

Development of probing strategies to investigate metabolic flux of biofuel production pathways in Clostridia.

Alexander Tomas Wichlacz, MSci (Hons)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2017



Abstract

Currently, fossil fuels contribute a large number of high value chemicals that are used on a daily basis. Crude oil is cracked to give a number of high value chemicals, including vehicle fuels as well as chemicals and solvents that are used daily both commercially and industrially. However, fossil fuel reserves are in decline, with research going into alternatives to obtain these useful chemicals, one of which is biofuels.

Biofuels can be generated in a number of ways, one of which is the fermentation of acetogenic bacteria, microorganisms that generate acetate as a product of anaerobic metabolism. *Clostridium autoethanogenum* is an acetogen that can grow on one carbon gases as its feedstock, and can be used to generate valuable chemicals, with scope to develop the range of metabolic products further.

One aim of this project was to investigate the metabolic flux through pathways of the bacterium using isotopically labelled compounds, which would be assessed by mass spectrometry and NMR. Following on from this, design of inhibitors for the enzymes of the pathways with a view to drive the metabolic

processes towards higher value chemical compounds by 'switching off' other branches of the pathway.

Putative small molecule mimics of acetyl-CoA, SNAC thioesters, were synthesised and tested for uptake and activity in whole cell growth experiments with *C. autoethanogenum*, and determined to be unsuccessful. Further to this, compounds were designed and synthesised to replace pantothenic acid in the growth media, which were not tested in growth experiments.

A library of inhibitor compounds was synthesised and tested against recombinantly purified acetate kinase. A number of compounds were shown to inhibit the enzyme, and the mode of inhibition was determined, as well as IC_{50} and K_i values for each.

This project operated as part of a larger GASCHEM project in the Synthetic Biology Research Centre at the University of Nottingham.

Acknowledgements

For their help and guidance with my PhD project, I would like to acknowledge and thank the following:

My supervisor Prof. Neil R. Thomas for the opportunity to undertake this project, which has been a both testing and rewarding experience.

For their helpful insight, Prof. David A. Barrett and Dr Klaus Winzer.

Dr Huw E. L. Williams for his work collaborating on NMR studies, as well as his guidance and teaching of NMR techniques.

Ronja Breitkopf for her collaboration on the acetate kinase work and the fumarate metabolomics NMR project, as well as her greatly appreciated friendship.

Dr Francesco Zamberlan, Dr James L. Krupa and Dr Terry May for their teaching, guidance and friendship during my PhD studies. In addition, the other members of C28 for their camaraderie on a day-to-day basis.

For mass spectrometry analysis of samples to corroborate with NMR results, Dr Sarah Schatschneider and Laudina Safo.

For their assistance in method development and inhibitor synthesis, Crystal Quist, Freya F. H. Taday and Diana Kołos.

The other collaborators of the GASCHEM project for guidance and friendship during the project.

Finally, the University of Nottingham, LanzaTech and the BBSRC for the funding opportunity to complete this project.

Abbreviations

1,2-DCE	1,2-Dichloroethane
2,3-BD	2,3-Butanediol
2,3-BDH	2,3-Butanediol dehydrogenase (enzyme)
ack	Acetate kinase (gene)
ACK	Acetate kinase (enzyme)
adh	Alcohol dehydrogenase (gene)
ADH	Alcohol dehydrogenase (enzyme)
ALDC	Acetolactate decarboxylase (enzyme)
ALS	Acetolactate synthase (enzyme)
ATP	Adenosine triphosphate
BCD	Butyryl-CoA dehydrogenase (enzyme)
C. acetobutylicum	*Clostridium acetobutylicum*
C. autoethanogenum	*Clostridium autoethanogenum*
C. beijerinckii	*Clostridium beijerinckii*
C. ljungdahlii	*Clostridium ljungdahlii*
C. ragsdalei	*Clostridium ragsdalei*
CDI	Carbonyldiimidazole
CoA	Co-enzyme A
CODH	Carbon monoxide dehydrogenase (enzyme)

