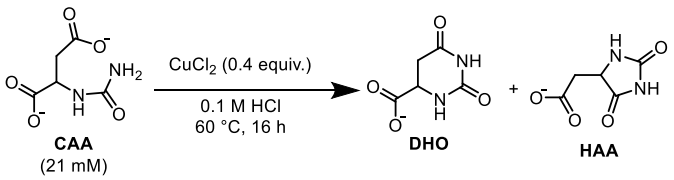


Table S2 Entry 5



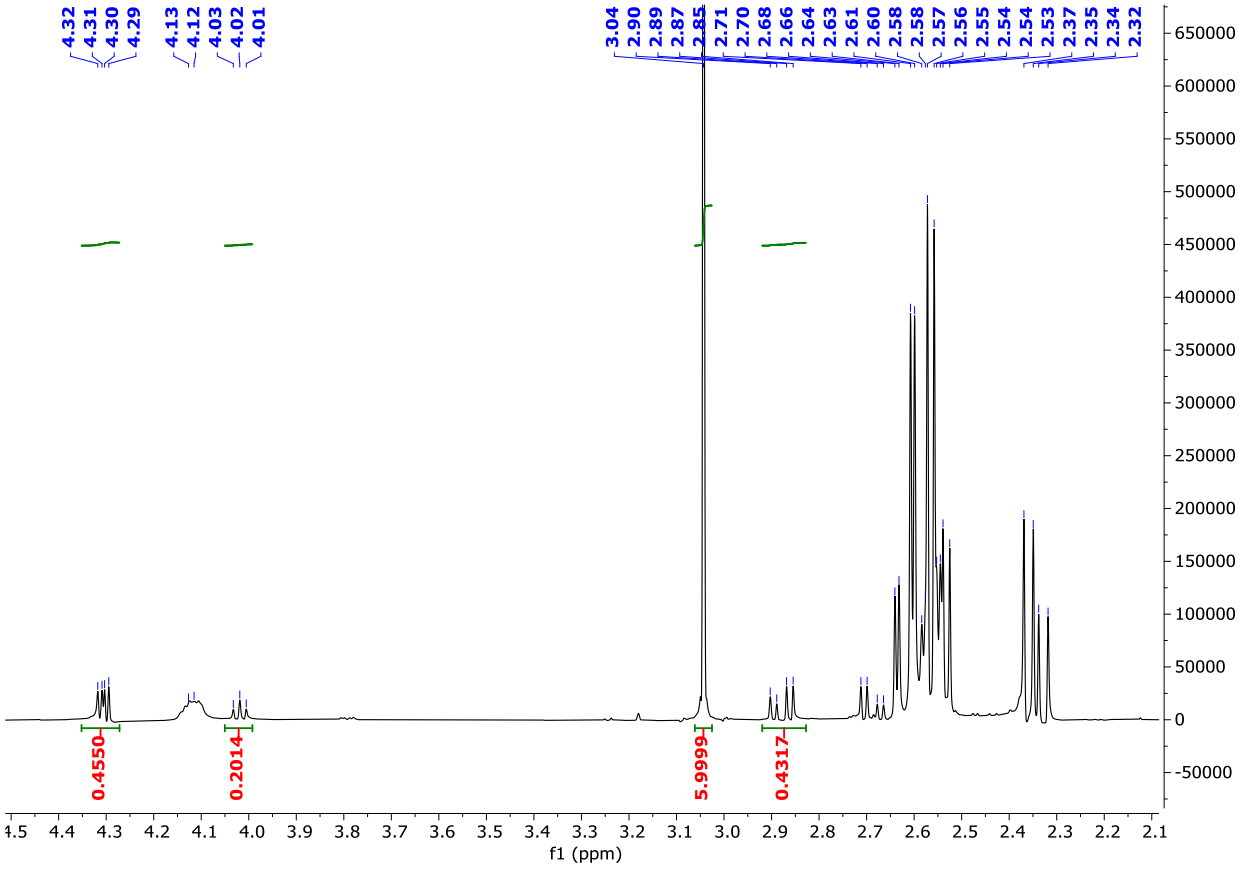
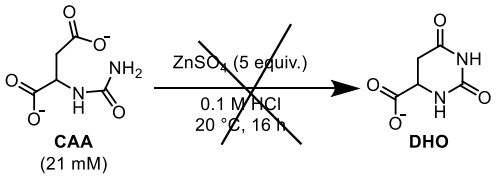


Table S2 Entry 6



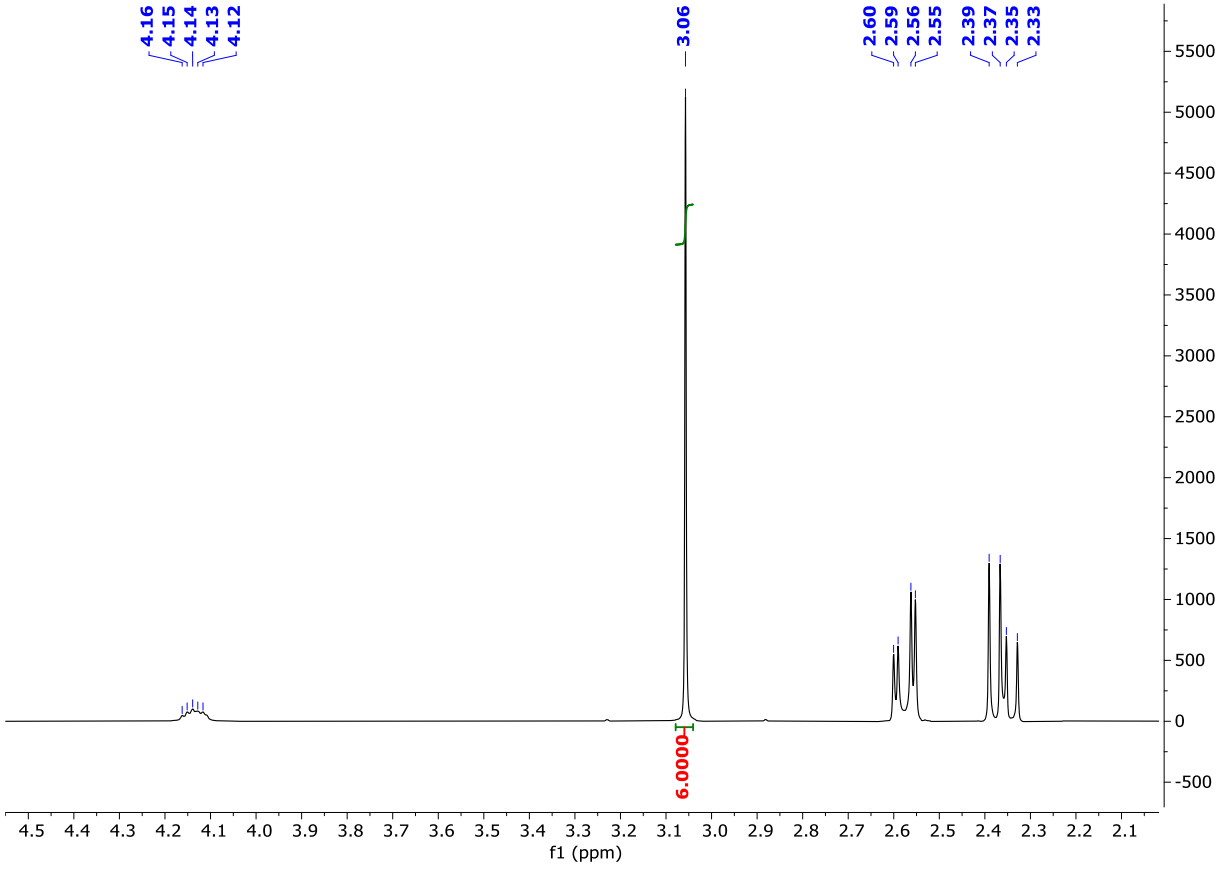
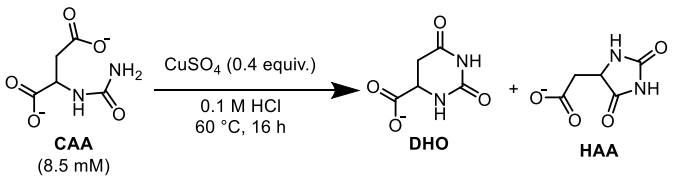


Table S2 Entry 7



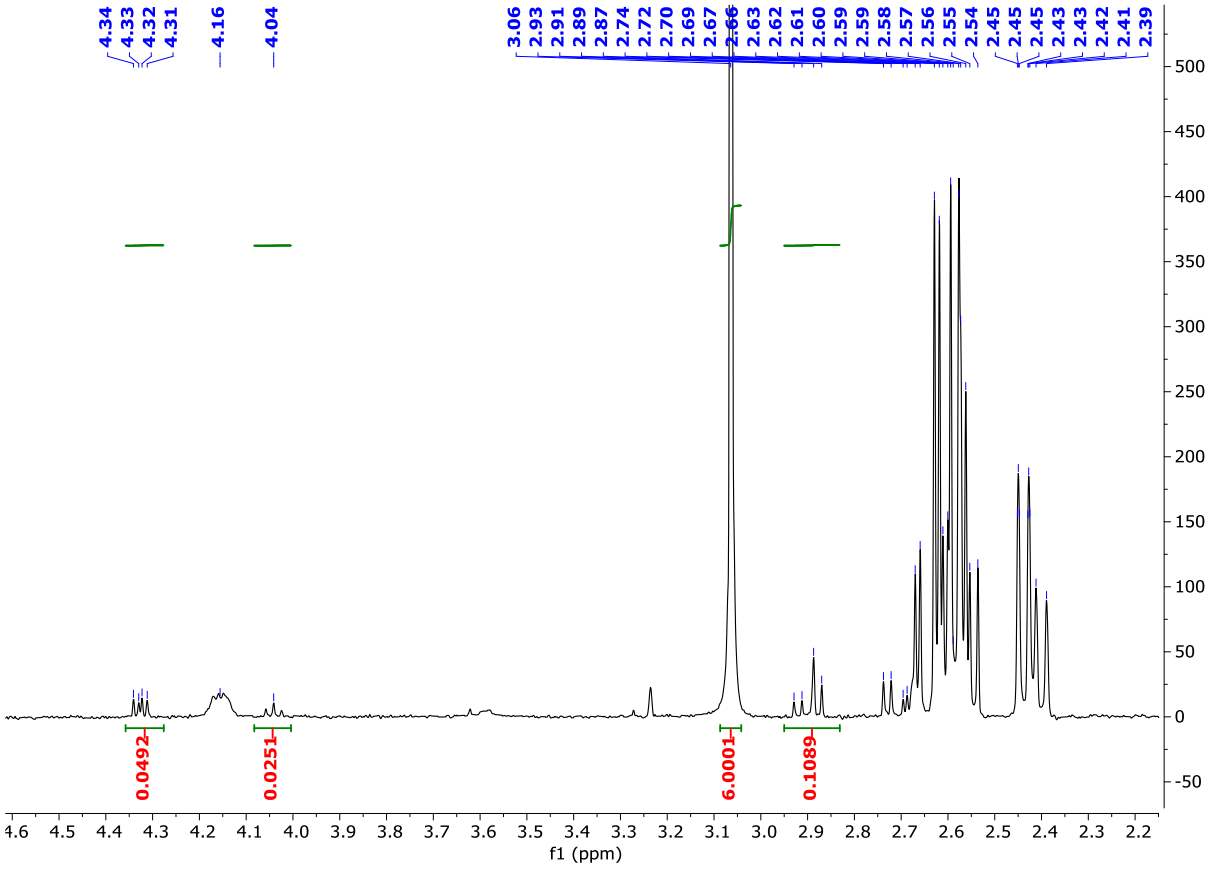
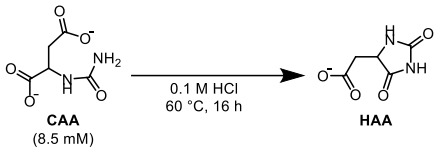


Table S2 Entry 8



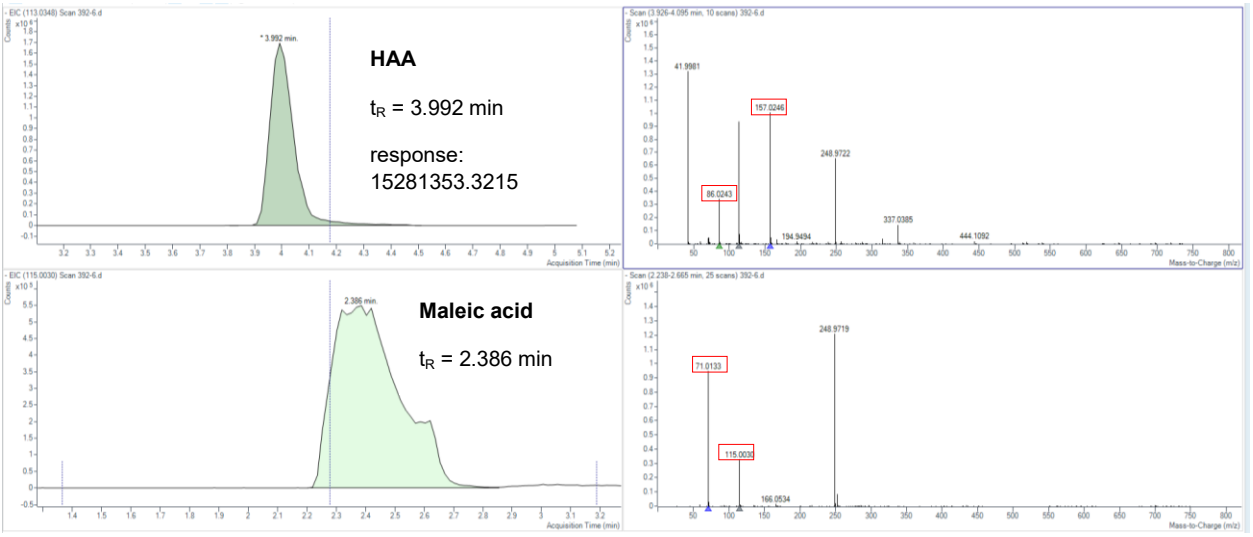
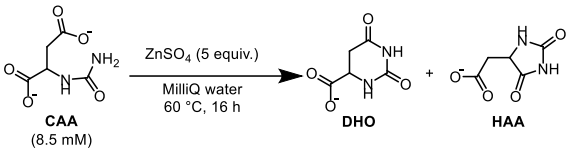
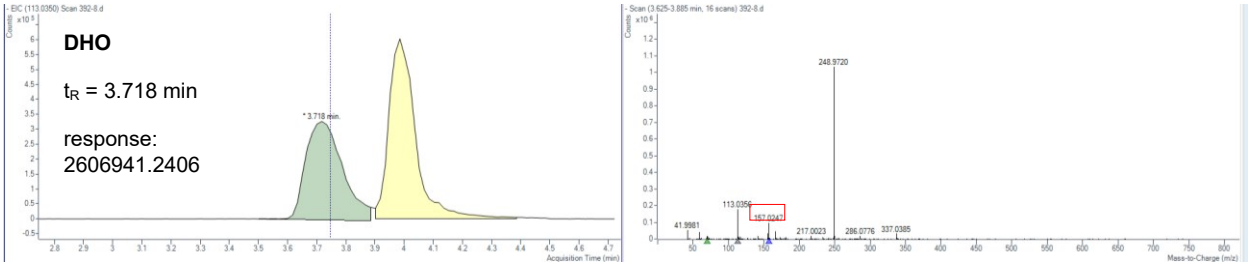
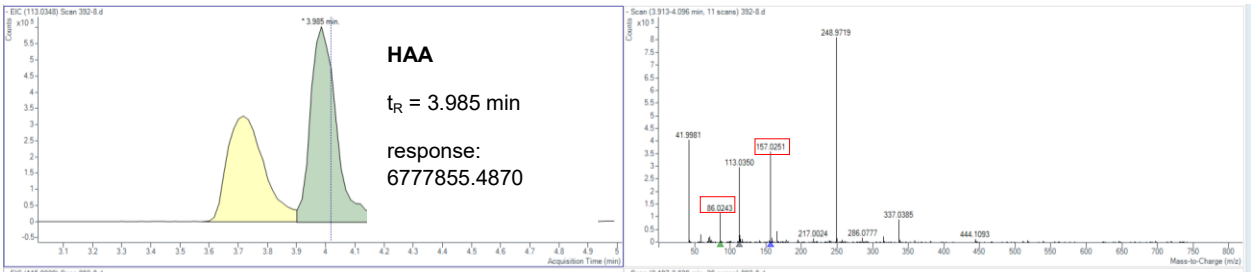


Table S2 Entry 9







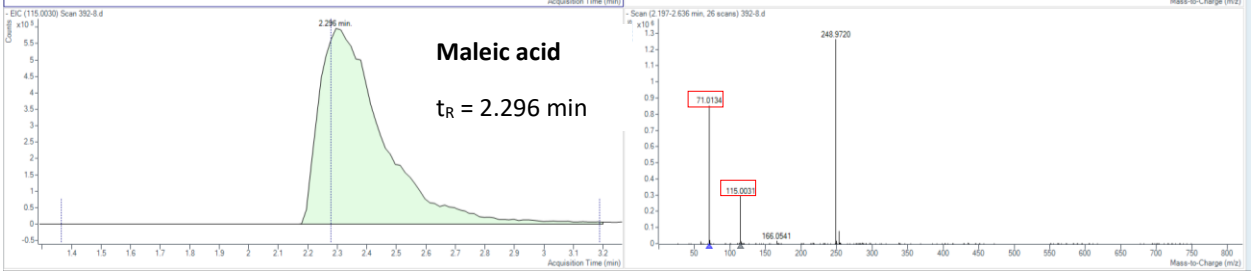
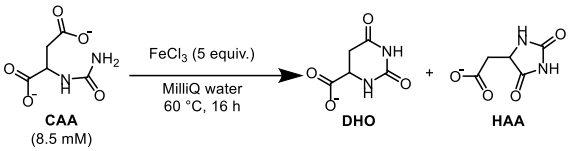
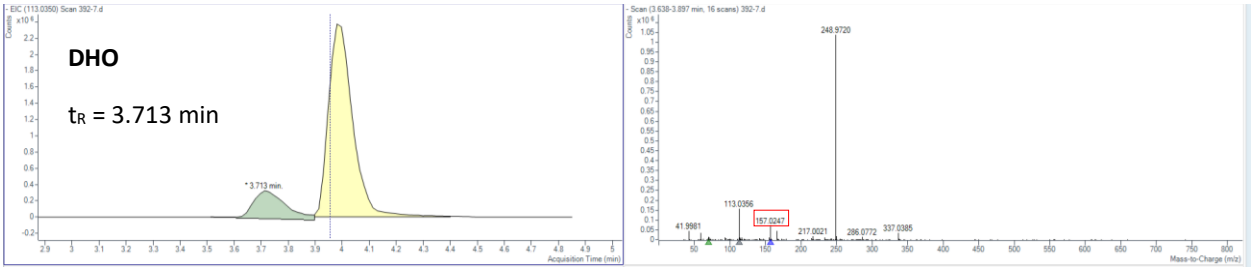
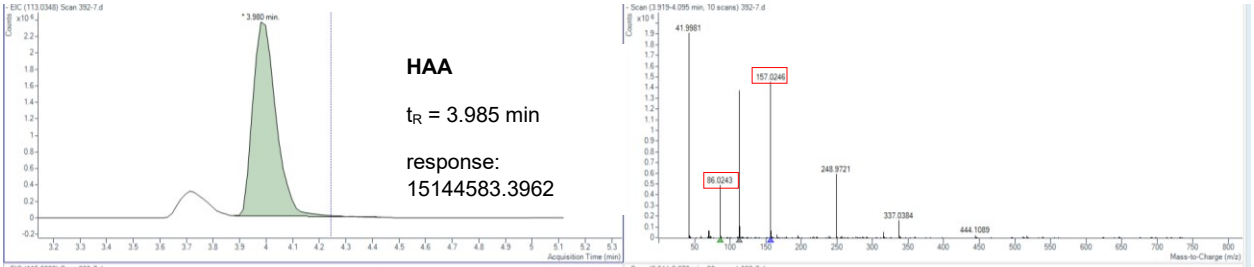


Table S2 Entry 10







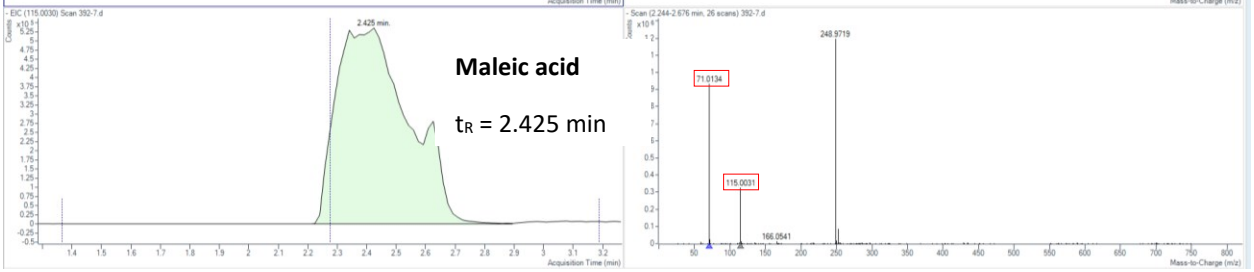
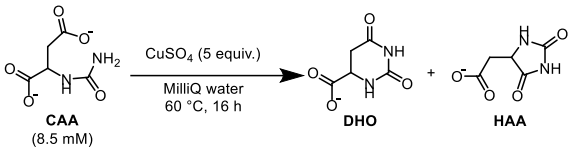


Table S2 Entry 11



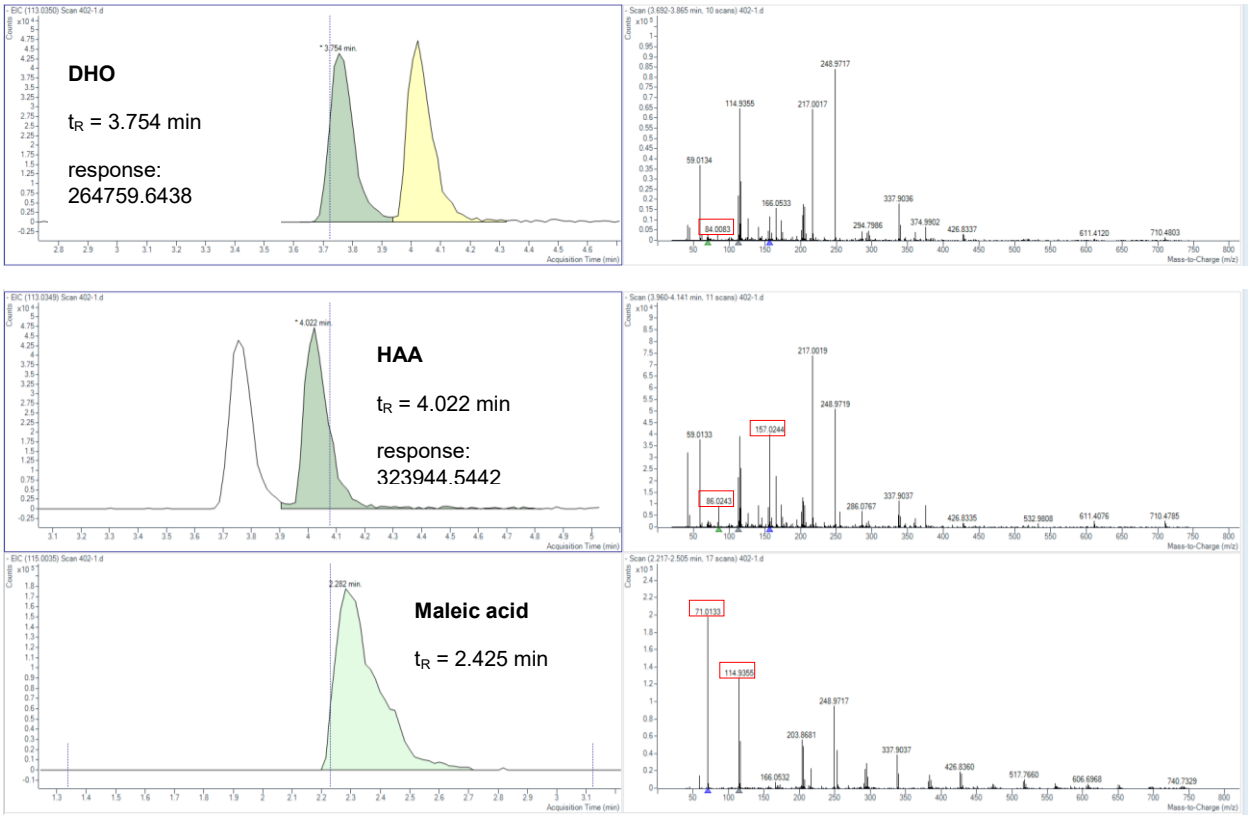
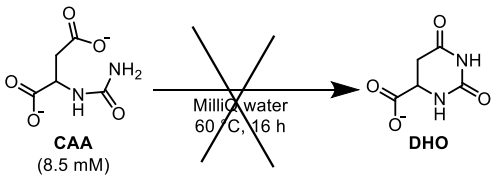