CRT	3-hydroxybutyryl-CoA dehydratase (enzyme)
DCC	Dicyclohexylcarbodiimide
DCU	Dicyclohexylurea
DCM	Dichloromethane
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSS	3-(Trimethylsilyl)-1-propanesulfonic acid sodium salt
E. coli	*Escherichia coli*
EDC.HCl	1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride
ESI$^+$	Positive electrospray ionisation
ESI$^-$	Negative electrospray ionisation
ETBE	Ethyl *tert*-butyl ether
FRD	Fumarate Reductase (enzyme)
GC-MS	Gas chromatography mass spectrometry
HBD	Hydroxybutyric dehydrogenase (enzyme)
HMDB	Human Metabolome Database
HPLC	High-performance liquid chromatography

HRMS	High resolution mass spectrometry
LB	Luria broth
LC-MS	Liquid chromatography mass spectrometry
ldh	Lactate dehydrogenase (gene)
LDH	Lactate dehydrogenase (enzyme)
MES	2-(N-morpholino)ethanesulfonic acid
NBT/BCIP	Nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl-phosphate
NHS	*N*-hydroxysuccinimide
NMR	Nuclear magnetic resonance
NPK	Nitrogen Phosphorous Potassium fertiliser
OD	Optical density
PCR	Polymerase chain reaction
PETC-MES	2-ethanesulfonic acid media
pfl	Pyruvate formate lyase (gene)
PFL	Pyruvate formate lyase (enzyme)
PFOR	Pyruvate:Ferredoxin oxidoreductase (enzyme)
PMP	Paramethoxyphenyl
pta	Phosphotransacetylase (gene)
PTA	Phosphotransacetylase (enzyme)

PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RNA	Ribonucleic acid
SAR	Specific activity remaining
SNAC	*N*-acetylcysteamine
Syngas	Synthesis gas
TBAF	Tetra-*n*-butylammonium fluoride
TBS	*Tert*-butyldimethylsilyl
TEMED	*N,N,N′,N′*-tetramethylethylenediamine
TEMPO	2,2,6,6-Tetramethylpiperidinyloxy
TFA	Trifluoroacetic acid
THF	Tetrahydrofolate
THL	Thiolase (enzyme)
TLC	Thin layer chromatography
TMS	Tetramethylsilane

Contents

Abstract ii

Acknowledgements iv

Abbreviations v

Contents ix

1.0 Introduction 1

1.1 Fossil fuel 1

1.2 Biofuels 3

1.2.1 Plant biomass feedstocks 4

1.2.2 Syngas 8

1.3 Acetogenic bacteria 12

1.4 *Clostridium autoethanogenum* 17

1.4.1 *C. autoethanogenum* metabolic products 19

1.5 Metabolic engineering in *C. autoethanogenum* 24

1.5.1 Feedstock and pH variability 24

1.5.2 Gas stripping 26

1.5.3 Genetic knockouts 27

1.6 Isotopic labelling and metabolic flux analysis 31

1.6.1 Acetyl-CoA 34

1.6.2 Metabolites of the reductive acetyl-CoA pathway 41

1.7 Enzyme inhibition 42

1.7.1 Acetate kinase 43

1.8 Project Aims 46

2.0 Results and Discussion 52

Chapter I – Metabolic flux in *C. autoethanogenum* 52

SNAC thioester synthesis 52

Whole cell growth experiments using acetyl-SNAC thioesters 57

^{13}C-Acetyl panthetheine synthesis 64

Fumarate growth experiments 70
Summary 83
Chapter II – Synthesis of acetate kinase inhibitors 85
Synthesis of putative β-ketophosphonate inhibitors 88
Synthesis of putative α-carbonylphosphonate inhibitors 98
Synthesis of putative phosphonic acid inhibitors 102
Synthesis of putative monomethyl phosphonate ester inhibitors 106
Summary 110
Chapter III – Acetate kinase 112
Purification of recombinant acetate kinase from *C. autoethanogenum* 114
Determination of kinetic parameters of acetate kinase 119
Inhibition of acetate kinase from *C. autoethanogenum* 126
Summary 146
3.0 Conclusions 148
4.0 Experimental 154
General procedures 154
General chemical procedures 154
General NMR procedures 155
General biological procedures 158
Whole cell lysis procedure 160
4.1 Pathway metabolites and mimic synthesis 161
4.2 Acetate kinase inhibitors 183
β-Ketophosphonate inhibitor synthesis 183
α-Carbonylphosphonate inhibitor synthesis 188
Phosphonic acid inhibitors 196
Monomethyl ester inhibitor synthesis 205
4.3 Acetate kinase recombinant purification 216
PCR amplification 217

Plasmid miniprep 219
ACK purification 221
SDS & Western Blot analysis 221
Enzyme kinetic parameter determination 224
Acetate kinase inhibition 226

5.0 References 228

6.0 Appendix 238

Acetate kinase gene and protein sequences 238

GenDB computational genome analysis 239

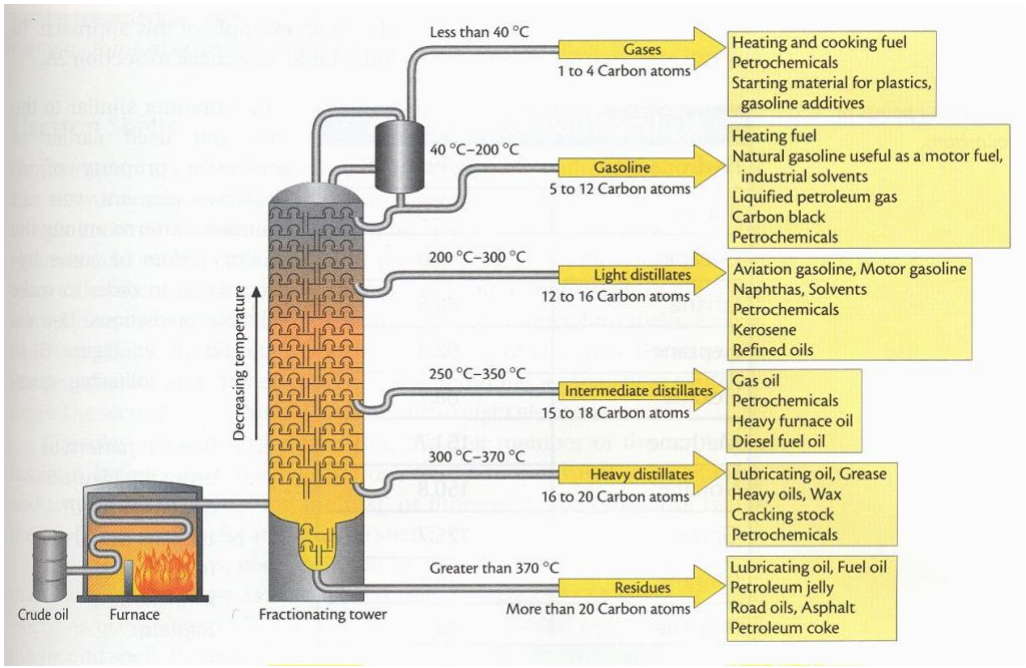
1.0 Introduction

1.1 Fossil fuel

Fossil fuels are a resource that we have been tapping into for a long time. However, diminishing fossil fuel levels, in conjunction with rising prices of oil, with few suitable alternatives, represents a current global problem. In 2009, Shafiee *et al.* predicted that we currently have enough oil reserves left, based upon current rate of consumption, to last until 2043.[1] This means alternatives need to be found in the near future so that we are no longer dependent upon fossil fuels at that point in time. In addition to their long-term non-sustainability, fossil fuels have a negative impact on the environment. This is due to the release of carbon dioxide and other pollutants when they are combusted. This has led to links to global warming *via* the greenhouse effect and significant climate change as direct effects of fossil fuel combustion.[2, 3]

One fossil fuel which generates useful fuels used on a daily basis is crude oil, which formed over millions of years from fossilised organic materials. These decompose over time to form a mixture of hydrocarbons which can be used as fuel.[4]

Crude oil is separated into its constituent parts *via* an industrial process known as cracking, shown in **Scheme 1.** This yields fuels for use in vehicles, such as petrol and diesel, but also gases such as propane and butane, as well as bitumen which is utilised in road surfacing and a number of other industrially useful or commercially viable products.[5]