Table S2 Entry 12



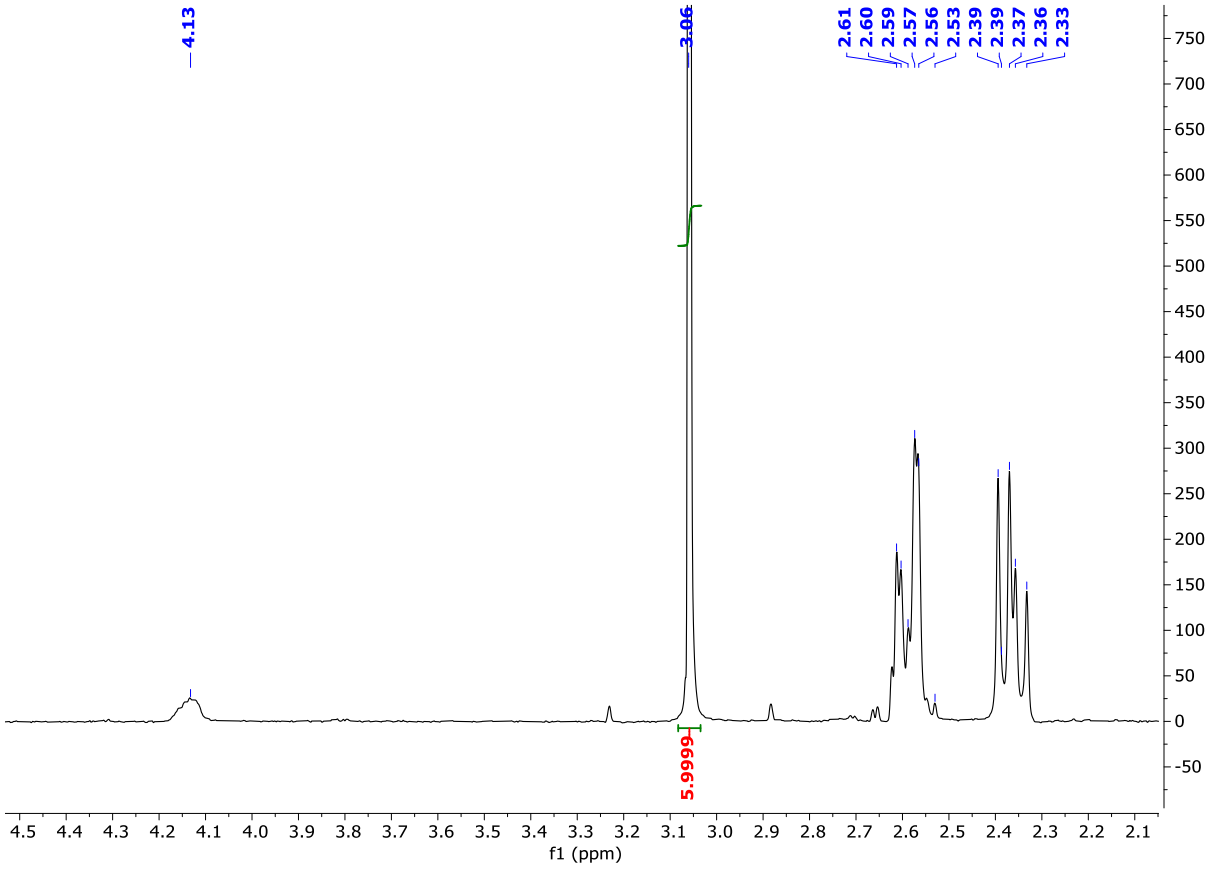
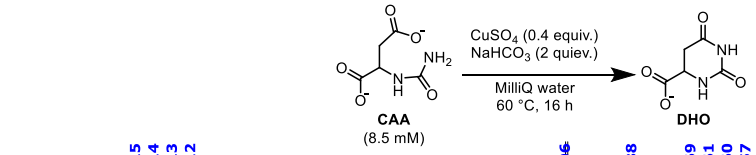
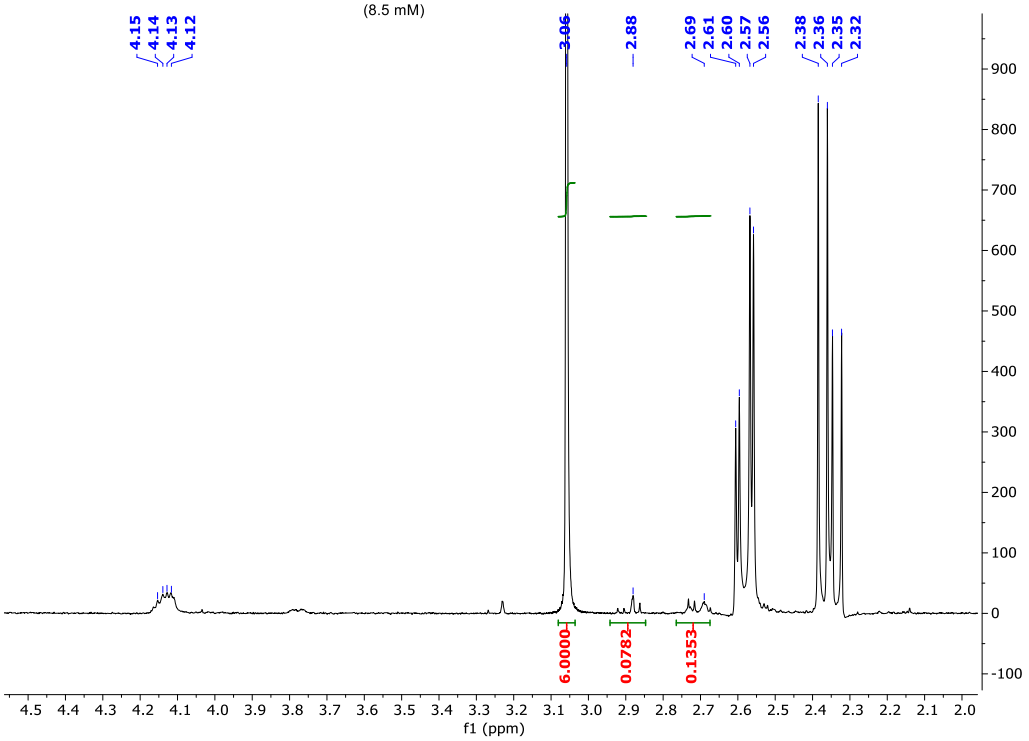


Table S2 Entry 13





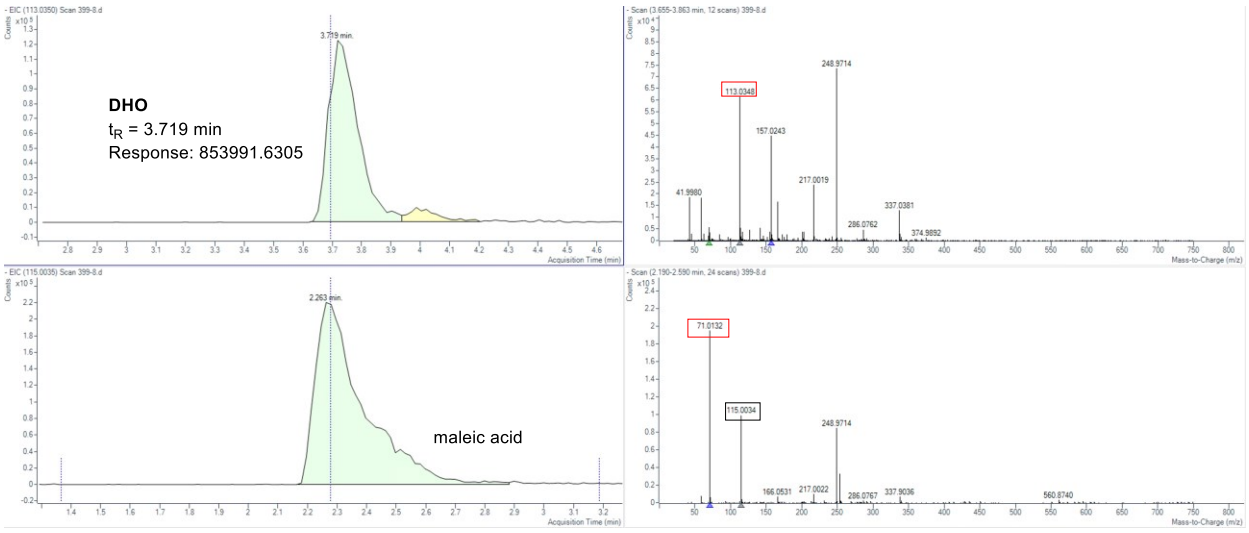
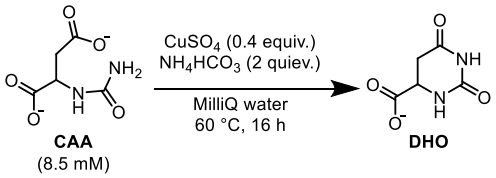


Table S2 Entry 14



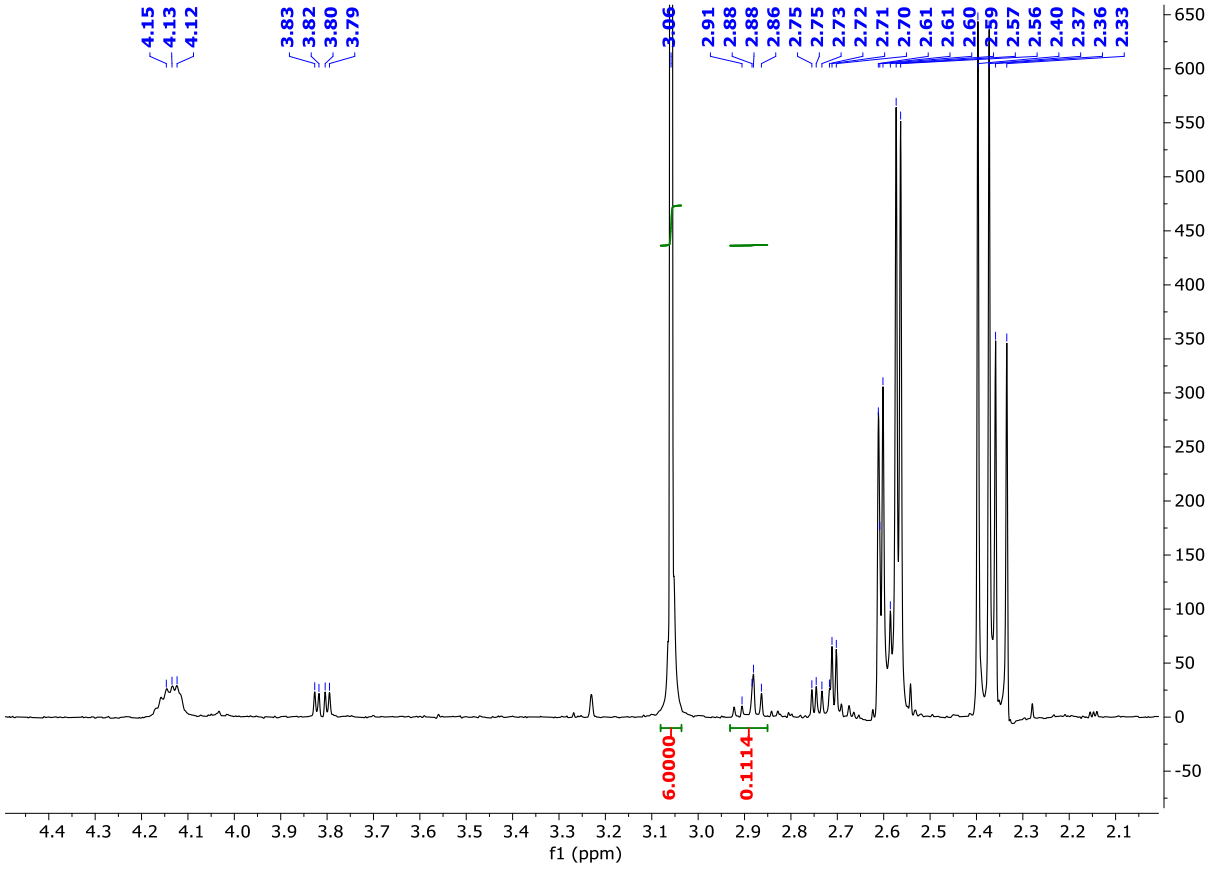
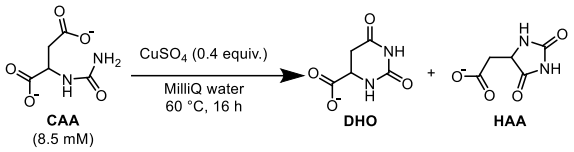


Table S2 Entry 15



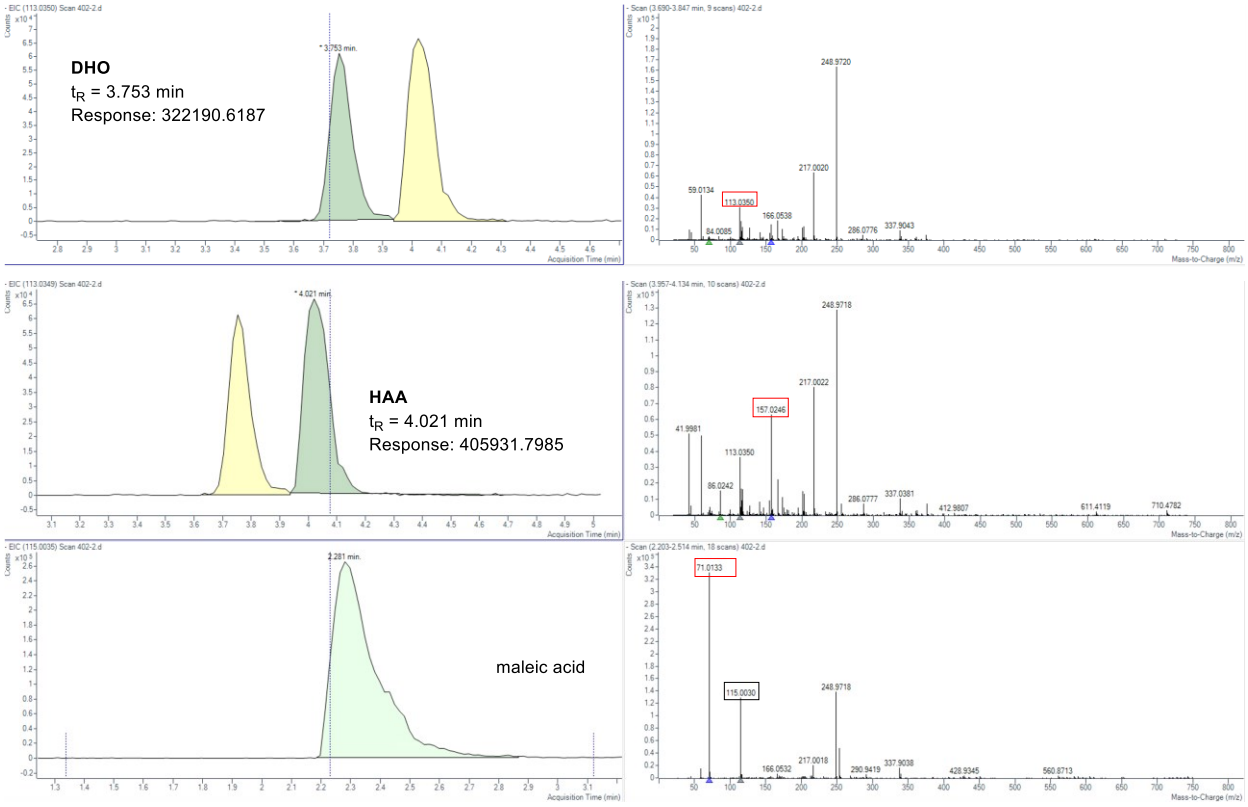
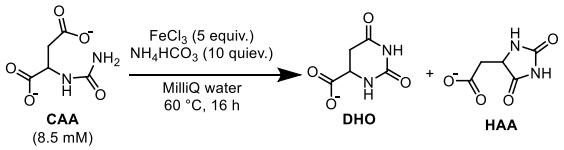
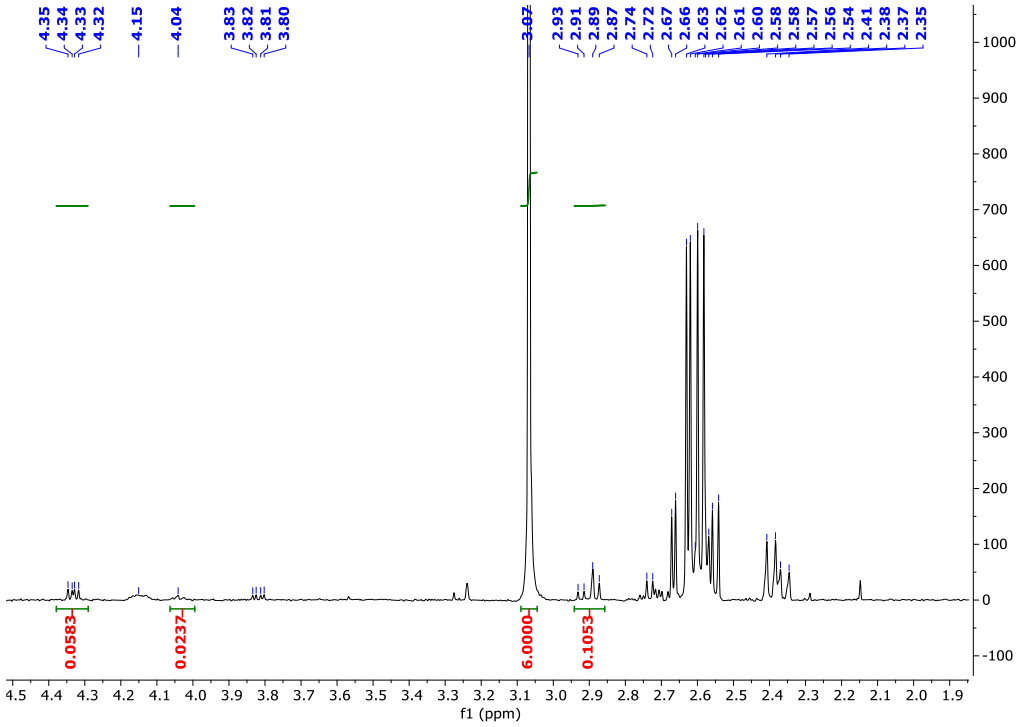
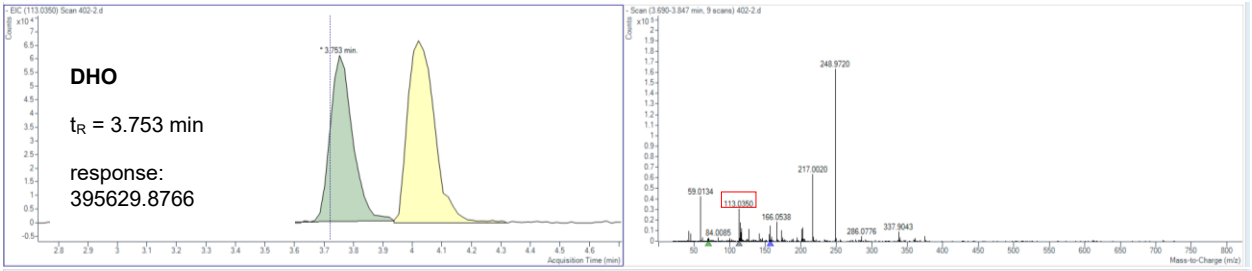
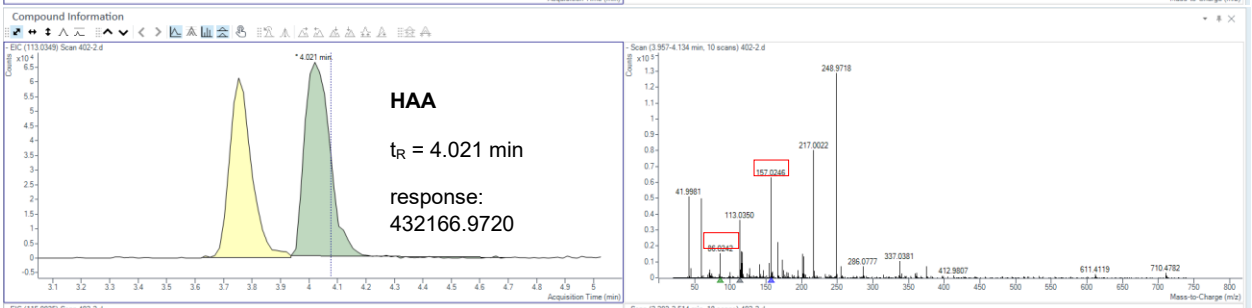


Table S2 Entry 16









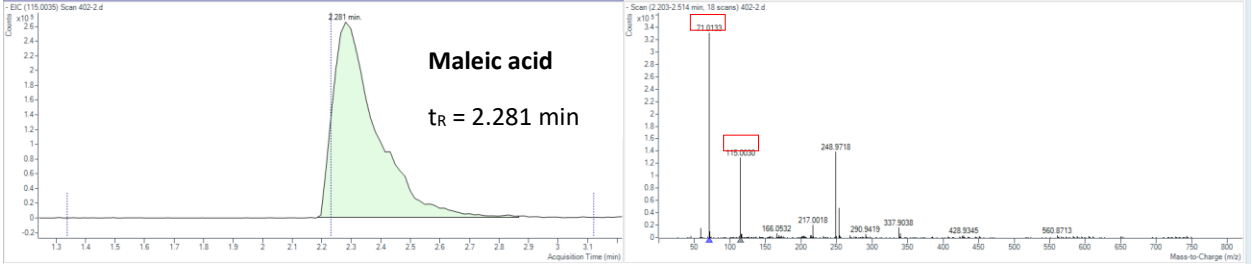
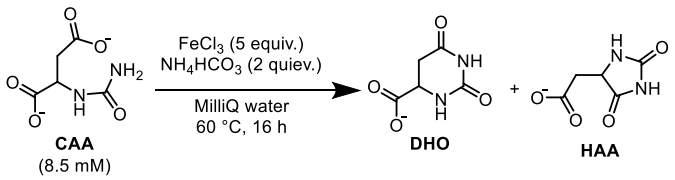


Table S2 Entry 17



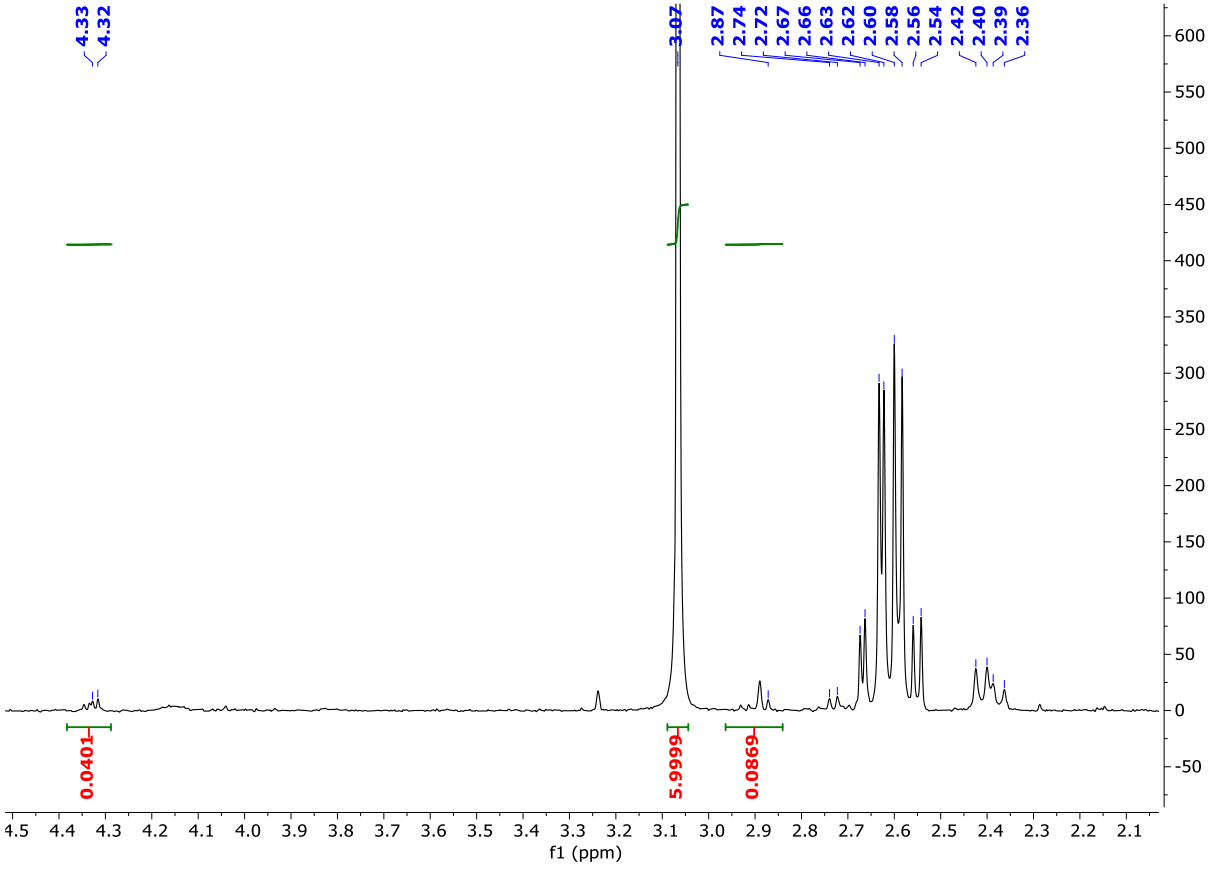
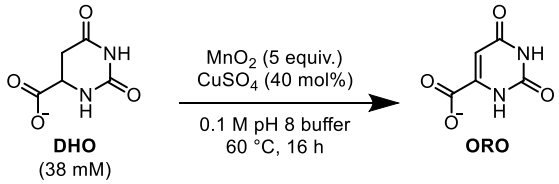


Table S3 Entry 1



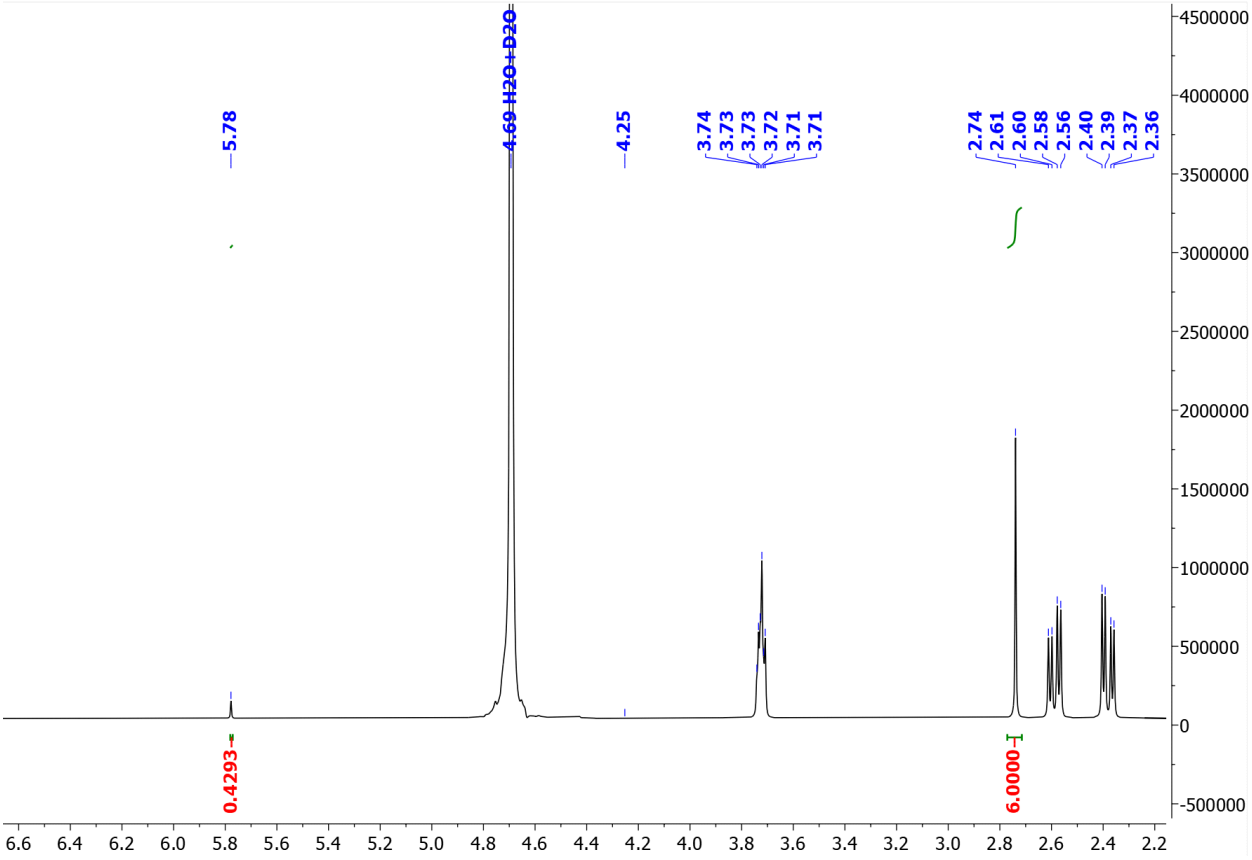
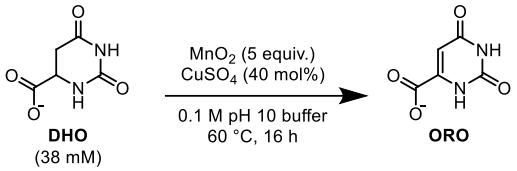