Scheme 1 – Fractional distillation column for industrial separation of crude oil.[6]

One potential alternative to crude oil would be the use of oil shale,[7] containing kerogens, which can be broken down into liquid hydrocarbons. It was estimated that globally there are between 4.8 to 5 trillion barrels of shale oil in 2010.[8] However, this does not contain petroleum or diesel fuels, meaning that it cannot be used for traditional vehicle fuel. Thus, at present, its

application is more applicable as a substitute for coal in power plant energy generation, a clear disadvantage to crude oil. In addition to its more limited use, its extraction cost is greater than that of crude oil, making it less economically viable on larger scale. As shale consists of hydrocarbons which produce energy by combustion, it also has the same inherent issue with release of greenhouse gases as fossil fuels do, making it of environmental concern also as a fossil fuel alternative.[9]

Due to the decline in fossil fuel reserves, the negative impact that these fuels have on the environment and the lack of alternatives at present, other sources of the products of fossil fuels are currently under investigation. One area that is of particular interest is biofuels, a more sustainable source of higher value chemical compounds, one method for generation of which is from waste that is already produced.

1.2 Biofuels

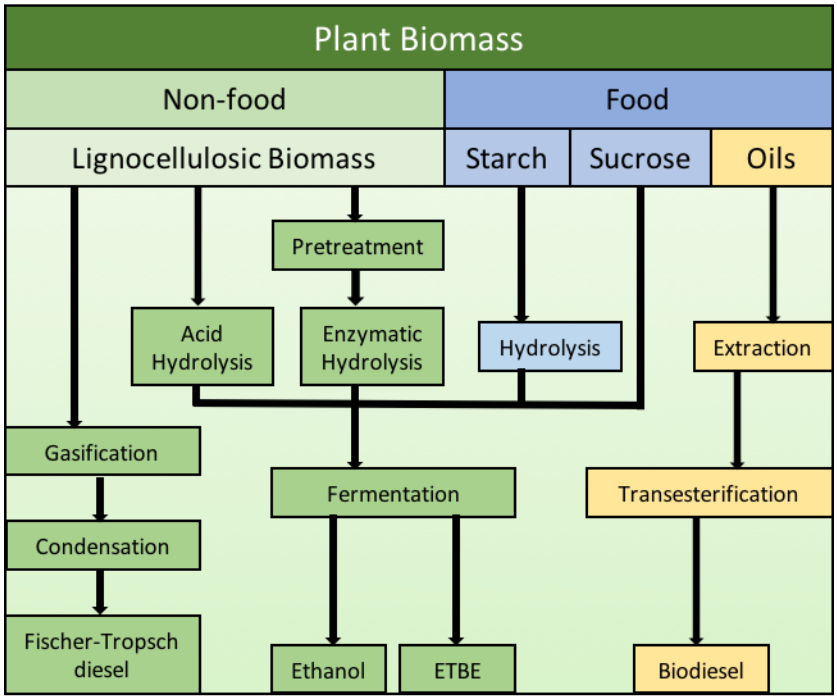
Biofuels can be obtained in a number of ways, but three of the main feedstocks for these processes come from gasification of biomass; synthesis gas, a mixture of carbon monoxide, hydrogen and carbon dioxide; and steel mill waste gas. All of these feedstocks, when utilised as the products of pyrolysis or

gasification, will consist of carbon monoxide and hydrogen as their major components, with some carbon dioxide also present, which can in some cases be utilised as part of the conversion process.

1.2.1 Plant biomass feedstocks

Biofuel production is one example of a key area where we can replace fossil fuels with a more sustainable resource with a highly renewable and sustainable potential. Advantages to biofuels are their more sustainable nature due to being derived from existing feedstocks or waste materials. There are two main classes of biofuel, which are known as first and second-generation biofuels. In relation to plant biomass, first generation biofuels are those which directly derive from food, whereas second generation fuels derive from non-food material such as lignocellulosic materials.[10] Due to the current situation with intermittent food shortages and sparsity of arable land in various parts of the world, widespread use of first generation biofuels would be of concern, as they take crops away from the food chain and divert them to use in energy generation.

Research interest has shifted to second generation biofuels, which utilise non-feedstock material such as lignin and other non-edible plant materials as a biomass energy source.[10]



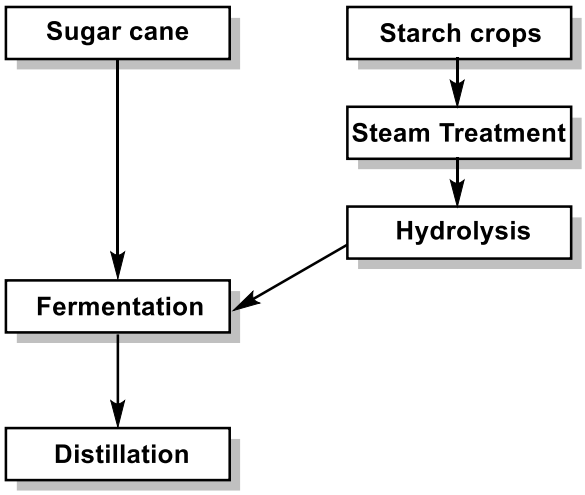
Scheme 2 – A depiction of a number of the current routes to first and second-generation biofuels utilizing food and non-food biomass.

As seen in **Scheme 2**, biodiesel and bioethanol are key biofuels of interest as part of biofuel production. Biodiesel is diesel derived from transesterification of fatty acid esters, formed by combination of animal and vegetable fats with short-chain alcohols.[11] These reactions require catalysis, usually utilising a metal catalyst. One process used to generate diesel using a metal catalyst is the Fischer-Tropsch process, which uses cobalt/iron catalysis to form hydrocarbons from gasified biomass.[12, 13]

In addition to concerns regarding food shortages, there is also an environmental issue regarding use of NPK fertilisers on crops, as this could lead to an increase in release of other gases into the atmosphere if the feedstock is combusted, in a process such as that to generate Fischer-Tropsch diesel, shown in **Scheme 2**. These gases include nitrous oxide, which would lead to an increased contribution to global warming. This is because the combustion process used to generate fuels from the crops themselves is designed to use carbon monoxide and/or carbon dioxide, where other gases are removed in a clean-up step and not utilised in the process (not shown).[14-16]

Bioethanol can be derived from processes similar to biodiesel, as seen in **Scheme 2**. The crops usually used for bioethanol production tend to be derived from fermentation of sugars. Examples of these would be sugar cane, sugar beet or starch-containing crops.[10] These (with the exception of starch) can be utilised by microorganisms without the need for breakdown or hydrolysis, which saves energy costs associated with other methods. The excess cell biomass that is generated from growth of these bacteria can also be used as an animal

feedstock, giving this process a two-fold advantage.[17] For routes utilising starch, a process similar to that of large scale alcoholic spirit consumption is employed. The starch is made into a mash by addition of water, followed by grinding, which forms a slurry. This is then cooked and hydrolysed enzymatically, followed by addition of yeast for fermentation as depicted in **Scheme 3**.[10]



Scheme 3 – Bioethanol production from traditional sugar-containing crops.

The major disadvantage to this process is that it is producing a first-generation biofuel and is therefore removing food crops, meaning that there are fewer available as feed for animal life.[10]

Another option is fuel production from waste gas, which can utilise gasified waste to produce biofuels *via* either microbial fermentation or metal catalysis.[13, 18] This is a useful alternative

to first-generation biofuels, as these can also be used in conjunction with second-generation biofuel feedstocks and methods, such as gasification and fermentation. The products that are generated can then be either sold as they are when formed, or further reacted to give more commercially viable products.

1.2.2 Syngas

In addition to the aforementioned feedstock of plant biomass, biofuels can also be obtained from syngas or steel mill waste gas. These processes do not require gasification of biomass as a second plant can be built and connected to the first, which can utilise the waste carbon monoxide and/or carbon dioxide produced as by-products. This means that both carbon monoxide and carbon dioxide can be ideal feedstocks for these processes, as the advantages are two-fold, as gases which contribute to the greenhouse effect are converted into usable chemical products as well as utilising a route to fuels and compounds which do not impact on the plant biomass used in the food industry.