Table S3 Entry 2



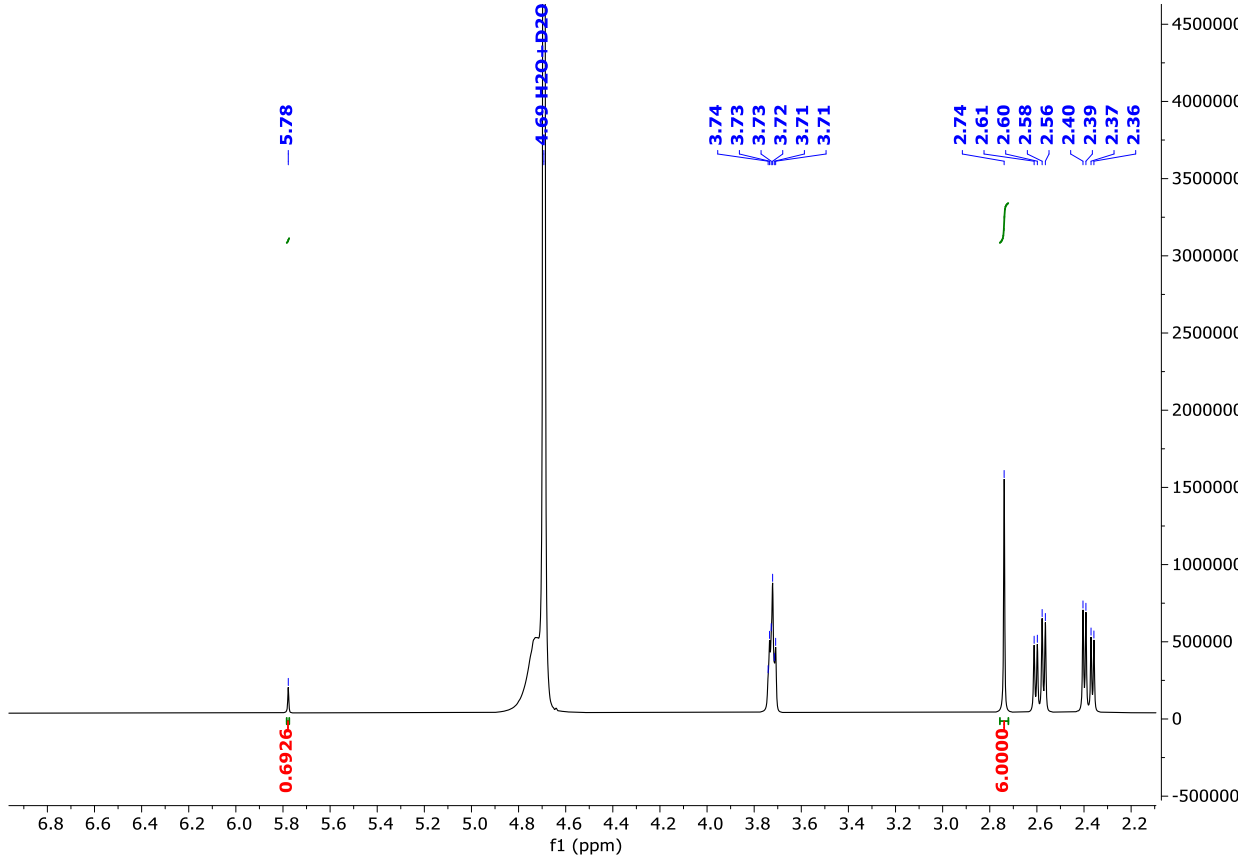
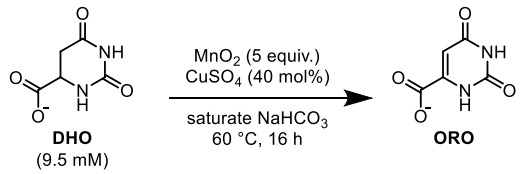


Table S3 Entry 3



DHO
(9.5 mM)

ORO

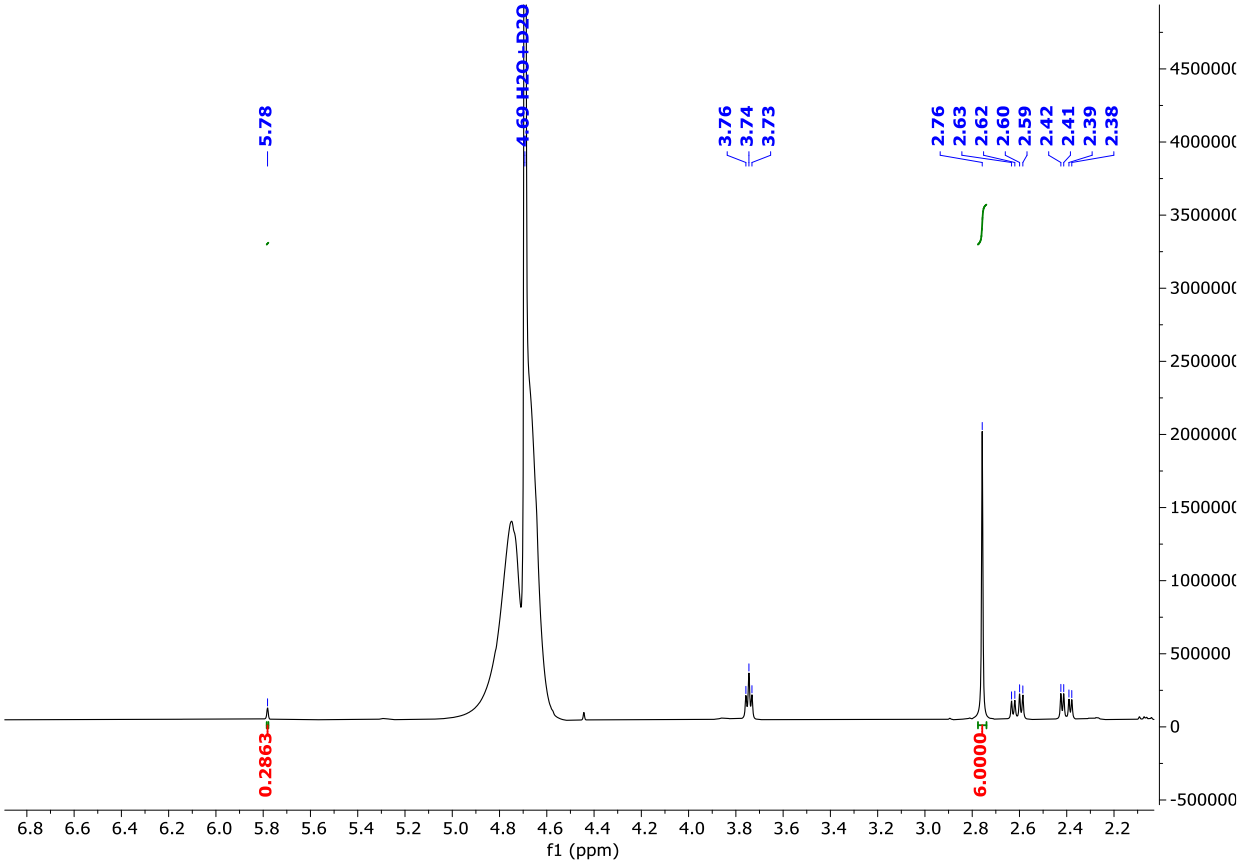
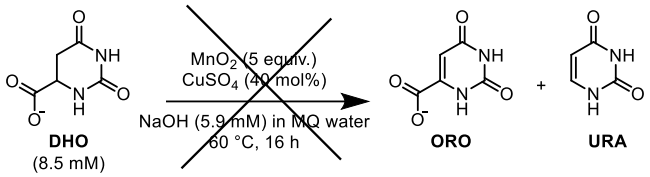


Table S3 Entry 4



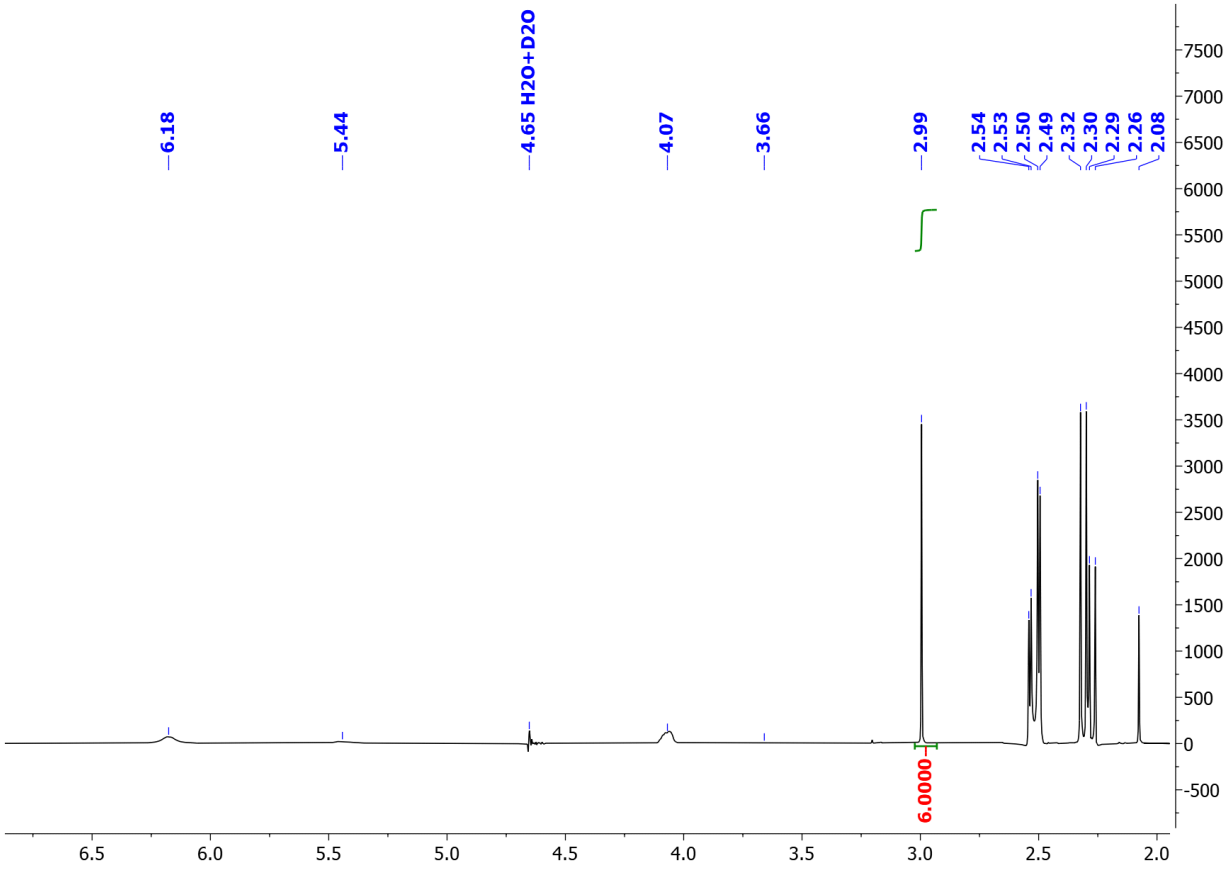
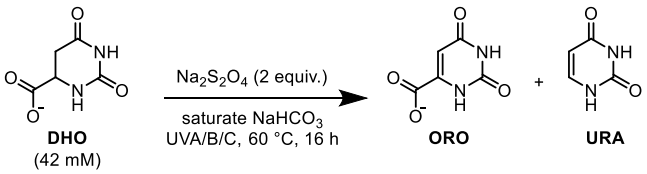


Table S3 Entry 5



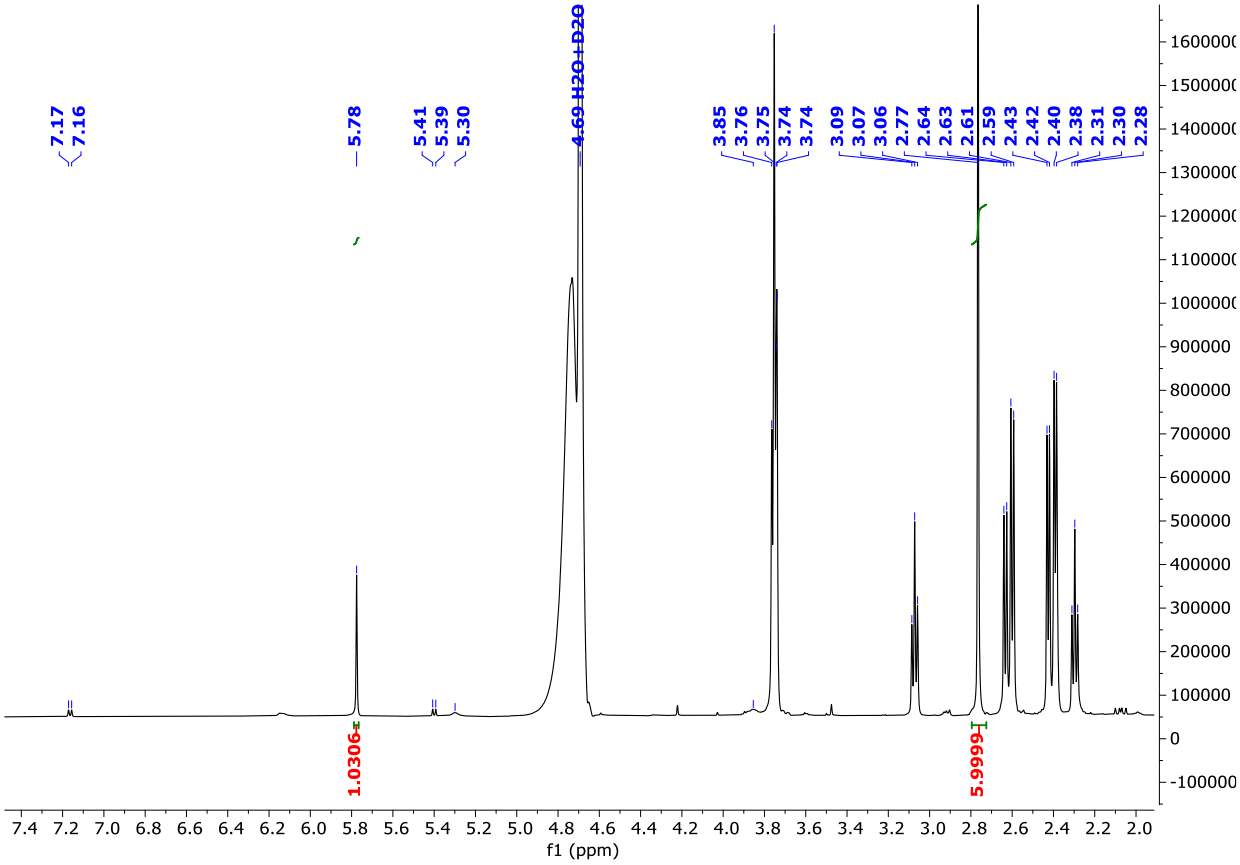
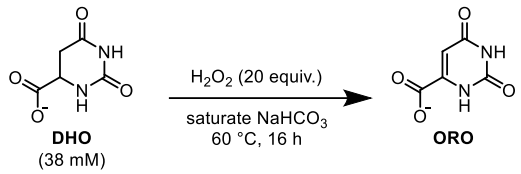


Table S3 Entry 6



DHO
(38 mM)

ORO

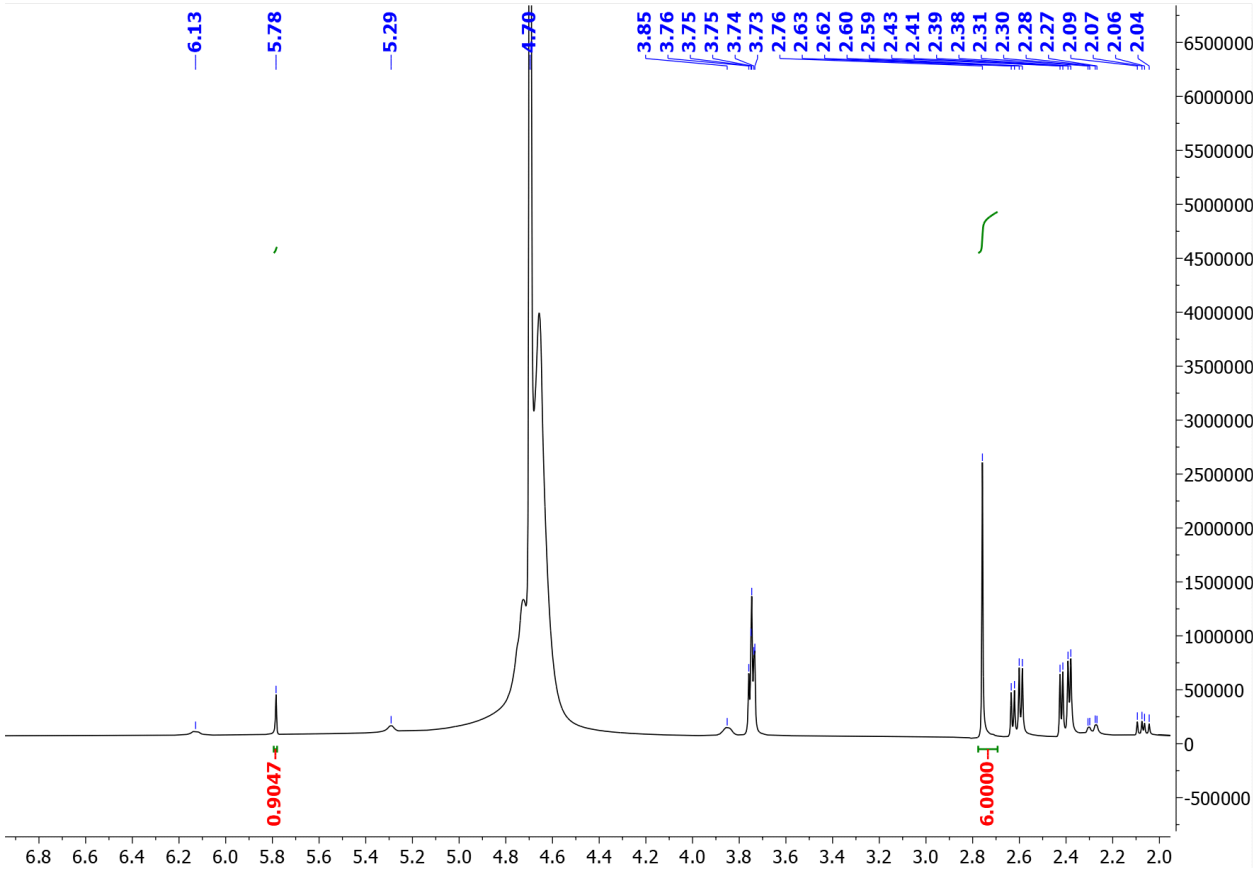


Table S3 Entry 7

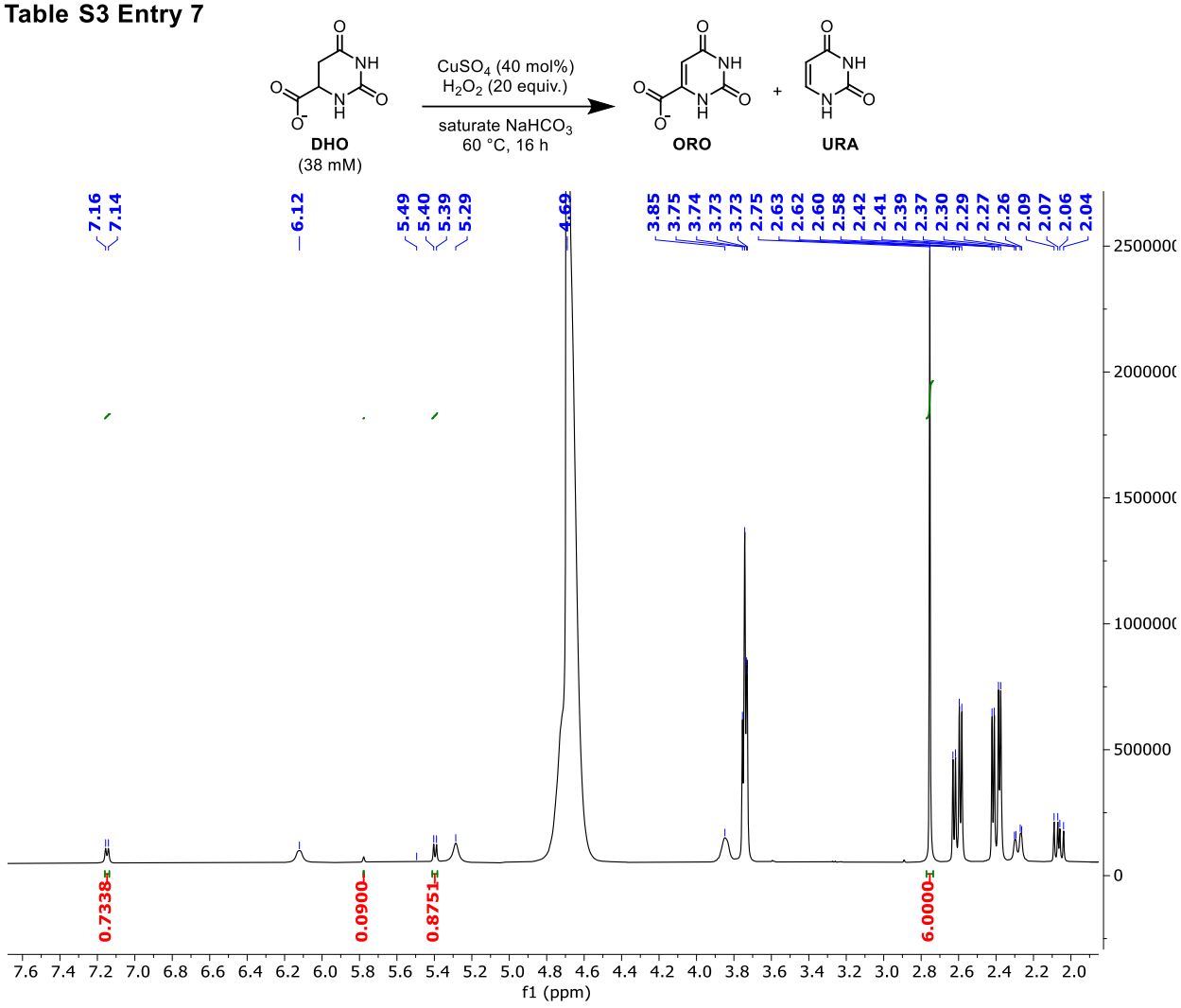
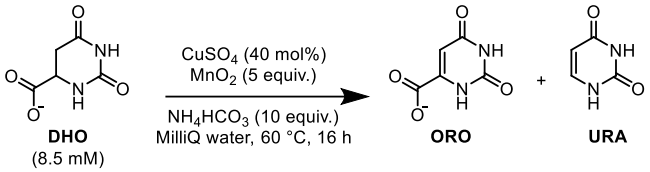


Table S3 Entry 8



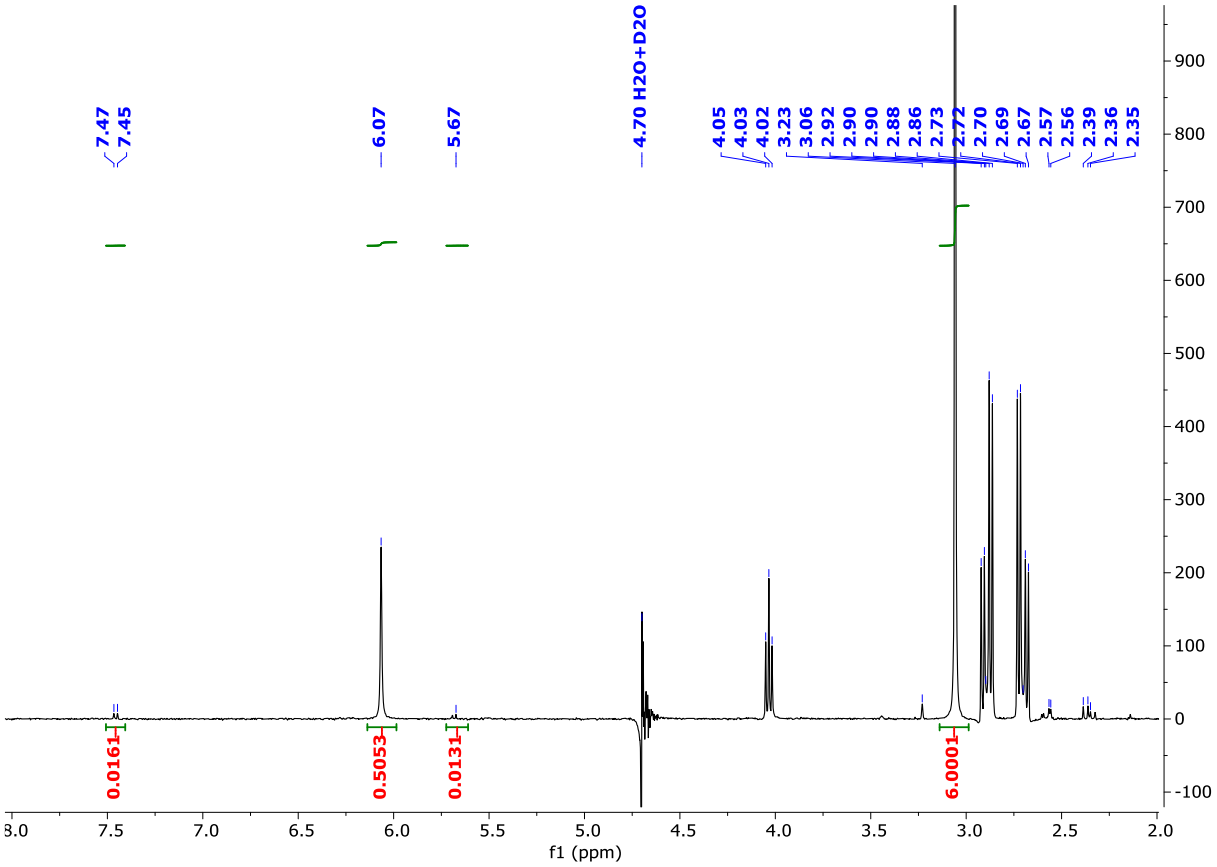
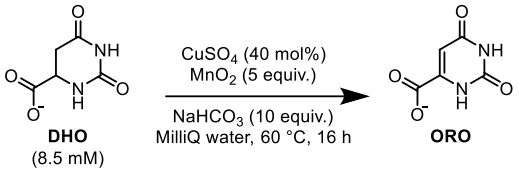