As mentioned previously, there are two major routes to fuel production which utilise syngas, a mixture of single-carbon waste gases and hydrogen. These gases, when mixed, can be

converted to fuels and other products which have mass market appeal. One route to these high value products is the use of metal catalysis, which traditionally uses a copper/zinc/alumina alloyed surface to catalyse product turnover. The conditions of these reactions require elevated temperatures (150 – 300 °C) and pressures (1 – 10s atm), which are both expensive and themselves require a source of fuel to maintain, which may have originally been derived from a non-sustainable feedstock such as fossil fuel.[13]

Metal catalysis

$$(2n + 1)\ H_2 + n\ CO \rightarrow C_nH_{2n+2} + n\ H_2O$$

Figure 1 – The Fischer-Tropsch equation, which produces alkanes as its major products, however alkenes and alcohols can also be obtained.

One of the major current processes which uses metal catalysis is Fischer-Tropsch biodiesel generation the equation for which can be seen in **Figure 1**. The Fischer-Tropsch process is catalysed by an iron or cobalt-based catalyst and is used to generate liquid hydrocarbons from syngas by thermochemical conversion.[13] This process can utilize any waste that can be gasified to form carbon monoxide and carbon dioxide. The advantage of this approach is the flexibility in the feedstock, in that any material that can be gasified can be used.[19] In an industrial process, biomass and other waste is gasified *via*

heating to approximately 700 °C. The waste is a mixture of carbon monoxide, hydrogen, carbon dioxide, nitrogen and methane, which, depending on the catalyst, creates a number of useful products.[20] The products can then be separated by distillation in a method similar to that used in crude oil distillation, yielding lighter fuels as well as tars and other products with higher boiling points.

One of the disadvantages of using metal catalysts is the expense of the elevated temperatures and pressures used for the process. Additionally, the metal catalysts used are often relatively expensive compared to other methods of making biofuels. One other disadvantage is the high purity of the syngas required to prevent catalyst poisoning, often requiring multiple gas cleaning steps before entering the Fischer-Tropsch process itself.[13]

Microbial fermentation

The other route of interest to these products from syngas is the use of microbial fermentation in bioreactors to generate solventogenic and acetogenic products from syngas utilising single carbon feedstocks.[21] This does not require an elevated temperature and pressure, which is advantageous for saving

both cost and energy, however the bacteria used are generally obligate anaerobes, which only grow in anoxic environments, therefore require reaction vessels devoid of oxygen.[22] Fermentation of these acetogenic bacteria has so far been shown to yield ethanol **8** and acetate **1** as major products, but there is scope to develop them to produce other commodities such as 2,3-butanediol **7**, as well as the potential for a range of other fuels and high value chemical compounds.[18, 21, 23]

The advantage of the use of syngas for these bacteria is the substantial availability of feedstock to produce a large volume of an environmentally sustainable fuel if an optimised process is developed. Syngas fermentation also offers a number of process advantages over other techniques in a similar area, as it has tremendous flexibility in feedstock as well as high selectivity, making it a robust process with high development potential, which is depicted in **Scheme 4**.[24,25]



Scheme 4 – A simplified process of biofuel production from feedstock, such as syngas, agricultural biomass or steel mill waste gas.

Carbon monoxide fermentation by bacteria would be an ideal way of both fixing carbon monoxide in the atmosphere, as well