Table S3 Entry 9



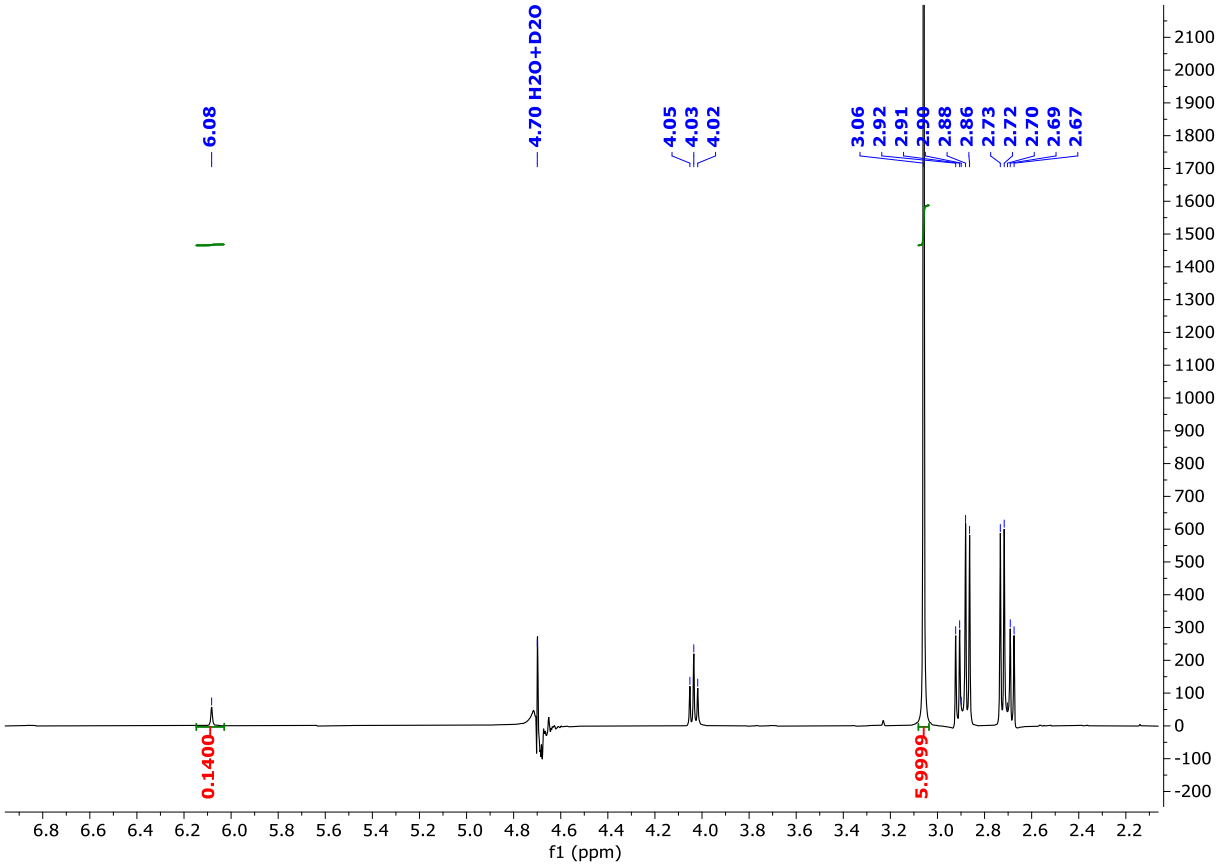
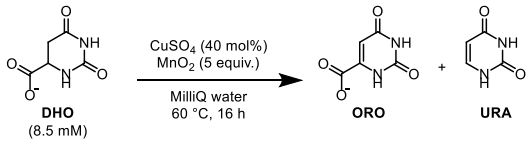
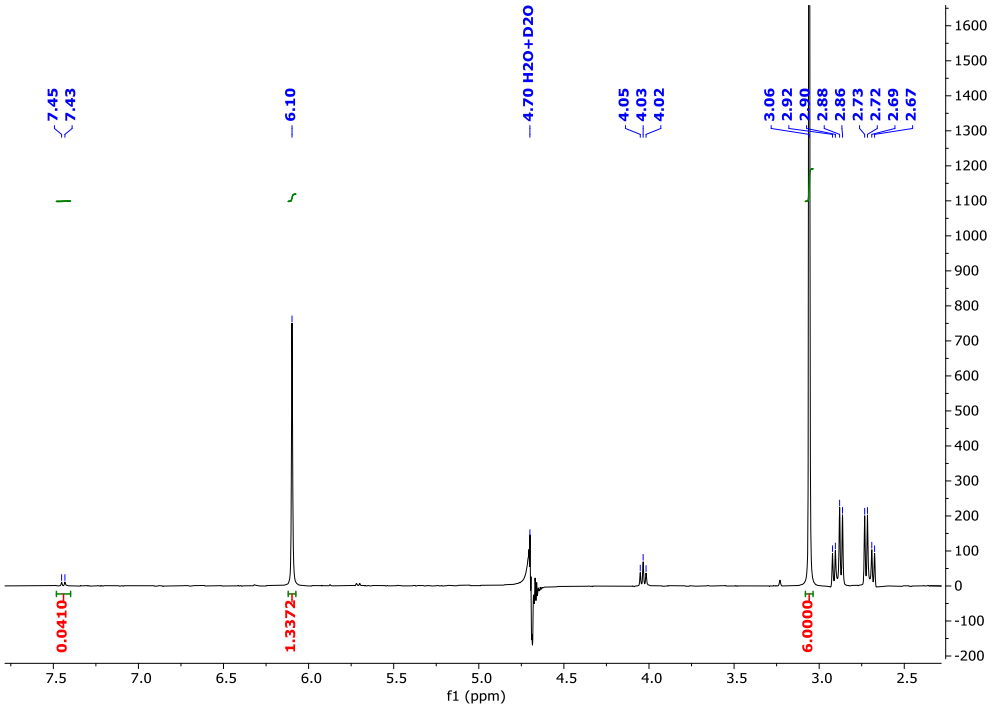
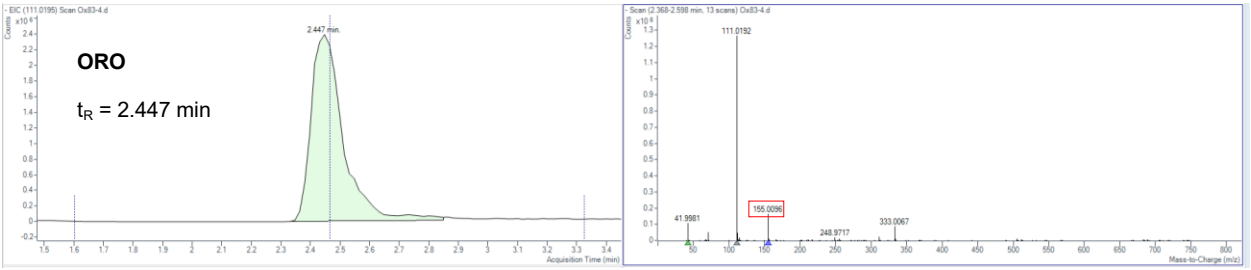
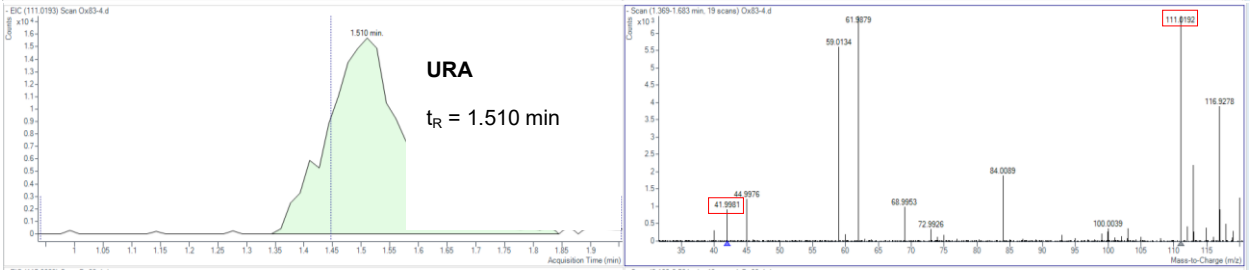


Table S3 Entry 10









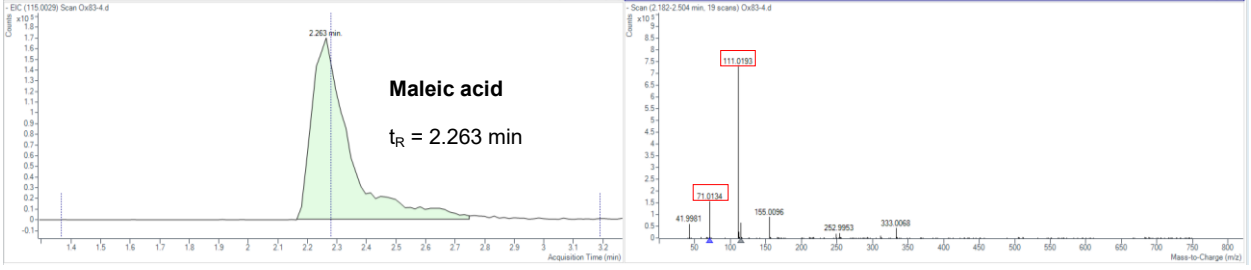
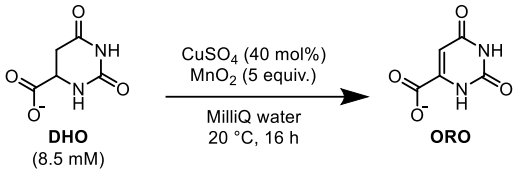


Table S3 Entry 11



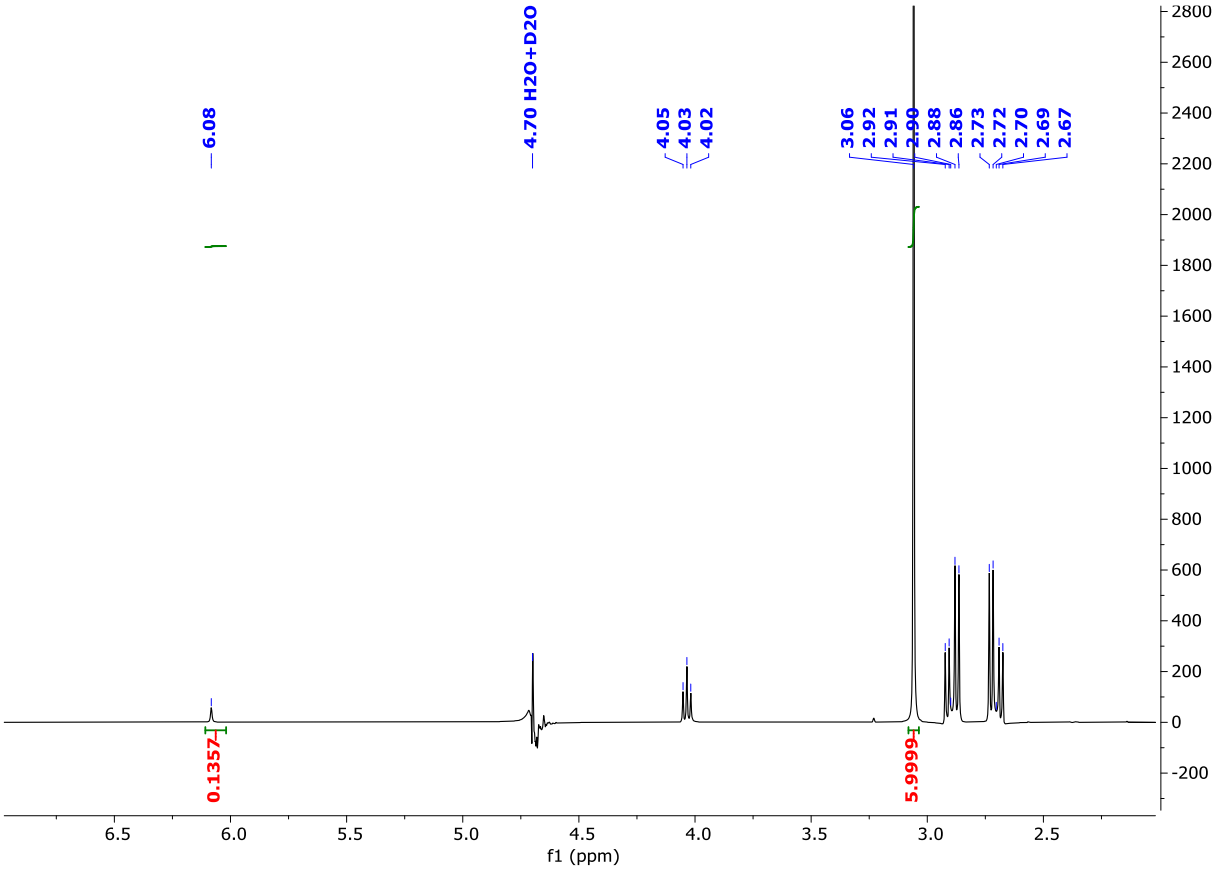
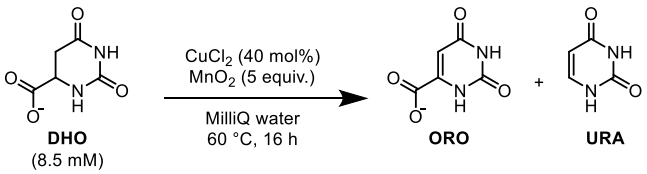


Table S3 Entry 12



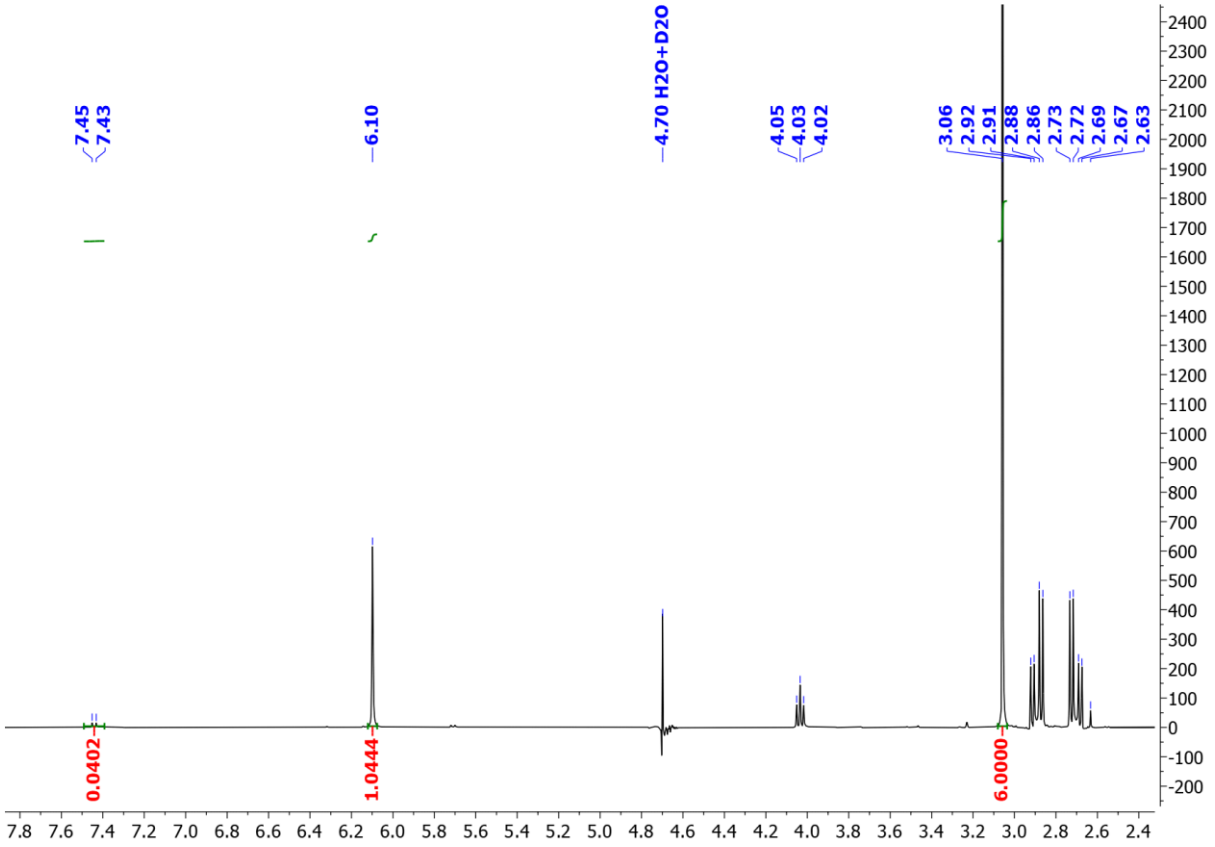
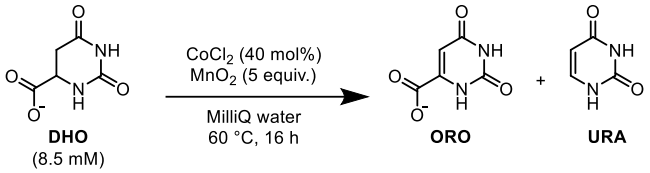


Table S3 Entry 13



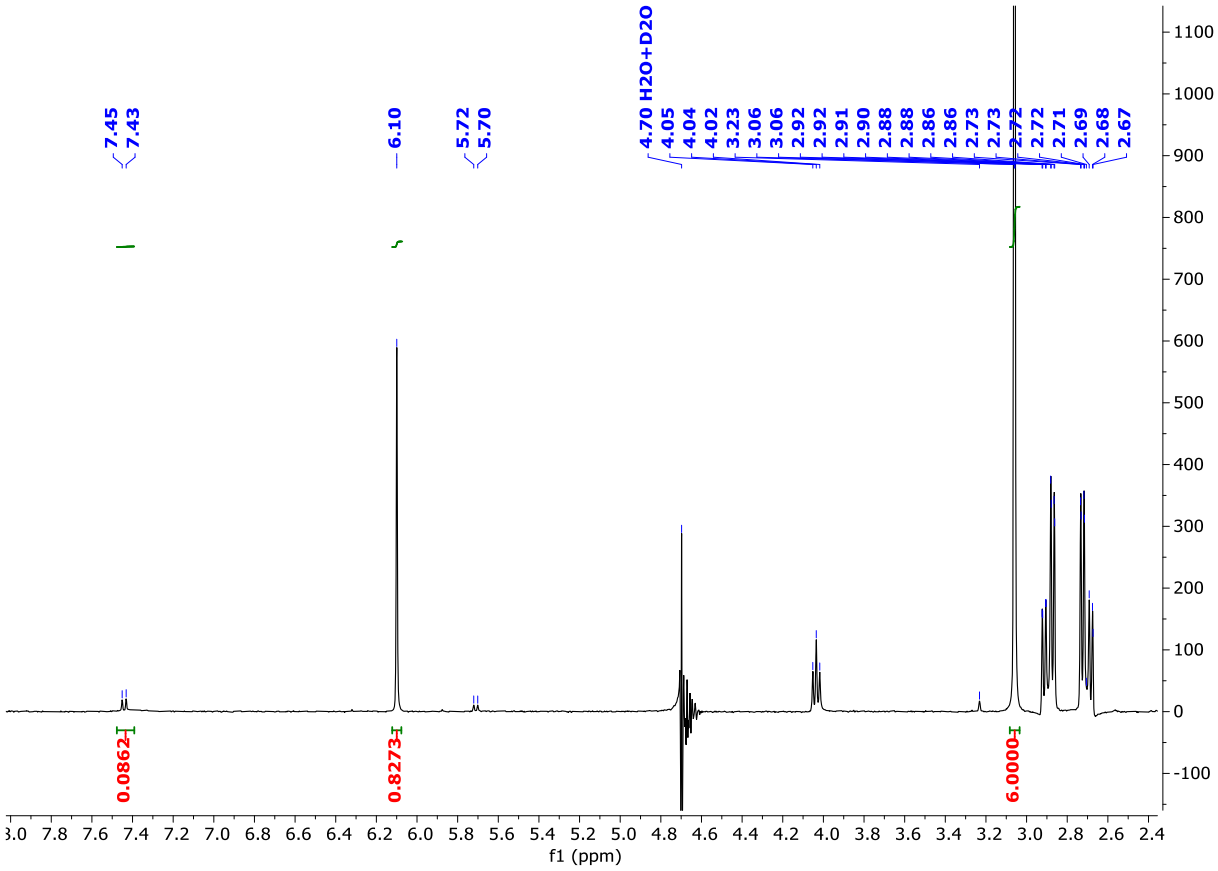
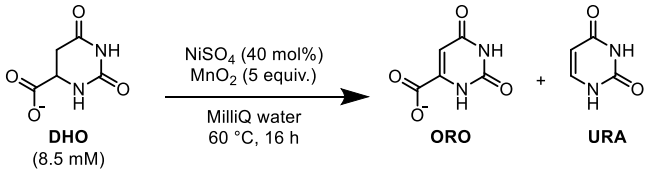


Table S3 Entry 14



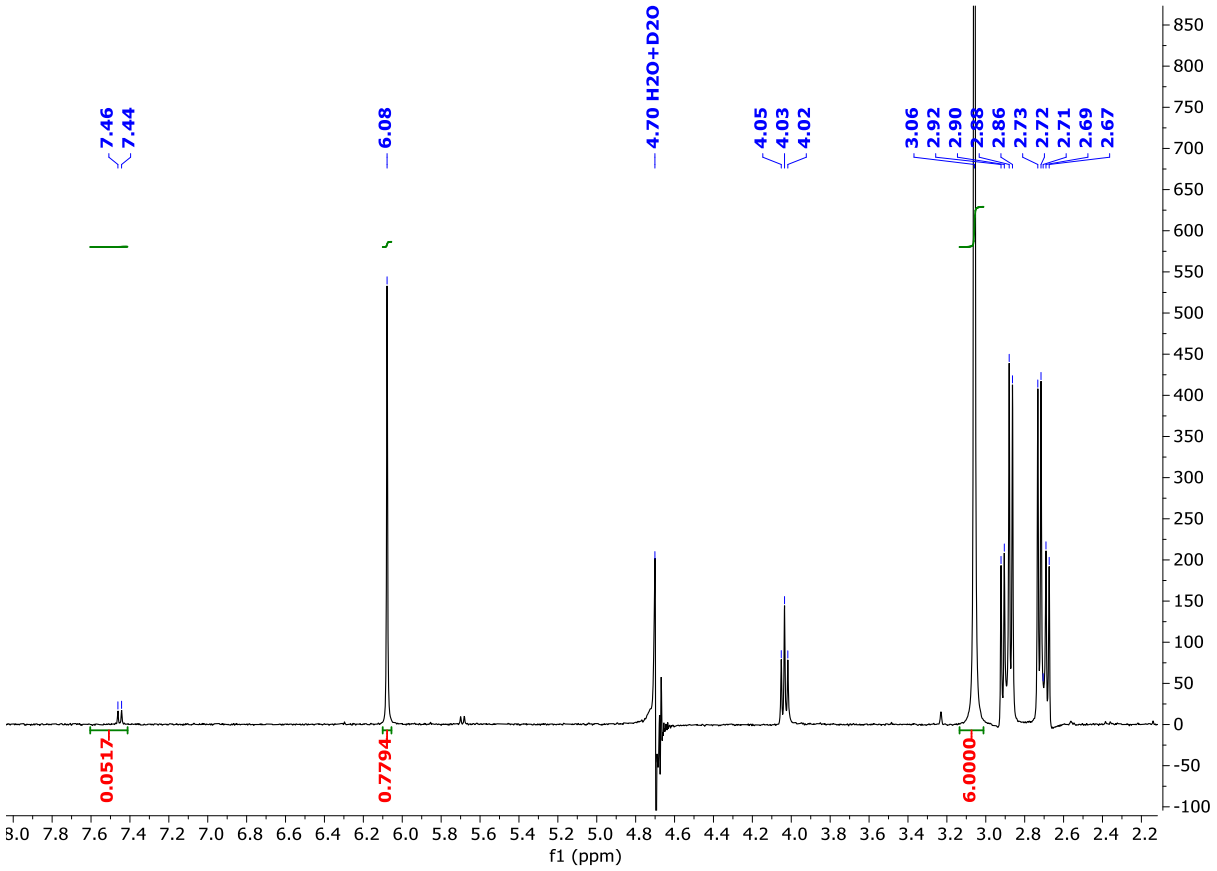
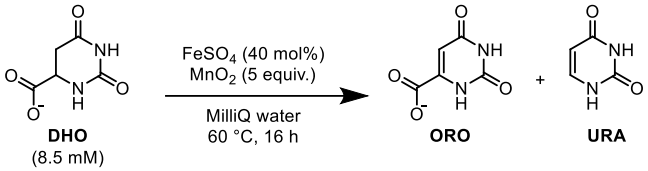