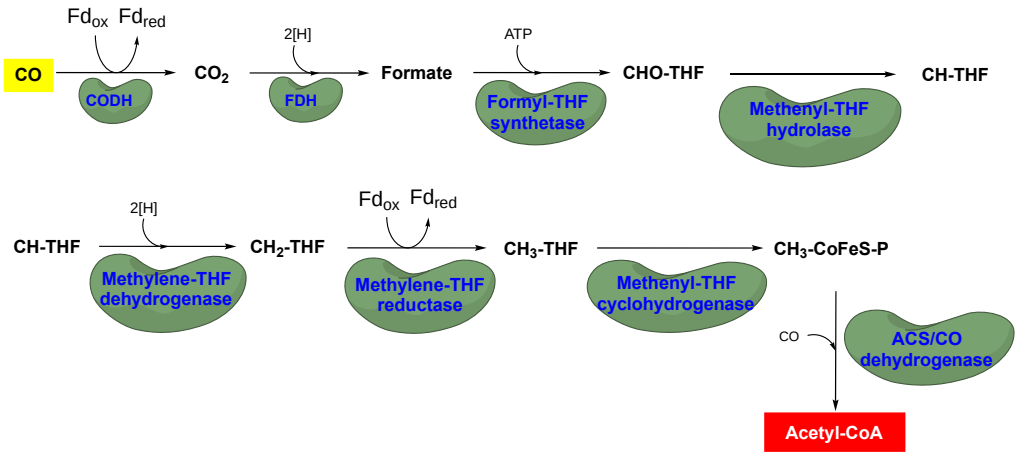
as developing a process to generate biofuels. Previous work at LanzaTech by Köpke *et al.* has determined a suitable process which, with development, may utilise bacteria which can do this. However, the optimisation of this process has not been completed and the results obtained as part of this work merely demonstrated the possibilities of the process.[21]

As part of microbial syngas fermentation, carbon monoxide is fermented by specific bacteria that do not find the gas toxic in the same way as the majority of organisms do. A group of these bacteria of interest, known as acetogens, can generate useful biofuels such as ethanol or butanol that can be metabolised from carbon monoxide and hydrogen *via* catabolic processes.[23]

1.3 Acetogenic bacteria

Acetogenic bacteria, by definition, are those who produce acetate **1** *via* the reductive acetyl-CoA pathway. However, this has been expanded to cover bacteria that can utilise carbon monoxide, carbon dioxide and hydrogen as feedstocks, in pathways such as the Wood-Ljungdahl pathway, depicted in **Scheme 5**. Acetogenic bacteria can metabolise different feedstocks, including sugars making them advantageous as

laboratory candidates for batch growth, and they form different products dependent on their growth media.[26]



Scheme 5 – The Wood-Ljungdahl pathway, showing the route which carbon monoxide is metabolized to form acetyl-CoA. CODH – carbon monoxide dehydrogenase; FDH – Formate dehydrogenase; THF – tetrahydrofolate; ACS – acetyl-CoA synthase; Fd – Ferredoxin.

Syngas agricultural biomass and steel mill waste gas all share the common characteristic of being feedstocks for organisms undergoing processes in the Wood-Ljungdahl pathway. Syngas which does not contain carbon dioxide can still be utilised by acetogenic bacteria which follow the Wood-Ljungdahl pathway. This is seen in the first step of the reductive Wood-Ljungdahl pathway (**Scheme 5**) as the organism generates carbon dioxide from carbon monoxide *via* carbon monoxide dehydrogenase (CODH) as CODH is a reversible reduction enzyme.