Table S3 Entry 15



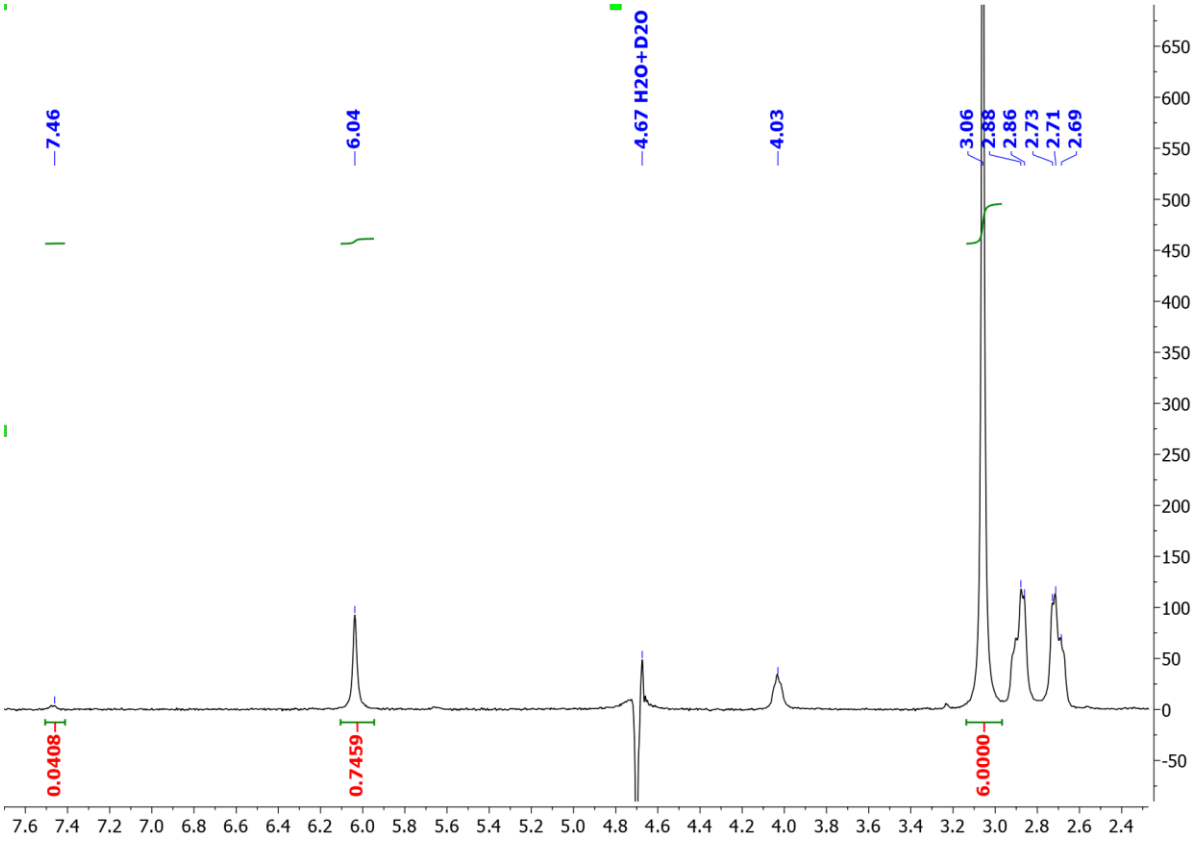


Table S3 Entry 16

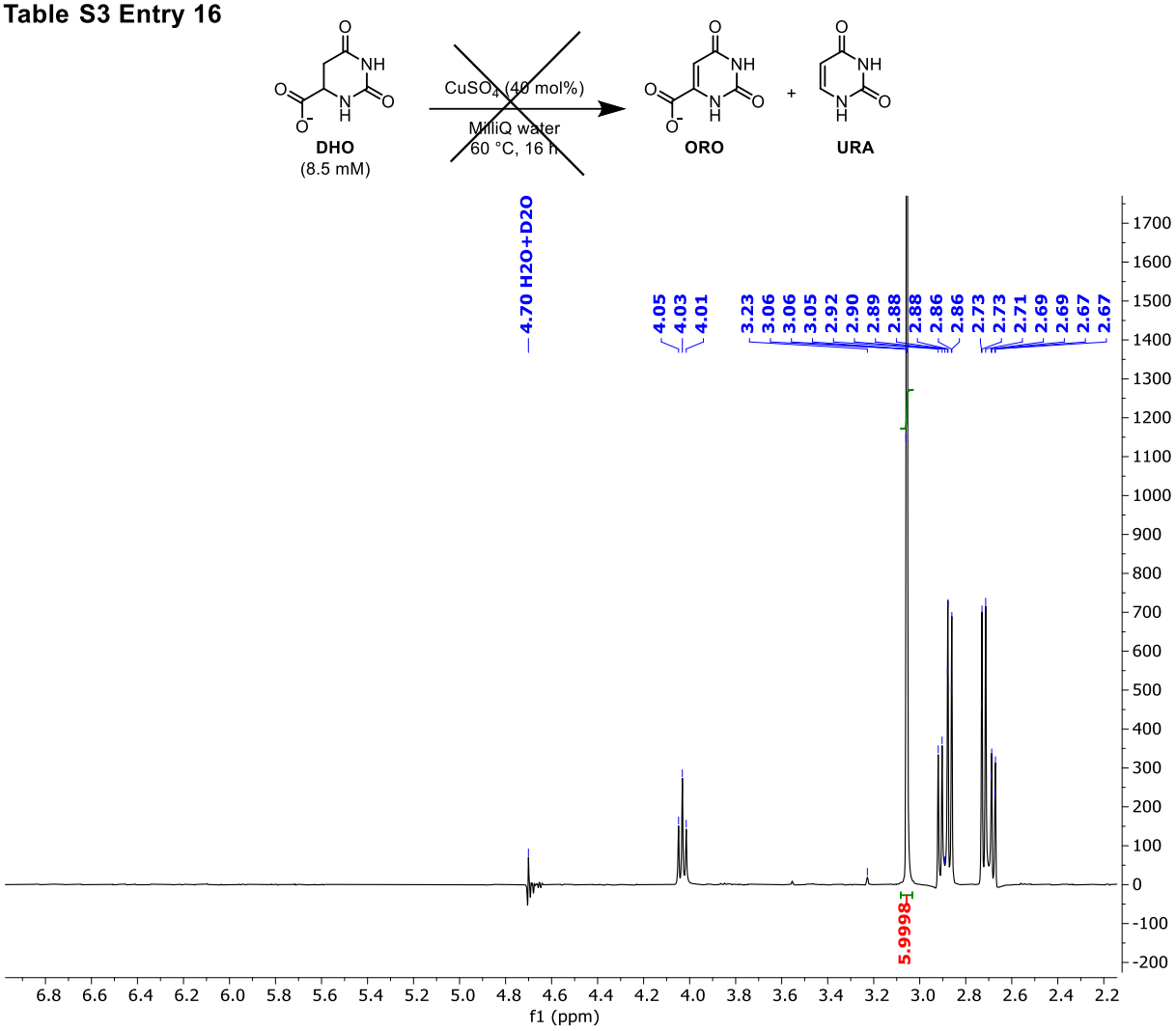


Table S3 Entry 17

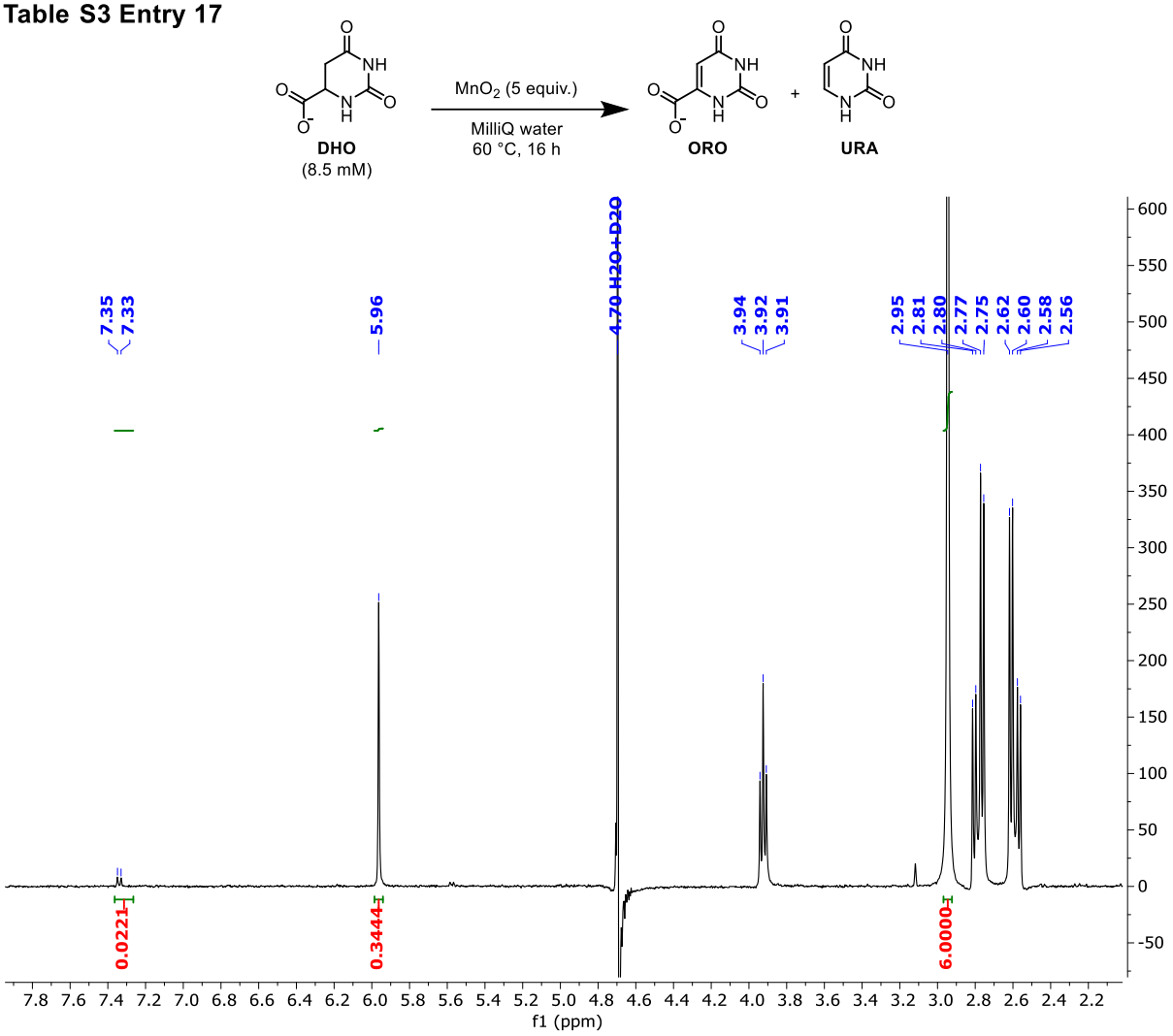


Table S3 Entry 18

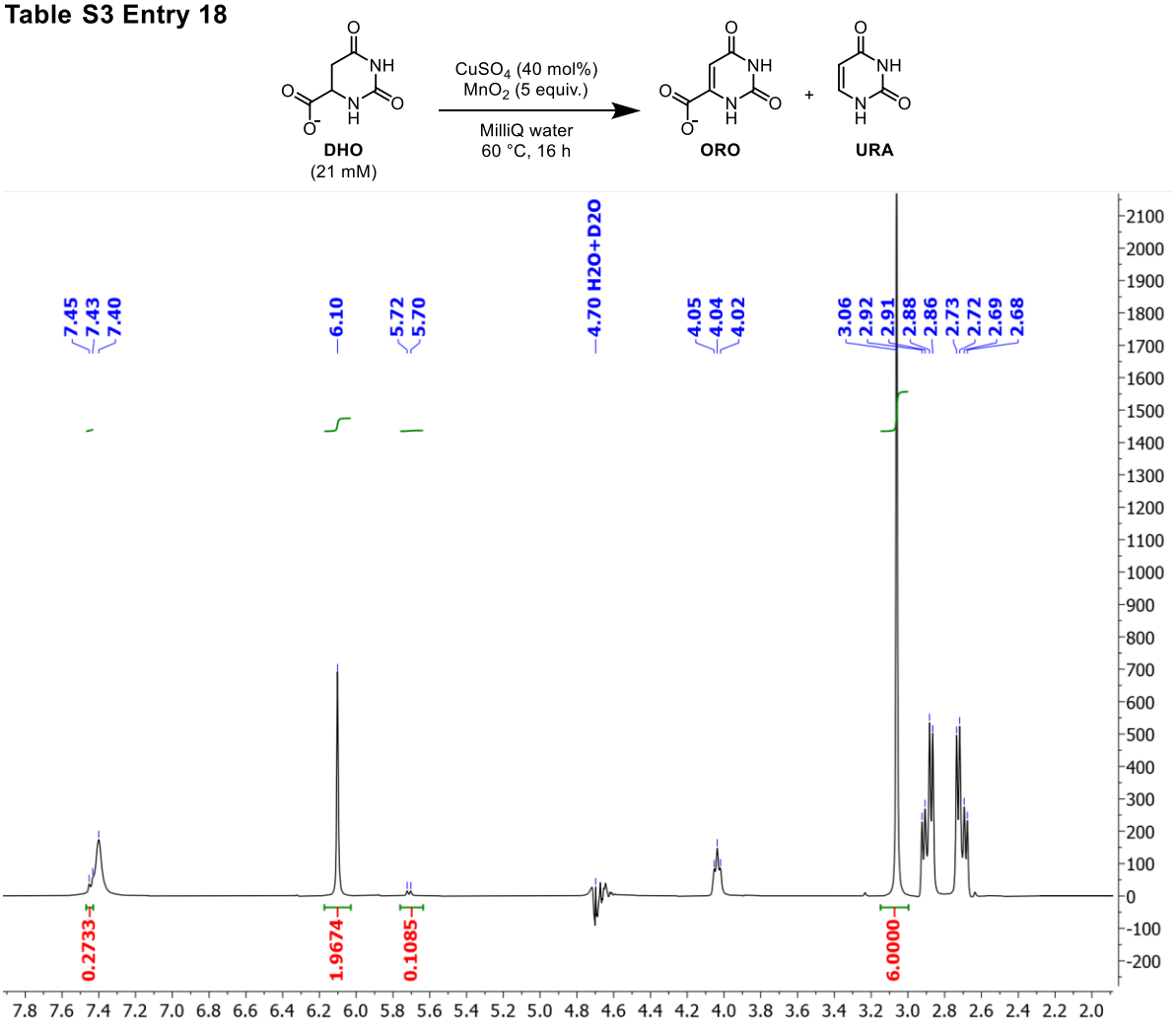


Table S4 Entry 1

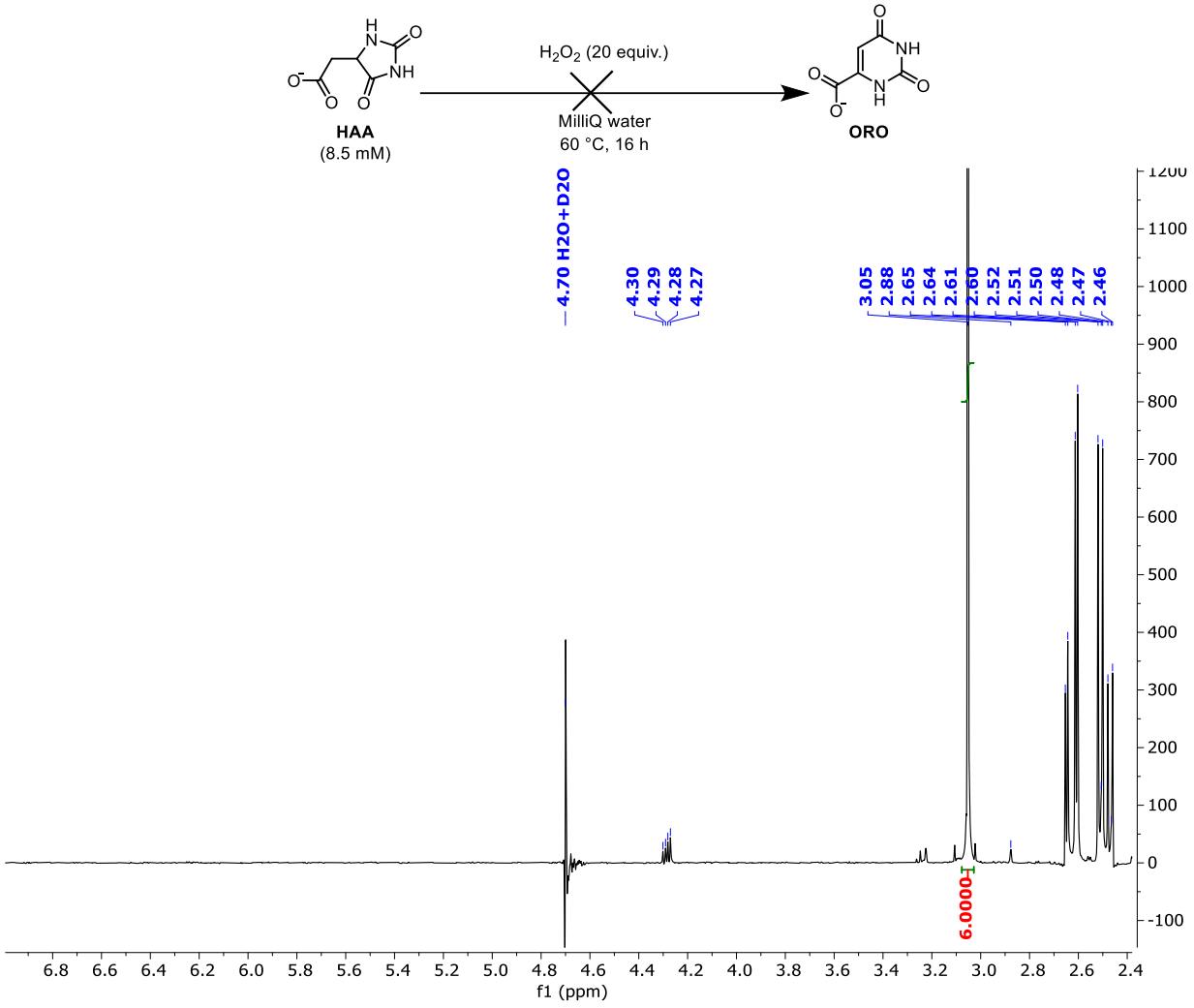
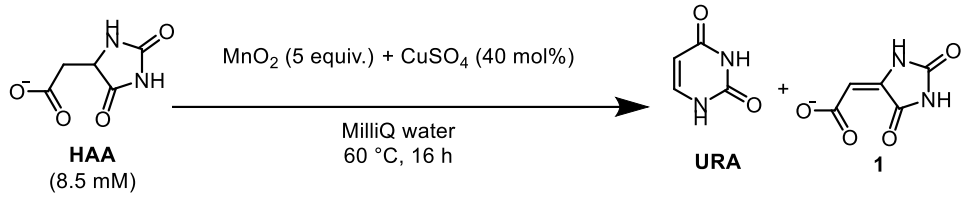
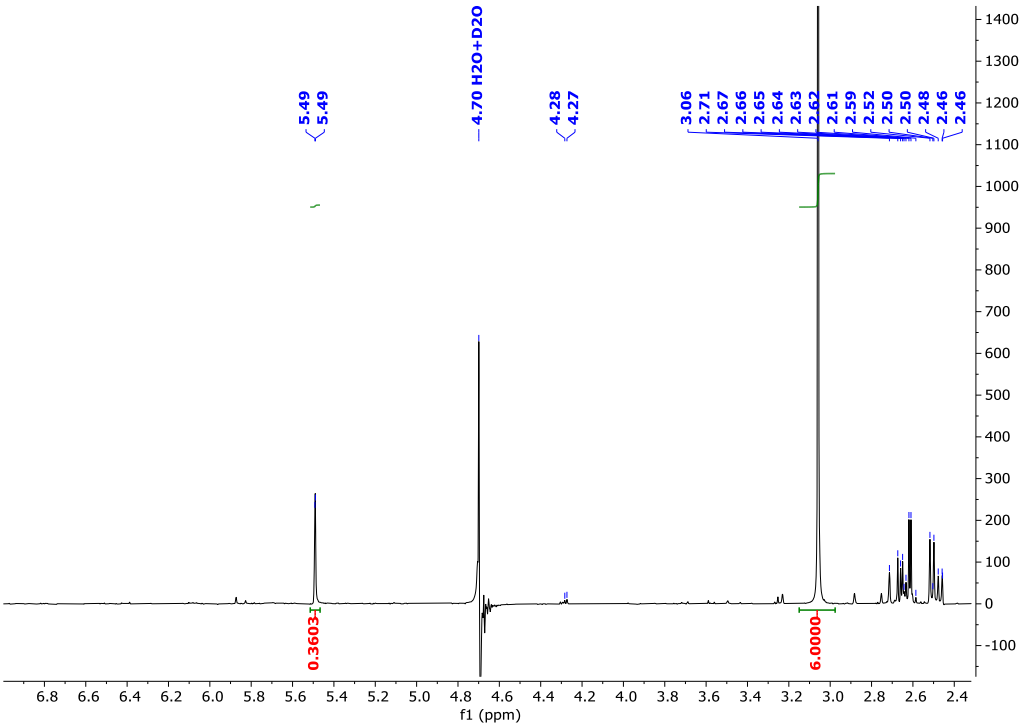
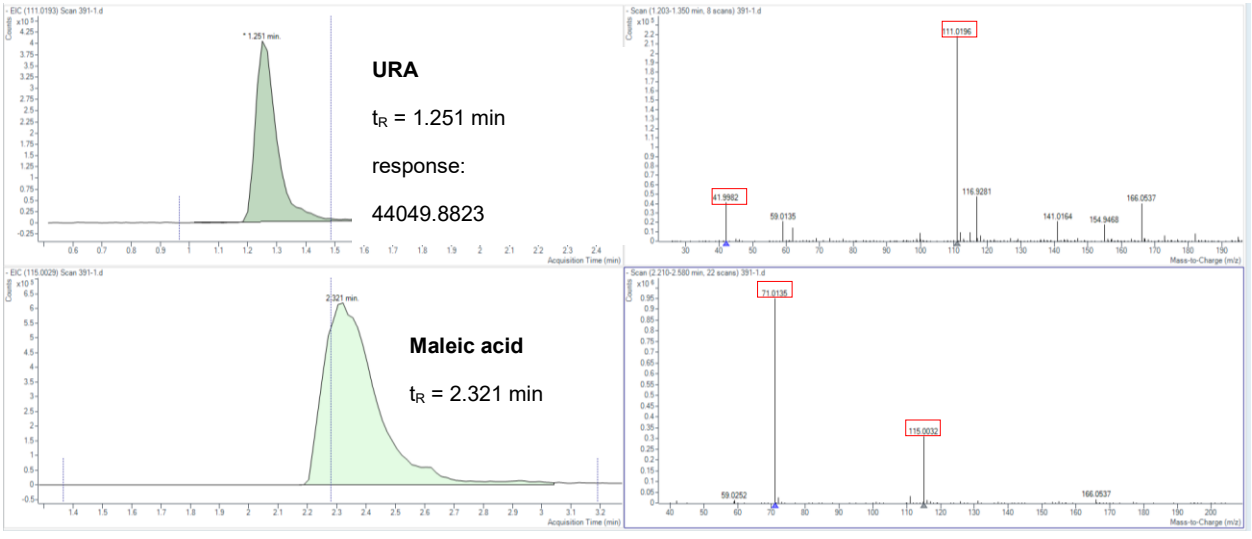


Table S4 Entry 2







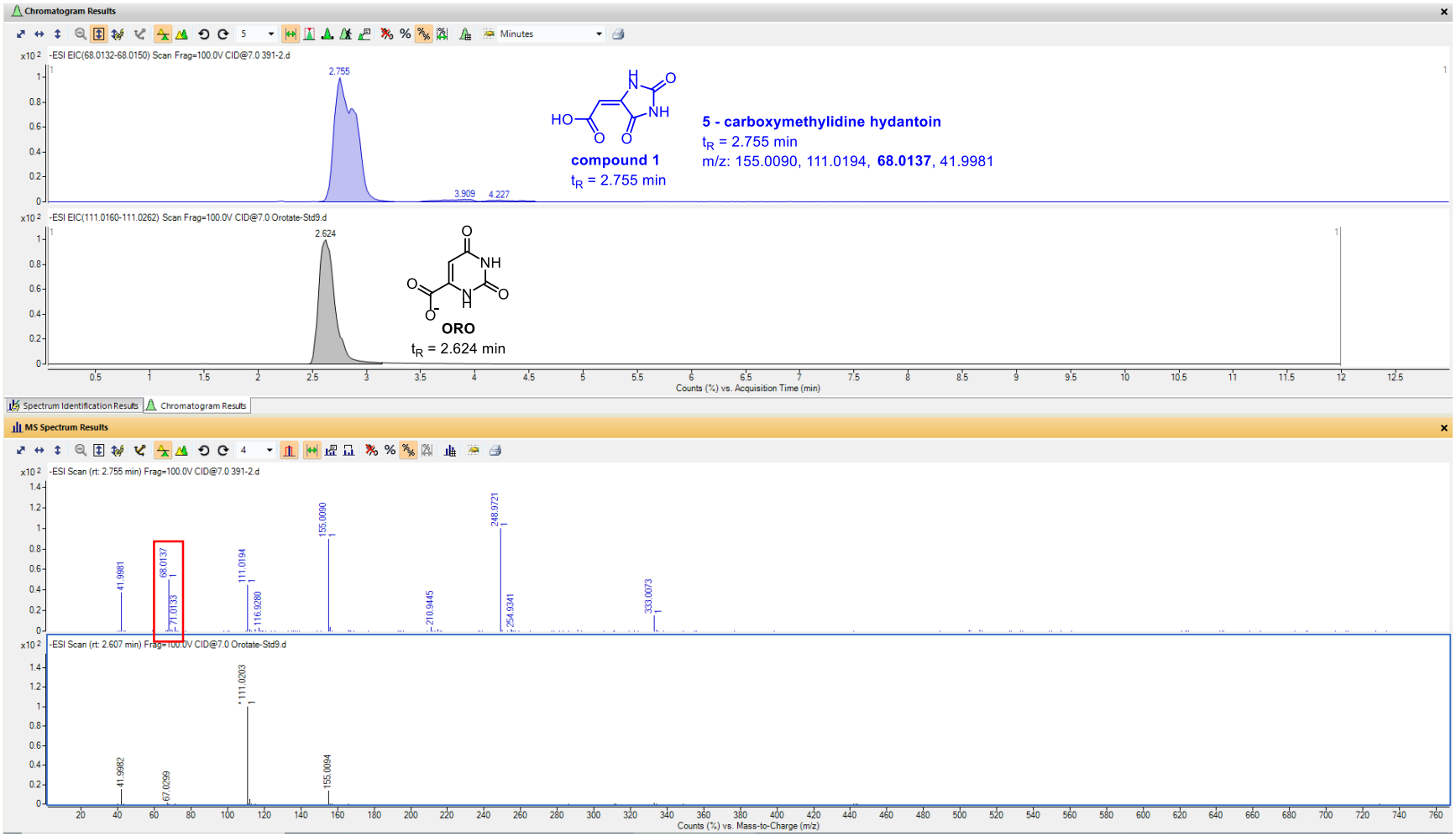
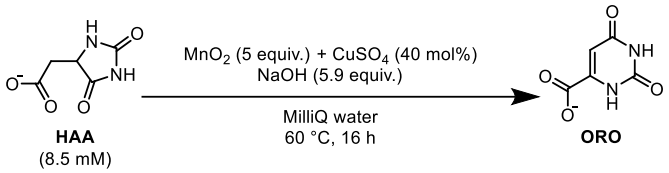

Chromatogram Results
-ESI EIC(68.0132-68.0150) Scan Frag=100.0V CID@7.0 391-2.d
2.755
3.909
4.227
compound 1
t_R = 2.755 min
5 - carboxymethylidine hydantoin
t_R = 2.755 min
m/z: 155.0090, 111.0194, 68.0137, 41.9981
-ESI EIC(111.0160-111.0262) Scan Frag=100.0V CID@7.0 Orotate-Std9.d
2.624
ORO
t_R = 2.624 min
Counts (%) vs. Acquisition Time (min)
Spectrum Identification Results
Chromatogram Results
MS Spectrum Results
-ESI Scan (rt: 2.755 min) Frag=100.0V CID@7.0 391-2.d
41.9981
68.0137
71.0133
111.0194
116.9280
155.0090
210.9445
248.9721
254.9341
333.0073
-ESI Scan (rt: 2.607 min) Frag=100.0V CID@7.0 Orotate-Std9.d
41.9982
67.0299
111.0203
155.0094
Counts (%) vs. Mass-to-Charge (m/z)

Table S4 Entry 3



HAA
(8.5 mM)

ORO

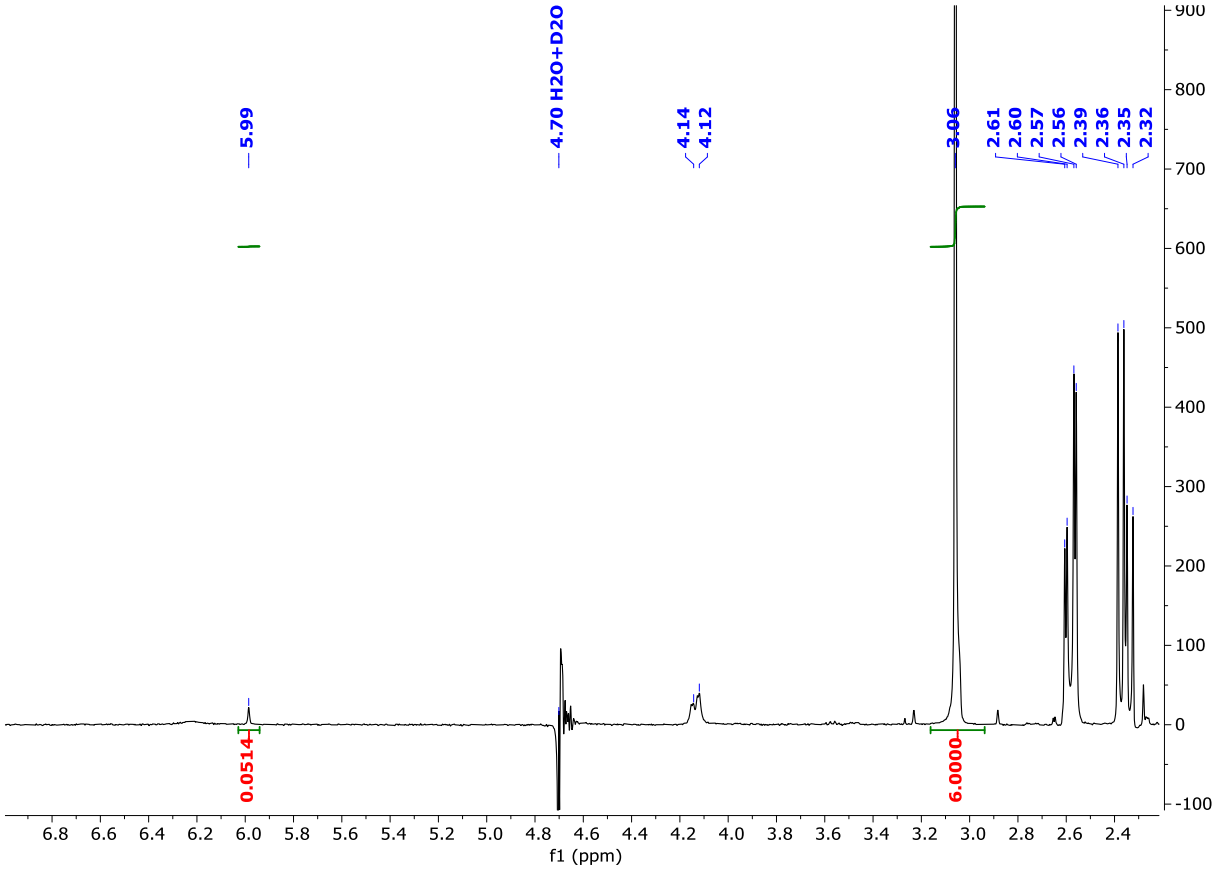
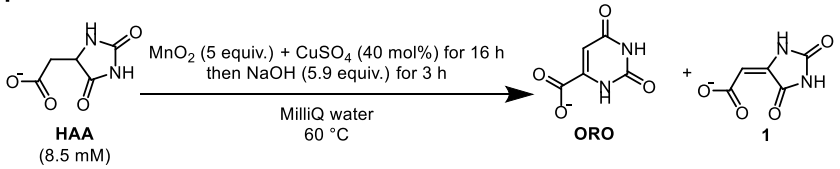
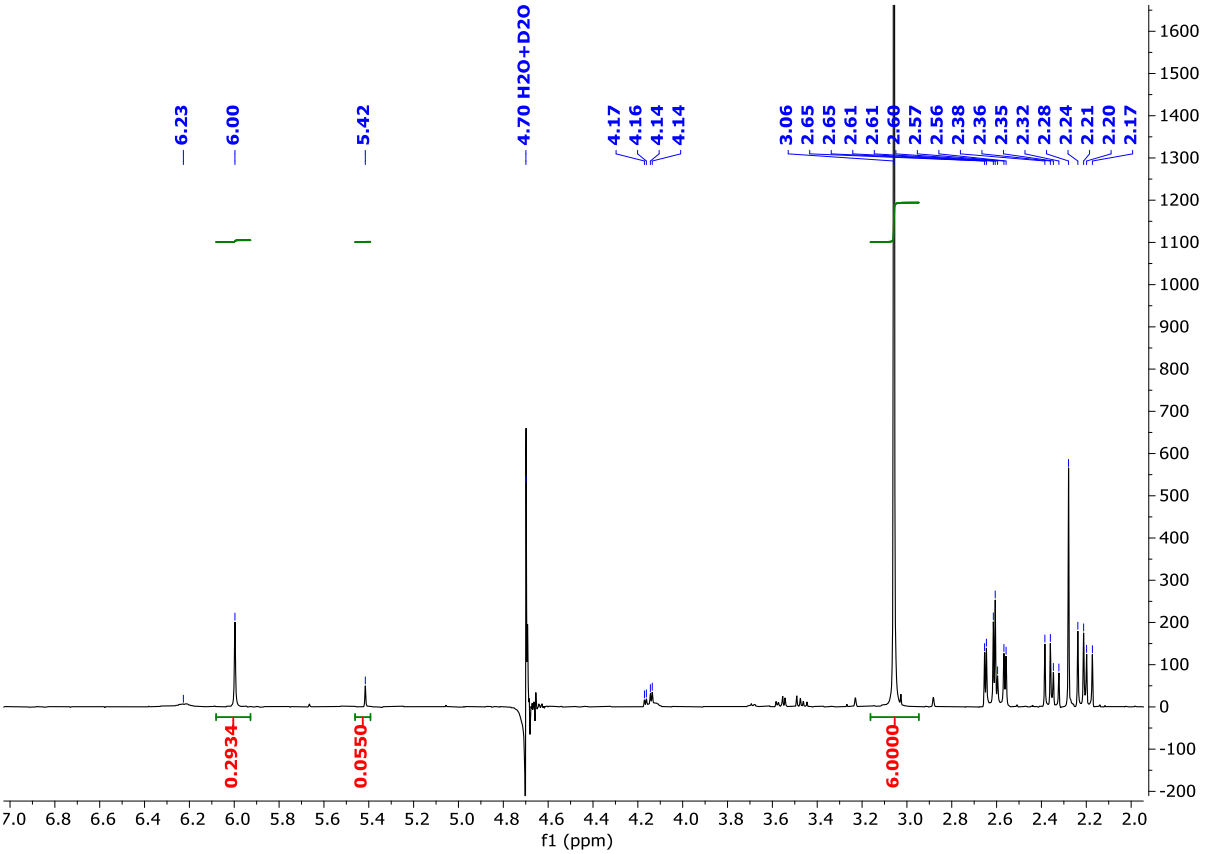


Table S4 Entry 4





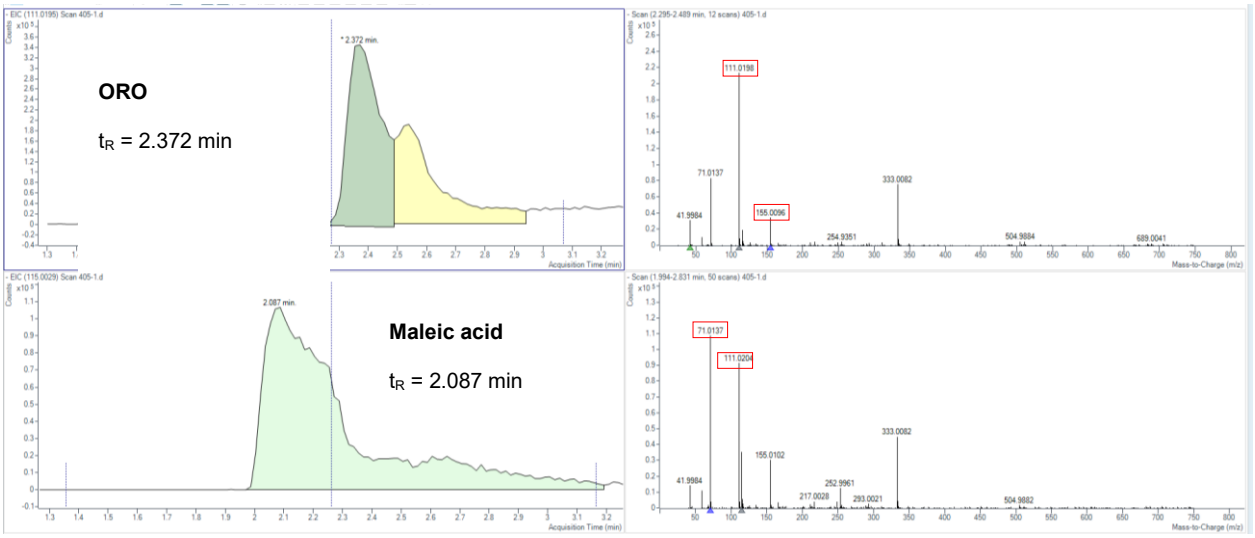


Table S4 Entry 5

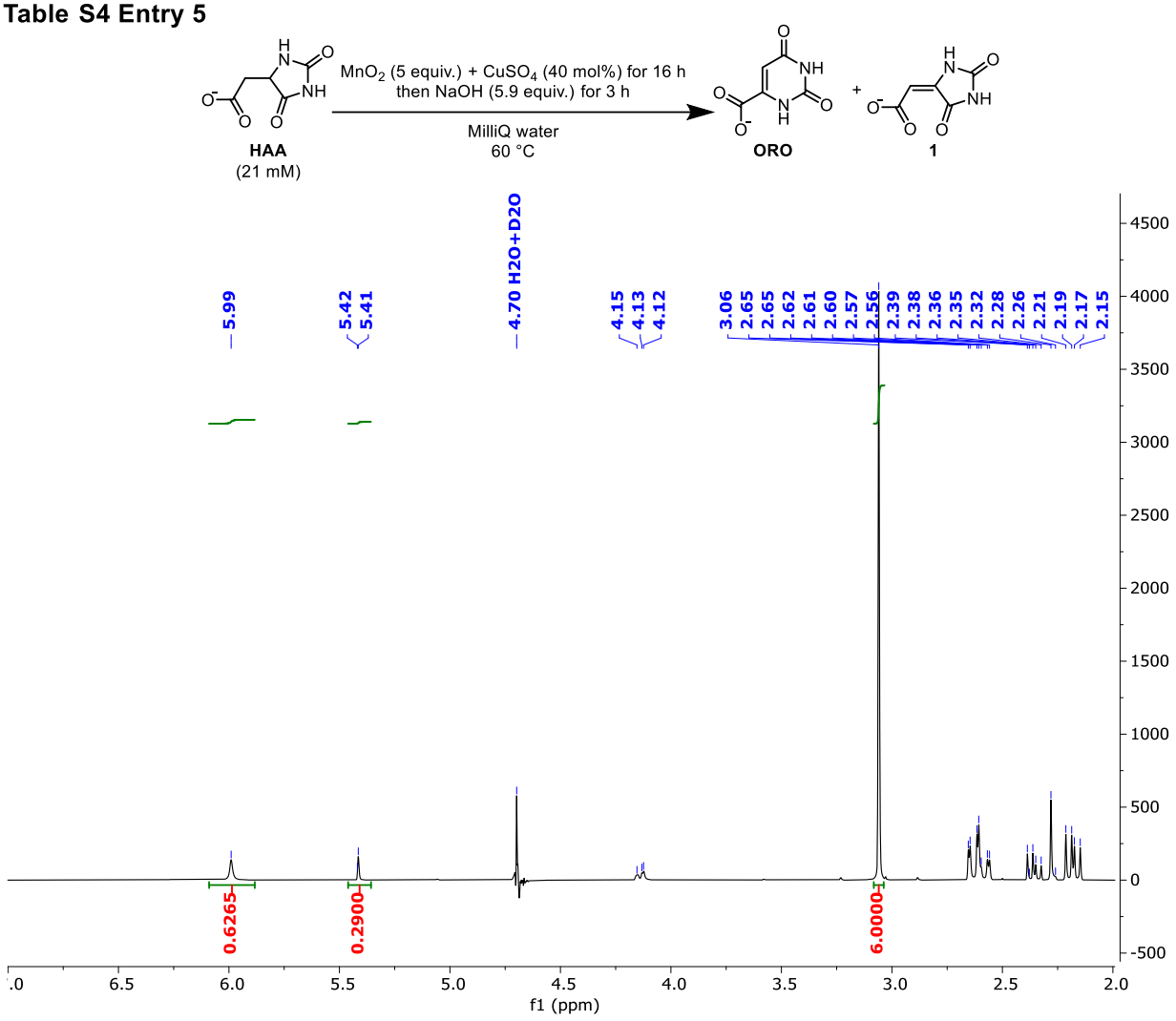


Table S5 Entry 1

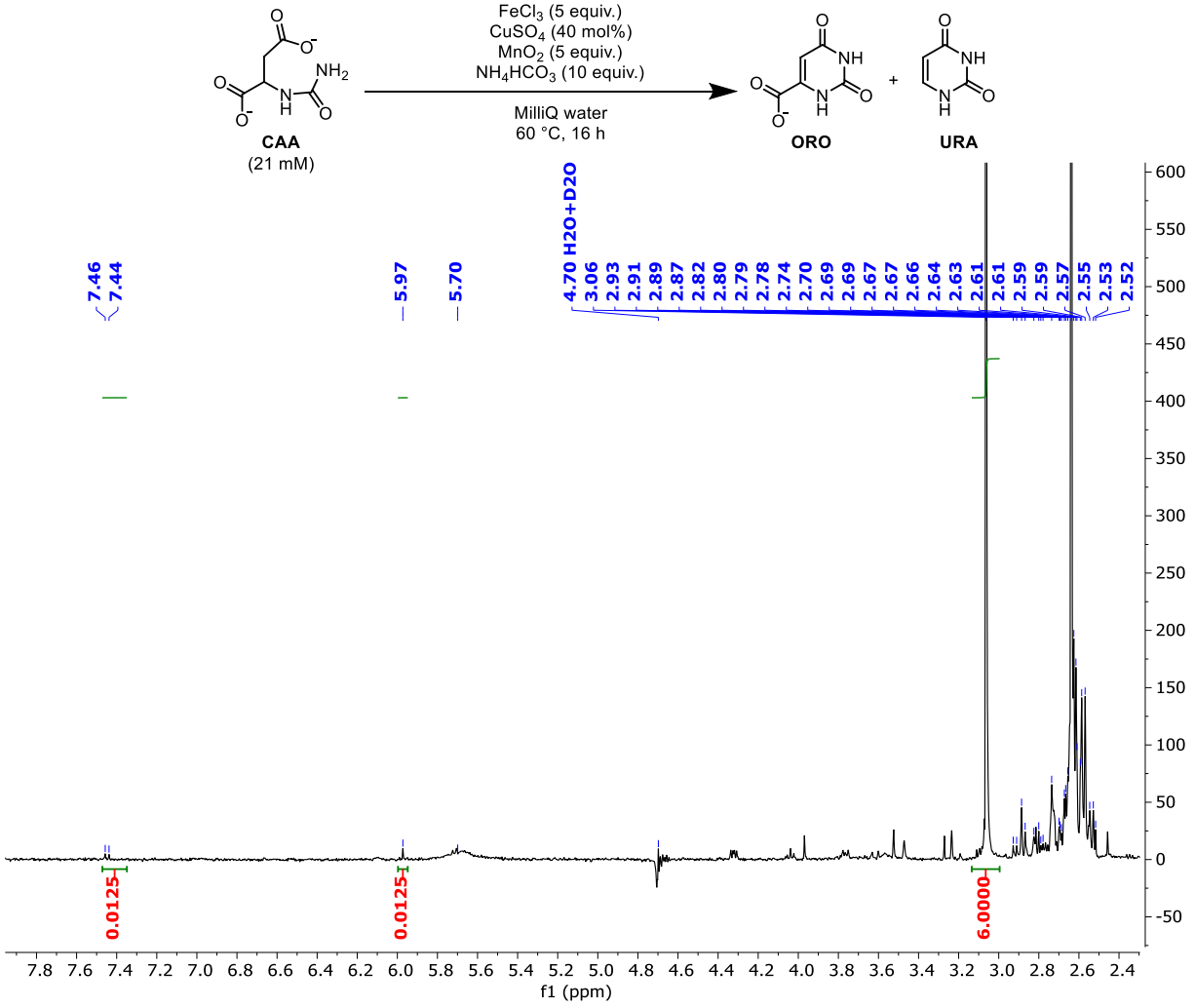
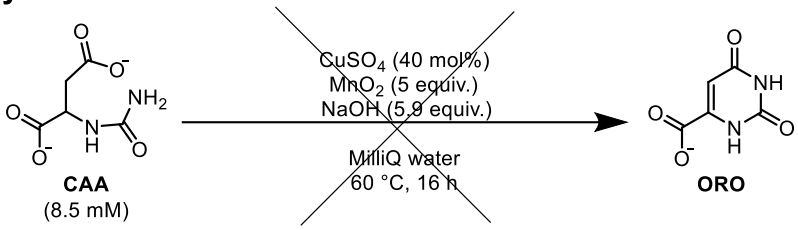


Table S5 Entry 2



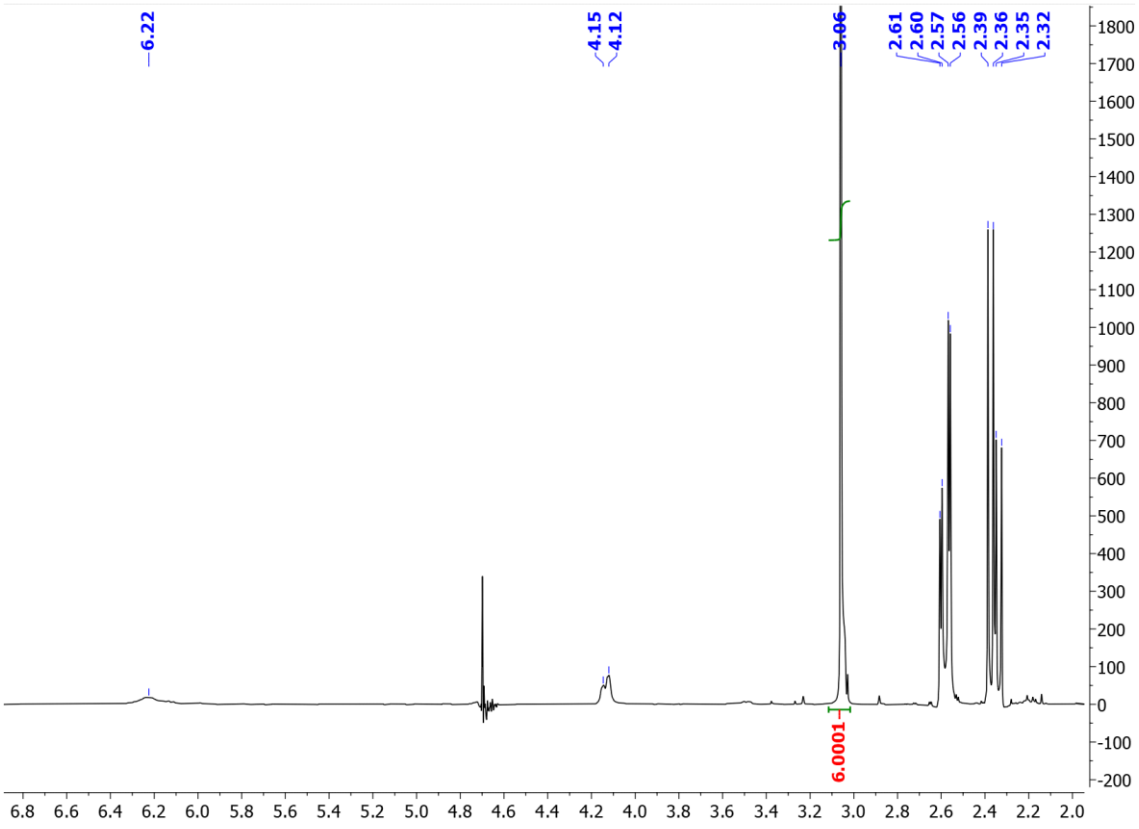


Table S5 Entry 3

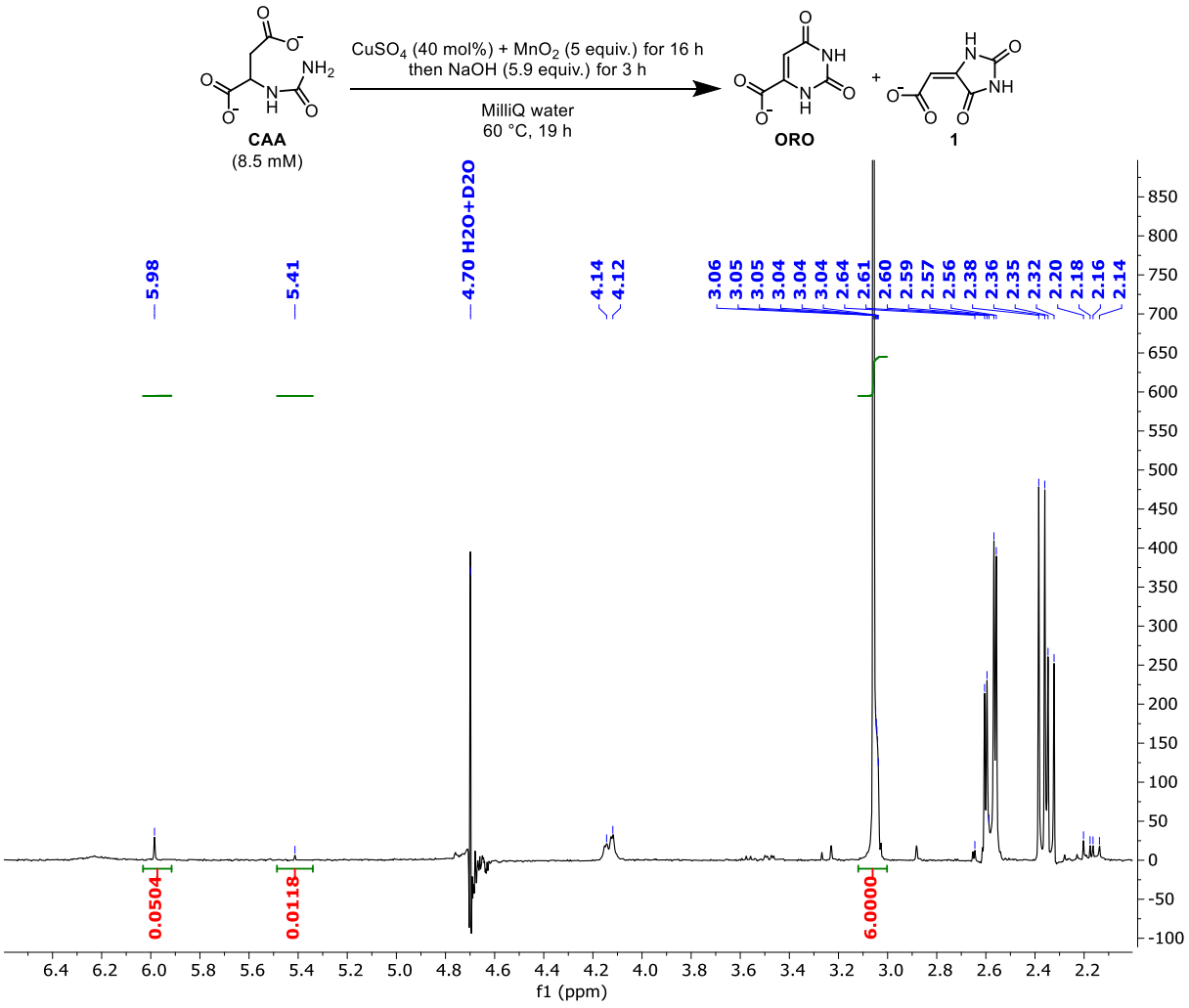


Table S5 Entry 4

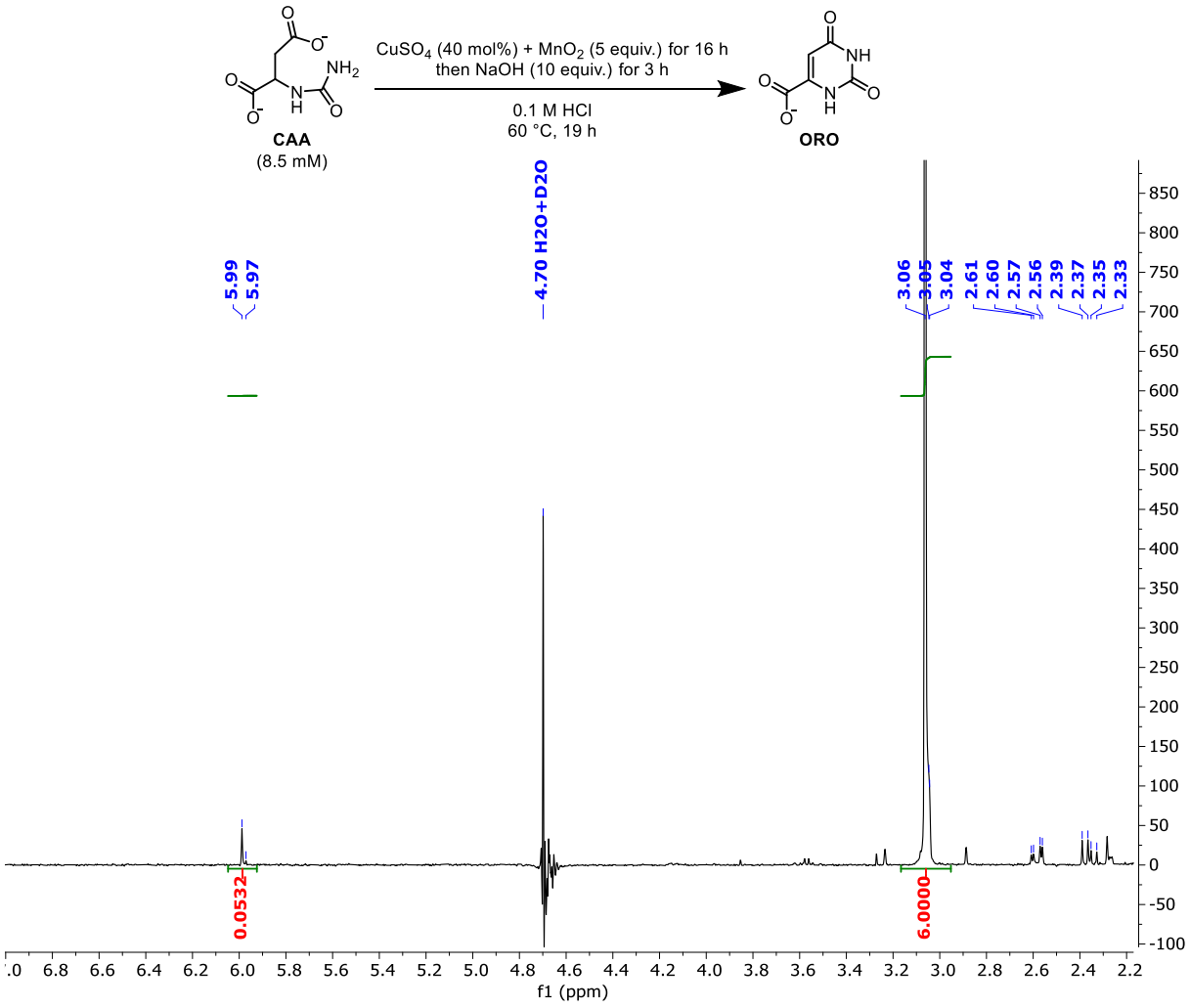
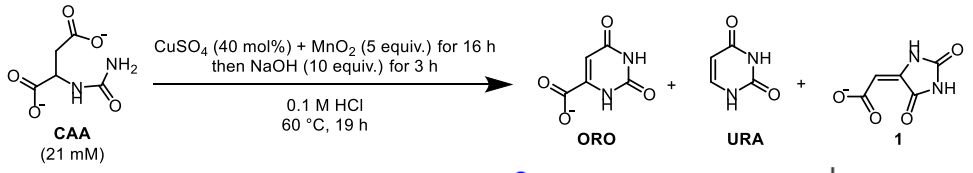
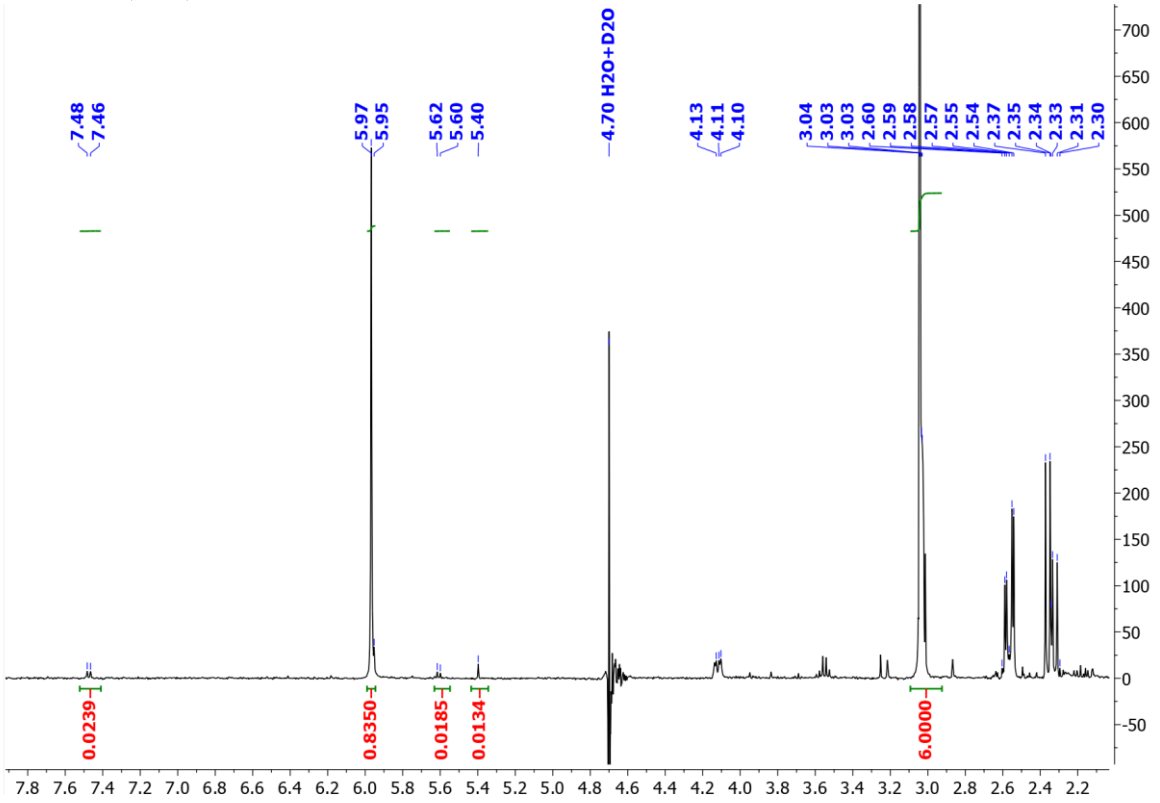
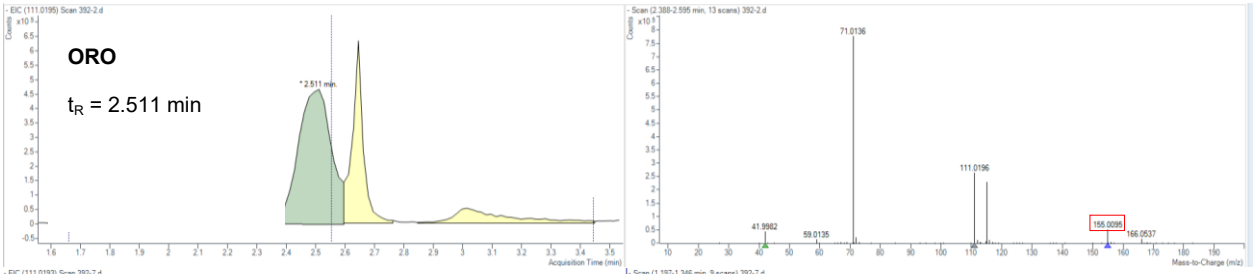
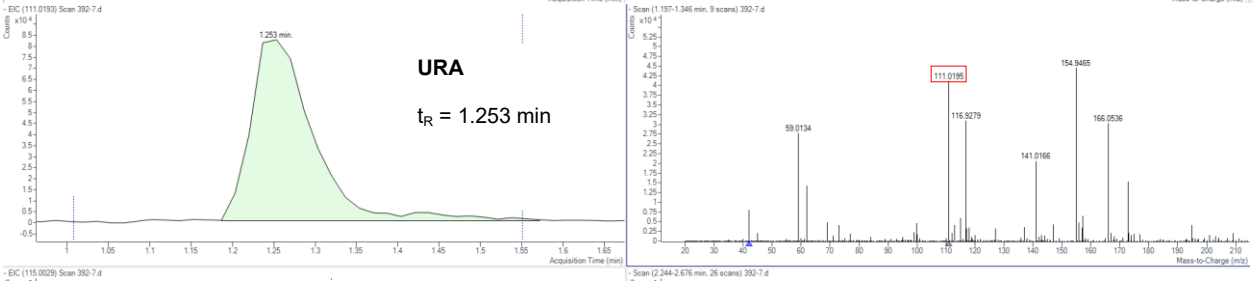


Table S5 Entry 5









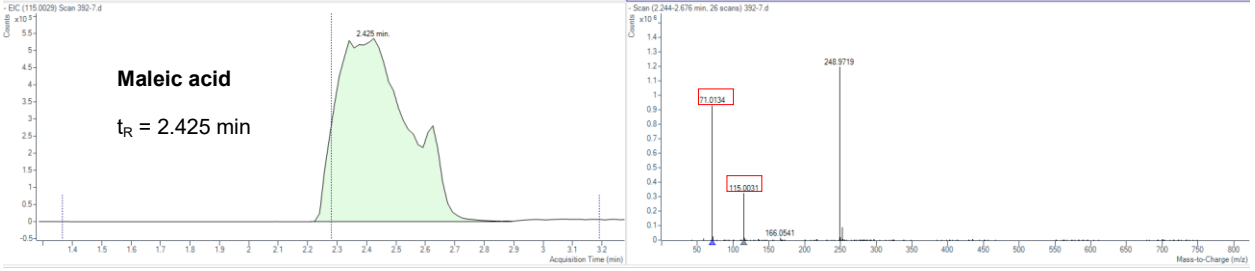
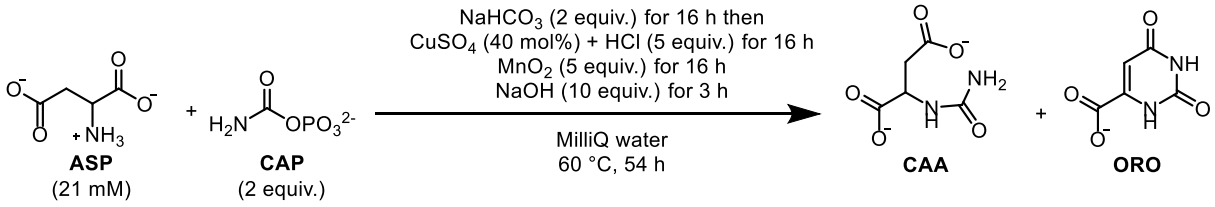
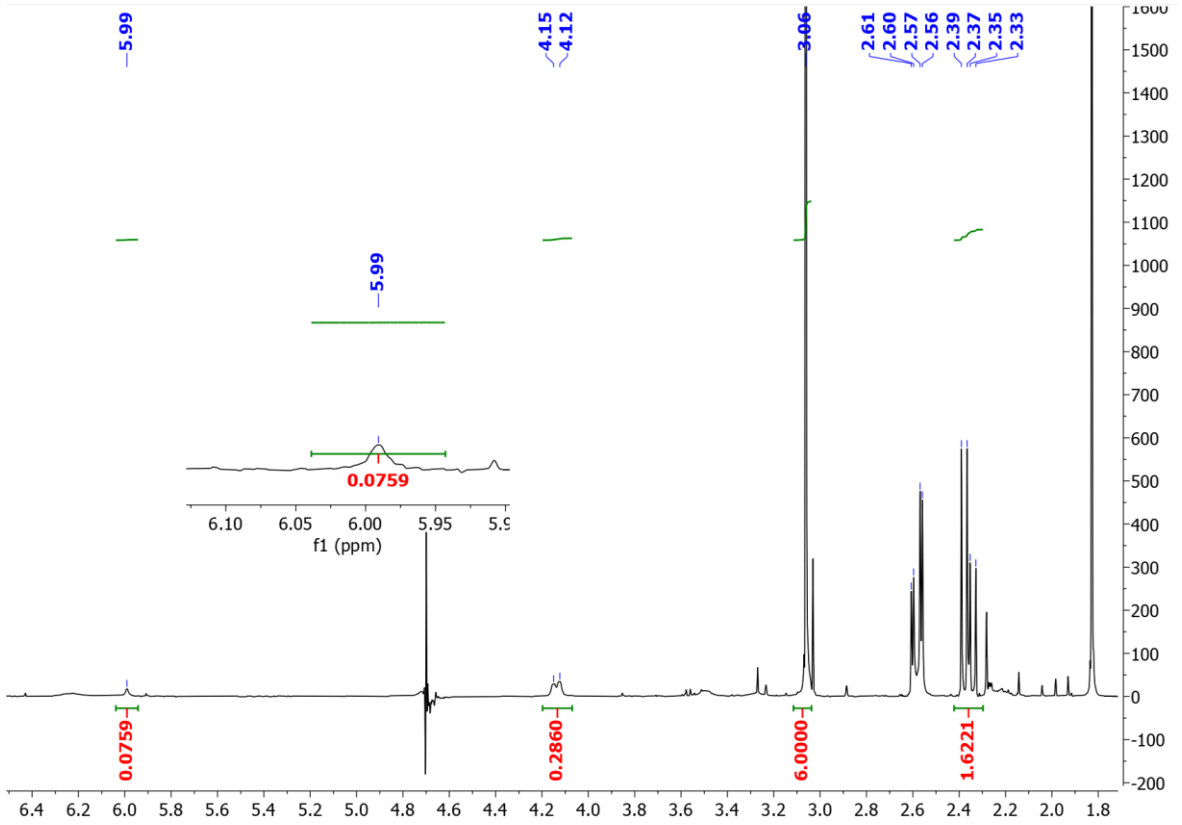
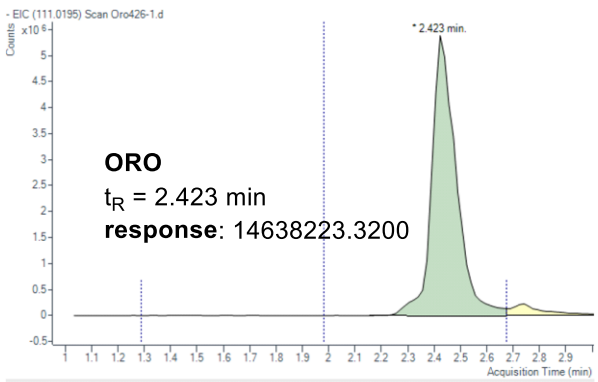
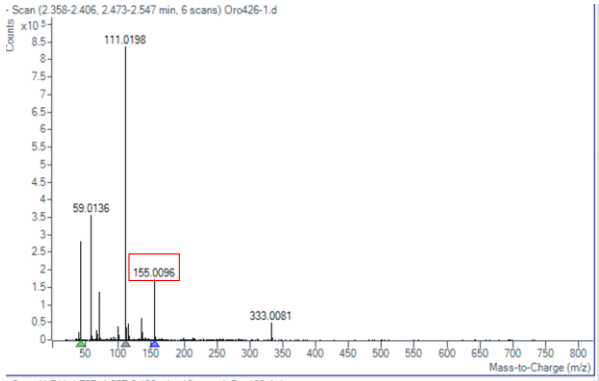


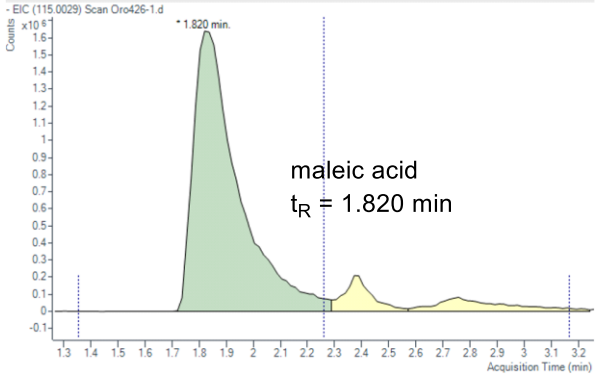
Table S5 Entry 6

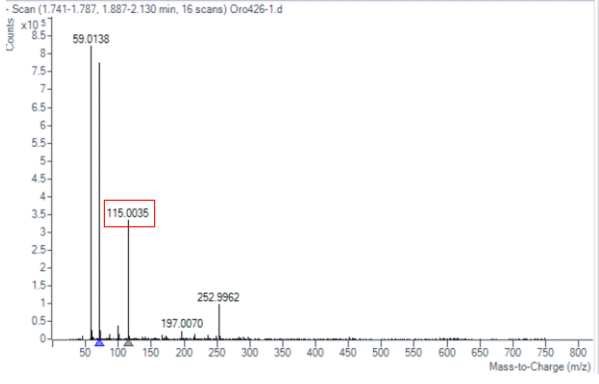












B. Bronsted acid catalyzed cyclopropanes and propargylic alcohols activation

Apart from prebiotic chemistry, during my doctoral studies I explored and finished up several additional projects that were started during my master's internships. As both these projects tackle a topic that is very different from the rest of this thesis, I will summarize them separately below.

One of these projects involved a TfOH/HFIP system catalyzed activation of various organic reactions, including the ring-opening hydroarylation of non-activated cyclopropanes and the dehydrative hydroarylation of trifluoromethylated propargylic alcohols. HFIP plays an important role in organic synthesis owing to its unique properties in enhancing Brønsted or Lewis acid catalysis, although its toxicity to the environment remains debatable. The combination and interaction between TfOH and HFIP provides a simple yet efficient Brønsted acid catalyst system, which allows electronically deactivated substrates to participate in hydroarylation reactions under mild conditions, even without transition metals or strong Lewis acids.

In the hydroarylation reaction of mono-substituted cyclopropanes using catalytic amount of TfOH in HFIP, two different mechanisms appear to be in operation for the C-C bond cleavage. The aryl group-substituted cyclopropanes furnish branched products and proceed through a carbocation intermediate, whereas the electron deficient substrates bearing a carbonyl group lead to linear products through homo-conjugate addition pathway. This part of results was published in *Chemical Science*.[163]

Another study I performed on the CF_3-substituted propargylic alcohols using the same TfOH/HFIP catalytic system indicates that, by tuning the reaction time and the temperature for tertiary alcohols, the produced allenes rearrange into indenes. By tuning the structure of the substrates, 2*H*-chromenes and tri-substituted allylic trifluoromethanes were successfully obtained from secondary alcohols. Insights into the mechanism suggested that the conversion of allenes to indenes was consistent with a Nazarov-type cyclization, and the bis-arylation to the allylic trifluoromethane proceeded through the double formation of carbocation intermediates. The reaction produces water as the only stoichiometric byproduct. This method has implications for the synthesis of pharmaceutical compounds which contain a CF_3 group. This work was published in *The Journal of Organic Chemistry*.[164]

References

[1] Tirard, S. (2014). Abiogenesis. In Encyclopedia of Astrobiology (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 1–1.
[2] Wöhler, F. (1828). Ueber künstliche Bildung des Harnstoffs. Ann. der Phys. und Chemie *88*, 253–256.
[3] Pasteur, L. (1857). Mémoire sur la fermentation appelée lactique. Compt. Rend. *45*, 913–916.
[4] Darwin, Charles (1859), On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.
[5] Oparin, Alexander I. (1924). "V. Abiogenic Organic-Chemical Evolution of Carbon Compounds". The Origin of Life on the Earth
[6] Miller, S.L. (1953). A Production of Amino Acids Under Possible Primitive Earth Conditions. Science *117*, 528 -529.
[7] Oró J. (1960) Synthesis of adenine from ammonium cyanide. Biochem. Biophys. Res. Commun. *2,* 407-412
[8] Fox, S.W., and Matsuno, K. (1983). Genetic takeover and the mineral origins of life. Trends Biochem. Sci. *8*, 341–342.
[9] Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. Cell *31*, 147–157.
[10] Orgel, L.E. (1968). Evolution of the genetic apparatus. J. Mol. Biol. *38*, 381–393.
[11] Gilbert, W. (1986). Origin of life: The RNA world. Nature *319*, 618–618.
[12] Ferris, J.P., Huang, C.-H., and Hagan, W.J. (1988). Montmorillonite: A multifunctional mineral catalyst for the prebiological formation of phosphate esters. Orig. Life Evol. Biosph. *18*, 121–133.
[13]Eschenmoser, A., and Dobler, M. (1992). Warum Pentose- und nicht Hexose-Nucleinsäuren? Teil I. Einleitung und Problemstellung, Konformationsanalyse für Oligonucleotid-Ketten aus 2,3-Dideoxyglucopyranosyl-Bausteinenowie Betrachtungen zur Konformation von A- und B-DNS. Helv. Chim. Acta *75*, 218–259.
[14] Nielsen PE, Egholm M, Berg RH, Buchardt O (1991). "Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide". Science. *254* (5037), 1497–500.
[15] Prigogine, Ilya (1978). "Time, Structure and Fluctuations". Science. 201 (4358): 777–785.
[16] Woese, C.R.; Magrum, L.J.; Fox, G.E. (1978). "Archaebacteria". J Mol Evol. *11* (3): 245–51.
[17] Woese, Carl (June 1998). "The universal ancestor". Proceedings of the National Academy of Sciences of the United States of America. *95* (12): 6854–6859.
[18] Koonin, E. V. (2008). "The deep archaeal roots of eukaryotes". Molecular Biology and Evolution. *25* (8): 1619–1630. K. Gull (2010). "Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes". Proceedings of the Royal Society B: Biological Sciences. *278* (1708): 1009–1018.
[19]Weiss MC, Sousa FL, Mrnjavac N, et al (2016) The physiology and habitat of the last universal common ancestor. Nat Microbiol. 1:16116.
[20] "What is a hydrothermal vent?". National Ocean Service. National Oceanic and Atmospheric Administration.
[21] Russell, M. J., Hall, A. J. and Turner, D. (1989), *In vitro* growth of iron sulphide chimneys: possible culture chambers for origin-of-life experiments. Terra Nova, *1*, 238-241.
[22] Martin, W., and Russell, M.J. (2007). On the origin of biochemistry at an alkaline hydrothermal vent. Philos. Trans. R. Soc. B Biol. Sci. *362*, 1887–1926.
[23] Kelley, D., Karson, J., Blackman, D. et al. (2001). An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N. Nature *412*, 145–149.
[24] De Duve, Christian. (1995). "The Beginnings of Life on Earth." American Scientist, *83*, pp. 428–37.
[25] Wachtershauser, G. (1990). "Evolution of the first metabolic cycles". Proceedings of the National Academy of Sciences. *87* (1): 200–204.
[26] Kamila B. Muchowska, Sreejith J. Varma, and Joseph Moran, (2020). "Nonenzymatic Metabolic Reactions and Life's Origins", Chemical Reviews *120* (15), 7708-7744
[27] B. Rasmussen, J.R. Muhling, and W.W. Fischer. (2021). Greenalite Nanoparticles in Alkaline Vent Plumes as Templates for the Origin of Life. Astrobiology. 246-259.
[28] Cleaves, H. J., Chalmers, J. H., Lazcano, A., Miller, S. L., & Bada, J. L. (2008). A Reassessment of Prebiotic Organic Synthesis in Neutral Planetary Atmospheres. Origins of Life and Evolution of Biospheres, *38*(2), 105–115.
[29] Martin Eric, The principle of continuity in origins of life research. ISSN 2045-5577.
[30] Strecker, Adolph (1850). "Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper". Annalen der Chemie und Pharmacie. 75 (1): 27–45.
[31] Powner, M.W., Gerland, B., and Sutherland, J.D. (2009). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459, 239–242.

[32] Kasting, James & Catling, David. (2003). Evolution of a Habitable Planet. Annual Review of Astronomy and Astrophysics. 41. 429-463.
[33] Powner MW, Sutherland JD, Szostak JW (2010) Chemoselective multicomponent one-pot assembly of purine precursors in water. J Am Chem Soc *132*:16677 16688.
[34] Becker, S.; Feldmann, J.; Wiedemann, S.; Okamura, H.; Schneider, C.; Iwan, K.; Crisp, A.; Rossa, M.; Amatov, T.; Carell, T. (2019). Unified Prebiotically Plausible Synthesis of Pyrimidine and Purine RNA Ribonucleotides. Science *366*, 76−82.
[35] Canavelli P, Islam S, Powner MW (2019) Peptide ligation by chemoselective aminonitrile coupling in water. Nature 571:546–549.
[36] Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. (2015) Common Origins of RNA, Protein and Lipid Precursors in a Cyanosulfidic Protometabolism. Nat. Chem., *7* (4), 301–307.
[37] Vay KL, Salibi E, Song EY, et al. (2020). Nucleic Acid Catalysis under Potential Prebiotic Conditions. Chemistry – An Asian Journal 15(2):214–230.
[38] Orgel LE. (2004). Prebiotic Chemistry and the Origin of the RNA World. Crit Rev Biochem Mol, 39(2):99–123.
[39] Russell MJ, et al. (2014). The drive to life on wet and icy worlds. Astrobiology.*14*:308–343.
[40] Ruiz-Mirazo K, Briones C, de la Escosura A. (2017). Chemical roots of biological evolution: the origins of life as a process of development of autonomous functional systems. *Open Biol. 7:*170050.
[41]Ferris JP. (1993). Catalysis and Prebiotic RNA Synthesis. Origins Life Evol B; 23(5–6):307–315.
[42] Kun, Á., Papp, B. & Száthmary, E. (2008). Computational identification of obligatorily autocatalytic replicators embedded in metabolic networks. Genome Biol. *9,* R51.
[43] Muchowska, K. B. et al. (2017). Metals promote sequences of the reverse Krebs cycle. Nat. Ecol. Evol. *1*, 1716–1721.
[44] Morowitz, Harold. (1992). Beginnings of cellular life. New Haven: Yale University Press.
[45] Morowitz, N. H. (1945). On the Evolution of Biochemical Syntheses. Proc. Natl. Acad. Sci. *31*, 153−157.
[46] Granick, S. (1957). Speculations on the Origins and Evolution of Photosynthesis. Ann. N. Y. Acad. Sci. *69,* 292−308.
[47] Yčas, M. (1974). On Earlier States of the Biochemical System. J. Theor. Biol. *44*, 145−160.
[48] Jensen, R. A. (1976). Enzyme Recruitment in Evolution of New Function. Annu. Rev. Microbiol. *30*, 409−425.
[49] Lazcano, A.; Miller, S. L. (1999). On the Origin of Metabolic Pathways. J. Mol. Evol. *49*, 424−431.
[50] Noda-Garcia, L.; Liebermeister, W.; Tawfik, D. S. (2018). Metabolite− Enzyme Coevolution: From Single Enzymes to Metabolic Pathways and Networks. Annu. Rev. Biochem. *87*, 187−216.
[51] Rosenberg, J.; Commichau, F. M. (2019). Harnessing Underground Metabolism for Pathway Development. Trends Biotechnol. *37*, 29−37.
[52] Smith, E., and Morowitz, H.J. (2004). Universality in intermediary metabolism. Proc. Natl. Acad. Sci. *101*, 13168–13173.
[53] Zahnle, K., Schaefer, L., and Fegley, B. (2010). Earth's earliest atmospheres. Cold Spring Harb. Perspect. Biol. *2*, 1–17.
[54] Lyons, T. W.; Reinhard, C. T.; Planavsky, N. J. (2014). The Rise of Oxygen in Earth's Early Ocean and Atmosphere. Nature *506*, 307−315.
[55] Wächtershäuser, G. (2006). From Volcanic Origins of Chemoautotrophic Life to Bacteria, Archaea and Eukarya. Philos. Trans. R. Soc., B, 361 (1474), 1787−1808.
[56] Mulkidjanian, A. Y. (2009). On the Origin of Life in the Zinc World: 1. Photosynthesizing, Porous Edifices Built of Hydrothermally Precipitated Zinc Sulfide as Cradles of Life on Earth. Biol. Direct. *4*, 26.
[57] Smith, E.; Morowitz, H. J. (2016).The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphere; 1st ed.; Cambridge University Press: New York.
[58] Evans, M.C., Buchanan, B.B., and Arnon, D.I. (1966). A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. Proc. Natl. Acad. Sci. U. S. A. *55*, 928–934.
[59] Weiss, M.C., Sousa, F.L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., and Martin, W.F. (2016). The physiology and habitat of the last universal common ancestor. Nat. Microbiol. *1*, 1–8.
[60] Braakman, R.; Smith, E. (2012).The Emergence and Early Evolution of Biological Carbon Fixation. PLoS Comput. Biol. 8.
[61] Schönheit, P.; Buckel, W.; Martin, W. F. (2016). On the Origin of Heterotrophy. Trends Microbiol. *24*, 12−25.
[62] Monod, J. (1971).Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology; Alfred A. Knopf: New York.
[63] Berg, I. A.; Kockelkorn, D.; Ramos-Vera, W. H.; Say, R. F.; Zarzycki, J.; Hügler, M.; Alber, B. E.; Fuchs, G. (2010). Autotrophic Carbon Fixation in Archaea. Nat. Rev. Microbiol. *8*, 447−460.
[64] Fuchs, G. (2011).Alternative Pathways of Carbon Dioxide Fixation: Insights into the Early Evolution of Life? Annu. Rev. Microbiol. *65*, 631−658.

[65] Sousa, F. L.; Martin, W. F. (2014). Biochemical Fossils of the Ancient Transition from Geoenergetics to Bioenergetics in Prokaryotic One Carbon Compound Metabolism. Biochim. Biophys. Acta, Bioenerg. 1837 (7), 964−981.
[66] Poehlein, A.; Schmidt, S.; Kaster, A.-K.; Goenrich, M.; Vollmers, J.; Thürmer, A.; Bertsch, J.; Schuchmann, K.; Voigt, B.; Hecker, M.; et al. (2012). An Ancient Pathway Combining Carbon Dioxide Fixation with the Generation and Utilization of a Sodium Ion Gradient for ATP Synthesis. PLoS One. *7*, No. e33439.
[67] Ferry, J.G. (1999). Enzymology of one-carbon metabolism in methanogenic pathways. FEMS Microbiol. Rev. *23*, 13–38.
[68] Martin, W. F. (2020). Older Than Genes: The Acetyl CoA Pathway and Origins. Front. Microbiol. *11*, 817.
[69] Ragsdale, S. W. (2008). Enzymology of the Wood-Ljungdahl Pathway of Acetogenesis. Ann. N. Y. Acad. Sci. *1125*, 129−136.
[70] Darnault, C.; Volbeda, A.; Kim, E. J.; Legrand, P.; Vernède, X.; Lindahl, P. A.; Fontecilla-Camps, J. C. (2003). Ni-Zn-[Fe4-S4] and Ni-Ni- [Fe4-S4] Clusters in Closed and Open α Subunits of Acetyl-CoA Synthase/Carbon Monoxide Dehydrogenase. Nat. Struct. Mol. Biol. 10 (4), 271−279.
[71] Buckel, W., and Thauer, R.K. (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. Biochim. Biophys. Acta - Bioenerg. *1827*, 94–113.
[72] Tanaka, K. (1995). Advances in Inorganic Chemistry. Adv. Inorg. Chem. *43*, 409−435.
[73] Huber, C.; Wä chtershäuser, G. (1997). Activated Acetic Acid by Carbon Fixation on (Fe,Ni)S Under Primordial Conditions. Science *276*, 245−247.
[74] Cody, G. D.; Boctor, N. Z.; Filley, T. R.; Hazen, R. M.; Scott, J. H. (2000). Primordial Carbonylated Iron-Sulfur Compounds and the Synthesis of Pyruvate. Science *289*, 1337−1340.
[75] Takahashi, H.; Liu, L. H.; Yashiro, Y.; Ioku, K.; Bignall, G.; Yamasaki, N.; Kori, T. (2006). CO2 Reduction Using Hydrothermal Method for the Selective Formation of Organic Compounds. J. Mater. Sci. *41*, 1585−1589.
[76] He, C.; Tian, G.; Liu, Z.; Feng, S. (2010). A Mild Hydrothermal Route to Fix Carbon Dioxide to Simple Carboxylic Acids. Org. Lett. *12*, 649−651.
[77] Liu, Y.; Chen, S.; Quan, X.; Yu, H. (2015). Efficient Electrochemical Reduction of Carbon Dioxide to Acetate on Nitrogen-Doped Nanodiamond. J. Am. Chem. Soc. *137*, 11631−11636.
[78] Roldan, A.; Hollingsworth, N.; Roffey, A.; Islam, H. U.; Goodall, J. B. M.; Catlow, C. R. A.; Darr, J. A.; Bras, W.; Sankar, G.; Holt, K. B.; et al. (2015). Bio-Inspired CO2 Conversion by Iron Sulfide Catalysts under Sustainable Conditions. Chem. Commun. *51*, 7501−7504.
[79] Varma, S. J.; Muchowska, K. B.; Chatelain, P.; Moran, J. (2018). Native Iron Reduces CO2 to Intermediates and End-Products of the Acetyl- CoA Pathway. Nat. Ecol. Evol. *2*, 1019−1024.
[80] Preiner, M.; Igarashi, K.; Muchowska, K. B.; Yu, M.; Varma, S. J.; Kleinermanns, K.; Nobu, M.; Kamagata, Y.; Tüysü z, H.; Moran, J.; et al. (2020). A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. Nat. Ecol. Evol. *4*, 534−542.
[81] Srinivasan, V.; Morowitz, H. J. (2009). Analysis of the Intermediary Metabolism of a Reductive Chemoautotroph. Biol. Bull. *217*, 222−232.
[82] Orgel, L. E. (2008). The Implausibility of Metabolic Cycles on the Prebiotic Earth. PLoS Biol. *6,* No. e18.
[83] Russell, M. J.; Martin, W. (2004). The Rocky Roots of the Acetyl-CoA Pathway. Trends Biochem. Sci. *29*, 358−363.
[84] Martin, W.; Russell, M. J. (2007). On the Origin of Biochemistry at an Alkaline Hydrothermal Vent. Philos. Trans. R. Soc., B. *362*, 1887− 1925.
[85] Olson, M. V.; Taube, H. (1970). Hydration and Isomerization of Coordinated Maleate. J. Am. Chem. Soc. *92*, 3236−3237.
[86] Rozelle, L. T.; Alberty, R. A. (1957).Kinetics of the Acid Catalysis of the Hydration of Fumaric Acid to Malic Acid. J. Phys. Chem. *61,* 1637−1640.
[87] Gahan, L. R.; Harrowfield, J. M.; Herlt, A. J.; Lindoy, L. F.; Whimp, P. O.; Sargeson, A. W. (1985). Metal Ion Promoted Hydration of Pendant Alkenes and Its Possible Relationship to Aconitase. J. Am. Chem. Soc. *107*, 6231−6242.
[88] Cody, G. D.; Boctor, N. Z.; Hazen, R. M.; Brandes, J. A.; Morowitz, H. J.; Yoder, H. S., Jr. (2001). Geochemical Roots of Autotrophic Carbon Fixation: Hydrothermal Experiments in the System Citric Acid, H2O-(±FeS)-(±NiS). Geochim. Cosmochim. Acta *65*, 3557− 3576.
[89] Zhang, X. V.; Martin, S. T. (2006). Driving Parts of Krebs Cycle in Reverse through Mineral Photochemistry. J. Am. Chem. Soc. *128*, 16032−16033.
[90] Wang, W.; Yang, B.; Qu, Y.; Liu, X.; Su, W. (201). FeS/S/FeS2 Redox System and its Oxidoreductase-like Chemistry in the Iron-Sulfur World. Astrobiology *11*, 471−476.
[91] Kitadai, N.; Nakamura, R.; Yamamoto, M.; Takai, K.; Yoshida, N.; Oono, Y. (2019). Metals Likely Promoted Protometabolism in Early Ocean Alkaline Hydrothermal Systems. Sci. Adv. *5*, No. eaav7848.
[92] Muchowska, K. B.; Varma, S. J.; Moran, J. (2019). Synthesis and Breakdown of Universal Metabolic Precursors Promoted by Iron. Nature *569*, 104−107.

[93] Schulze, U. (1995). Anaerobic Physiology of Saccharomyces Cerevisiae. Ph.D. Thesis . Technical University of Denmark.
[94] Liaw, S.-H.; Kuo, I.; Eisenberg, D. (1995). Discovery of the Ammonium Substrate Site on Glutamine Synthetase, A Third Cation Binding Site. Protein Sci. *4*, 2358−2365.
[95] McMurry, J. E.; Begley, T. P. (2016). The Organic Chemistry of Biological Pathways, 2nd ed.; Roberts and Company: Greenwood Village, CO.
[96] Shapiro BM, Stadtman ER (1970). "The Regulation of Glutamine Synthesis in Microorganisms". Annual Review of Microbiology. *24*: 501–524.
[97] McMurry, J. E., and Begley, T. P. (2016) The Organic Chemistry of Biological Pathways 2nd ed., Roberts and Company Publishers, Inc., Greenwood Village, Colorado.
[98] Taillades, J.; Beuzelin, I.; Garrel, L.; Tabacik, V.; Bied, C.; Commeyras, A. (1998). N-Carbamoyl-α-Amino Acids Rather than Free α- Amino Acids Formation in the Primitive Hydrosphere: A Novel Proposal for the Emergence of Prebiotic Peptides. Origins Life Evol. Biospheres *28*, 61−77.
[99] Tabushi, I.; Yabushita, Y.; Nakajima, T. (1976). Novel Amino Acids Synthesis Using NH 3 through Biogenetic-Type CO 2 Fixation. Tetrahedron Lett. *17*, 4343−4346.
[100] Huber, C.; Wä chtershä user, G. (2003). Primordial Reductive Amination Revisited. Tetrahedron Lett. *44*, 1695−1697.
[101] Kalson, N.-H.; Furman, D.; Zeiri, Y. (2017). Cavitation-Induced Synthesis of Biogenic Molecules on Primordial Earth. ACS Cent. Sci. *3*, 1041−1049.
[102] Sakurai, M.; Yanagawa, H. (1984). Prebiotic Synthesis of Amino Acids from Formaldehyde and Hydroxylamine in a Modified Sea Medium. Origins Life *14,* 171−176.
[103] Nakada, H. I.; Weinhouse, S. (1953).Non-Enzymatic Transamination with Glyoxylic Acid and Various Amino Acids. J. Biol. Chem. *204*, 831−836.
[104] Meisch, H.-U.; Hoffmann, H.; Reinle, W. (1978). Vanadium Catalysis in the Nonenzymatic Transamination of δ-Aminolevulinic Acid. Z. Naturforsch., C: J. Biosci. *33*, 623−628.
[105] Holanda, M. I. D.; Krumholz, P.; Chum, H. L. (1976). Transamination and Amine-Exchange Reactions in the System Iron(II)-Sodium Pyruvate-Aminomethylpyridine. I. Stoichiometry and Reaction Prod- ucts. Inorg. Chem. *15*, 890−893.
[106] Mayer RJ, Kaur H, Rauscher SA, et al. (2021). Mechanistic Insight into Metal Ion-Catalyzed Transamination. J Am Chem Soc. 143(45):19099–19111.
[107] Verhees, C. H.; Kengen, S. W. M.; Tuininga, J. E.; Schut, G. J.; Adams, M. W. W.; de Vos, W. M.; van der Oost, J. (2003). The Unique Features of Glycolytic Pathways in Archaea. Biochem. J. *375*, 231−246.
[108] Berg, I. A.; Kockelkorn, D.; Ramos-Vera, W. H.; Say, R. F.; Zarzycki, J.; Hügler, M.; Alber, B. E.; Fuchs, G. (2010). Autotrophic Carbon Fixation in Archaea. Nat. Rev. Microbiol. *8*, 447−460.
[109] Stincone, A.; Prigione, A.; Cramer, T.; Wamelink, M. M. C.; Campbell, K.; Cheung, E.; Olin-Sandoval, V.; Grü ning, N.-M.; Krü ger, A.; Alam, M. T.; et al. (2015). The Return of Metabolism: Biochemistry and Physiology of the Pentose Phosphate Pathway. Biol. Rev. *90*, 927−963.
[110] McMurry, J. E.; Begley, T. P. (2016). The Organic Chemistry of Biological Pathways, 2nd ed.; Roberts and Company: Greenwood Village, CO.
[111] Butlerow, A. (1861). Formation Synthetique d'une Substance Sucree. C. R. Acad. Sci. *53*, 145−147.
[112] Ritson, D.; Sutherland, J. D. (2012). Prebiotic Synthesis of Simple Sugars by Photoredox Systems Chemistry. Nat. Chem. *4*, 895− 899.
[113] Longo, L. M.; Petrović, D.; Kamerlin, S. C. L.; Tawfik, D. S. (2020). Short and Simple Sequences Favored the Emergence of N-Helix Phospho-Ligand Binding Sites in the First Enzymes. Proc. Natl. Acad. Sci. U. S. A. *117*, 5310−5318.
[114] Messner, C. B.; Driscoll, P. C.; Piedrafita, G.; De Volder, M. F. L.; Ralser, M. (2017). Nonenzymatic Gluconeogenesis-like Formation of Fructose 1,6-Bisphosphate in Ice. Proc. Natl. Acad. Sci. *114*, 7403−7407.
[115] Kofoed, J.; Machuqueiro, M.; Reymond, J.-L.; Darbre, T. (2004). Zinc− Proline Catalyzed Pathway for the Formation of Sugars. Chem. Commun. *0*, 1540−1541.
[116] Krishnamurthy, R.; Pitsch, S.; Arrhenius, G. Mineral Induced Formation of Pentose-2,4-Bisphosphates. Origins Life Evol. Biospheres *29,* 139−152.
[117] Pitsch, S.; Eschenmoser, A.; Gedulin, B.; Hui, S.; Arrhenius, G. (1995). Mineral Induced Formation of Sugar Phosphates. Origins Life Evol. Biospheres *25,* 297−334.
[118] Akouche, M.; Jaber, M.; Maurel, M.-C.; Lambert, J.-F.; Georgelin, T. (2017). Phosphoribosyl Pyrophosphate: A Molecular Vestige of the Origin of Life on Minerals. Angew. Chem., Int. Ed. *56*, 7920− 7923.
[119] Gull, M. (2014). Prebiotic Phosphorylation Reactions on the Early Earth. Challenges *5*, 193−212.
[120] Liu, Z.; Rossi, J.-C.; Pascal, R. (2019). How Prebiotic Chemistry and Early Life Chose Phosphate. Life *9*, 26−41.
[121] Karki, M.; Gibard, C.; Bhowmik, S.; Krishnamurthy, R. (2017). Nitrogenous Derivatives of Phosphorus and the Origins of Life: Plausible Prebiotic Phosphorylating Agents in Water. Life *7*, 32− 59.

[122] Keller, M. A.; Turchyn, A. V.; Ralser, M. (2014). Non-enzymatic Glycolysis and Pentose Phosphate Pathway-like Reactions in a Plausible Archean Ocean. Mol. Syst. Biol. *10*, 725−736.
[123] Keller, M. A.; Zylstra, A.; Castro, C.; Turchyn, A. V.; Griffin, J. L.; Ralser, M. (2016).Conditional Iron and PH-Dependent Activity of a Non- Enzymatic Glycolysis and Pentose Phosphate Pathway. Sci. Adv. *2*, No. e1501235.
[124] Petrychenko, V., Peng, BZ., de A. P. Schwarzer, A.C. et al. Structural mechanism of GTPase-powered ribosome-tRNA movement. *Nat Commun* 12, 5933 (2021).
[125] A.C. Ebrecht et al., The UDP-glucose pyrophosphorylase from Giardia lamblia is redox regulated and exhibits promiscuity to use galactose-1-phosphate. *Biochimica et Biophysica Acta* 1850 (2015) 88–96
[126] E. Strauss, C. Kinsland, Y. Ge, F.W. McLafferty, T.P. Begley. Phosphopantothenoylcysteine synthetase from Escherichia coli. Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J. Biol. Chem*, 276 (2001), 13513-13516
[127] Harrison, S. A.; Lane, N. Life as a Guide to Prebiotic Nucleotide Synthesis. *Nat. Commun*. 2018, 9, 5176.
[128]Powner MW, Sutherland JD, Szostak JW (2010) Chemoselective multicomponent one-pot assembly of purine precursors in water. J Am Chem Soc 132:16677 16688.
[129] Becker, S.; Feldmann, J.; Wiedemann, S.; Okamura, H.; Schneider, C.; Iwan, K.; Crisp, A.; Rossa, M.; Amatov, T.; Carell, T. Unified Prebiotically Plausible Synthesis of Pyrimidine and Purine RNA Ribonucleotides. *Science* 2019, 366, 76−82.
[130] Richert C (2018) Prebiotic chemistry and human intervention. Nat Commun 9:5177.
[131] Sutherland, J. D. The Origin of Life-Out of the Blue. *Angew. Chem., Int. Ed*. 2016, 55, 104−121.
[132] Harrison, S. A.; Lane, N. Life as a Guide to Prebiotic Nucleotide Synthesis. *Nat. Commun*. 2018, 9, 5176.
[133] Yadav, M.; Kumar, R.; Krishnamurthy, R. Chemistry of Abiotic Nucleotide Synthesis. *Chem. Rev*. 2020, 120, 4766−4805.
[134] Lagoja IM and Herdewijn P. One-Step Synthesis of Hypoxanthine from Glycinamide and Diformylurea. *Chem Biodivers* 2004;1(1):106–111.
[135] Lagoja IM and Herdewijn P. A Potential Prebiotic Route to Adenine from Hypoxanthine. *Chem Biodivers* 2005;2(7):923–927.
[136] Hudson, J. S.; Eberle, J. F.; Vachhani, R. H.; Rogers, L. C.; Wade, J. H.; Krishnamurthy, R.; Springsteen, G. A Unified Mechanism for Abiotic Adenine and Purine Synthesis in Formamide. *Angew. Chem., Int. Ed*. 2012, 51, 5134−5137.
[137] Fox, S. W.; Harada, K. Synthesis of Uracil under Conditions of a Thermal Model of Prebiological Chemistry. *Science* 1961, 133, 1923− 1924.
[138] Yamagata, Y.; Sasaki, K.; Takaoka, O.; Sano, S.; Inomata, K.; Kanemitsu, K.; Inoue, Y.; Matsumoto, I. Prebiotic Synthesis of Orotic Acid Parallel to the Biosynthetic Pathway. *Origins Life Evol. Biospheres* 1990, 20, 389−399.
[139] J. F. Nyc, H. K. Mitchell, Synthesis of Orotic Acid from Aspartic Acid. *J. Am. Chem. Soc*. 1947, 69, 1382–1384.
[140] Bruce MJ, Butler AR and Russell KV. Formation of N -Carbamoylaspartic Acid and Its Cyclisation to Orotic Acid. *J Chem Soc Perkin Transactions 2* 1994;0(2):319–321.
[141] Clay AP, Cooke RE, Kumar R, et al (2022) A Plausible Prebiotic One-Pot Synthesis of Orotate and Pyruvate Suggestive of Common Protometabolic Pathways. Angewandte Chemie Int Ed e202112572.
[142] Akouche, M.; Jaber, M.; Maurel, M.-C.; Lambert, J.-F.; Georgelin, T. Phosphoribosyl Pyrophosphate: A Molecular Vestige of the Origin of Life on Minerals. *Angew. Chem., Int. Ed*. 2017, 56, 7920− 7923.
[143] Kim H-J, Benner SA (2017) Prebiotic stereoselective synthesis of purine and noncanonical pyrimidine nucleotide from nucleobases and phosphorylated carbohydrates. *Proc National Acad Sci* 114:11315–11320.
[144] Krishnamurthy R, Guntha S, Eschenmoser A (2000) Regioselective α-phosphorylation of aldose in aqueous solution. *Angew Chem Int Ed*. 39:2281–2285.
[145] Nam I, Nam HG, Zare RN (2018) Abiotic synthesis of purine and pyrimidine ribonucleosides in aqueous microdroplets. *Proc National Acad Sci* 115:36–40.
[146] Jones ME, Lipmann F (1960) CHEMICAL AND ENZYMATIC SYNTHESIS OF CARBAMYL PHOSPHATE. *Proc National Acad Sci* 46:1194–1205.
[147] Liu Z, Wu L-F, Xu J, et al (2020) Harnessing chemical energy for the activation and joining of prebiotic building blocks. *Nat Chem* 12:1–6.
[148] Larralde R, Robertson MP and Miller SL. (1995). Rates of Decomposition of Ribose and Other Sugars: Implications for Chemical Evolution. Proc National Acad Sci. 92(18):8158–8160.
[149] Yi J, Kaur H, Kazöne W, et al (2022) A Nonenzymatic Analog of Pyrimidine Nucleobase Biosynthesis. Angewandte Chemie Int Ed. *61*:e202117211.
[150] R. I. Christopherson, M. E. Jones, (1979). Interconversion of carbamayl-L-aspartate and L-dihydroorotate by dihydroorotase from mouse Ehrlich ascites carcinoma. J. Biol. Chem. *254,* 12506–12512.

[151] Moore SA and Jencks WP. (1982). Reactions of Acyl Phosphates with Carboxylate and Thiol Anions. Model Reactions for CoA Transferase Involving Anhydride Formation. J Biol Chem. 257(18):10874–10881.
[152] Skoog MT and Jencks WP. (1984). Reactions of Pyridines and Primary Amines with N-Phosphorylated Pyridines. J Am Chem Soc. 106(24):7597–7606.
[153] Herschlag D and Jencks WP. (1990). The Effects of Magnesium (II), Hydrogen Bonding, and Steric Factors on Rate and Equilibrium Constants for Phosphoryl Transfer between Carboxylate Ions and Pyridines. J Am Chem Soc. 112(5):1942–1950.
[154] Li Y, Li Y, Liu Y, et al. (2020). Photoreduction of Inorganic Carbon(+IV) by Elemental Sulfur: Implications for Prebiotic Synthesis in Terrestrial Hot Springs. Sci Adv. 6(47):eabc3687.
[155] Mayhew, S. G. (2008). "The Redox Potential of Dithionite and SO−2 from Equilibrium Reactions with Flavodoxins, Methyl Viologen and Hydrogen plus Hydrogenase". European Journal of Biochemistry. 85 (2): 535–547.
[156] R. Ball, J. Brindley, (2019). The Power Without the Glory: Multiple Roles of Hydrogen Peroxide in Mediating the Origin of Life. Astrobiology *19*, 675–684.
[157] Coggins AJ and Powner MW. (2016). Prebiotic Synthesis of Phosphoenol Pyruvate by α-Phosphorylation-Controlled Triose Glycolysis. Nat Chem. 9(4):310–317.
[158] Rode BM and Suwannachot Y. (1999). The Possible Role of Cu(II) for the Origin of Life. Coordin Chem Rev. *190*, 1085–1099.
[159] Liu Z, Mariani A, Wu L, et al. (2018). Tuning the Reactivity of Nitriles Using Cu(Ii) Catalysis – Potentially Prebiotic Activation of Nucleotides. Chem Sci. 9(35):7053–7057.
[160] P. Giraudeau, V. Silvestre; S. Akoka, (2015). Optimizing water suppression for quantitative NMR-based metabolomics: a tutorial review, Metabolomics, *11*, 1041-1055.
[161] Gibard C, Bhowmik S, Karki M, et al. (2017). Phosphorylation, Oligomerization and Self-Assembly in Water under Potential Prebiotic Conditions. Nat Chem. 10(2):212–217.
[162] Krishnamurthy, R., Arrhenius, G. & Eschenmoser, A. (1999). Formation of Glycolaldehyde Phosphate from Glycolaldehyde in Aqueous Solution. Orig Life Evol Biosph *29*, 333–354.
[163] Richmond E, Yi J, Vuković VD, et al (2018) Ring-opening hydroarylation of monosubstituted cyclopropanes enabled by hexafluoroisopropanol. Chem Sci 9:6411–6416.
[164] Noël F, Vuković VD, Yi J, et al (2019) Catalytic Synthesis of Trifluoromethylated Allenes, Indenes, Chromenes, and Olefins from Propargylic Alcohols in HFIP. J Org Chem 84:15926–15947.