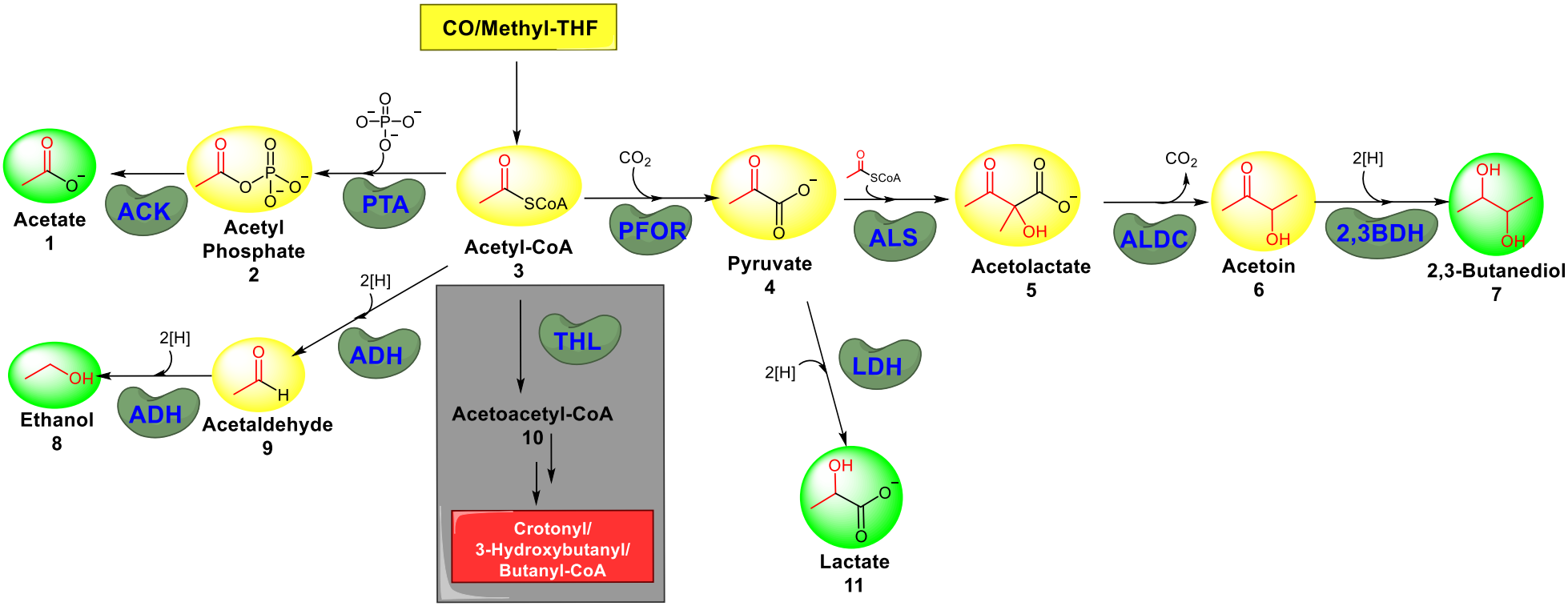
The acetogenic bacterium of interest in this project is *Clostridium autoethanogenum*, an anaerobic, Gram-positive

bacteria, which was first isolated in 1994 from rabbit faeces.[22] The reductive acetyl-CoA pathway, a simplified version of which can be seen in **Scheme 6** shows the major products produced by *C. autoethanogenum via* its natural metabolic pathways,[21] as well as the potential thiolase pathway. *C. autoethanogenum* is one of a number of acetogenic members of the *Clostridium* genus.

It has been shown previously in work performed by LanzaTech that *C. autoethanogenum*, along with its sister organisms *C. ljungdahlii* and *C. ragsdalei* can be used in syngas fermentation to generate biofuels.[21] In this study by Köpke *et al.*,[21] it was shown that four of the products of synthesis gas fermentation in these bacteria were ethanol **8**, lactate **11**, acetate **1** and 2,3-butanediol **7**. Its natural metabolic pathways convert carbon monoxide and carbon dioxide to acetate **1** and ethanol **8** in concentrations of approximately 30 mM and 20 mM respectively.[21]

However, it is known that there is the possibility for Clostridia to produce other commercially useful products, one example of which is 2,3-butanediol **7**. A simplified version of the reductve acetyl-CoA pathway which is present in *C.*

autoethanogenum, containing the major products and their metabolite intermediates, can be seen in **Scheme 6**.



Scheme 6 – Generation of major fermentation products in C. autoethanogenum showing the process of metabolism from one-carbon building blocks. The acetyl group is denoted in red to show its path. The labelled compounds of interest in the metabolic network are highlighted in yellow, with the major products highlighted in green. The prospective thiolase pathway is boxed in grey. ACK – acetate kinase; PTA – phosphotransacetylase; PFOR – pyruvate ferredoxin oxidoreducatse; ALS – acetolactate synthase; ALDC – acetolactate decarboxylase; 2,3BDH – 2,3-butanediol dehydrogenase; ADH – alcohol dehydrogenase; THL – thiolase; LDH – lactate dehydrogenase.

1.4 *Clostridium autoethanogenum*

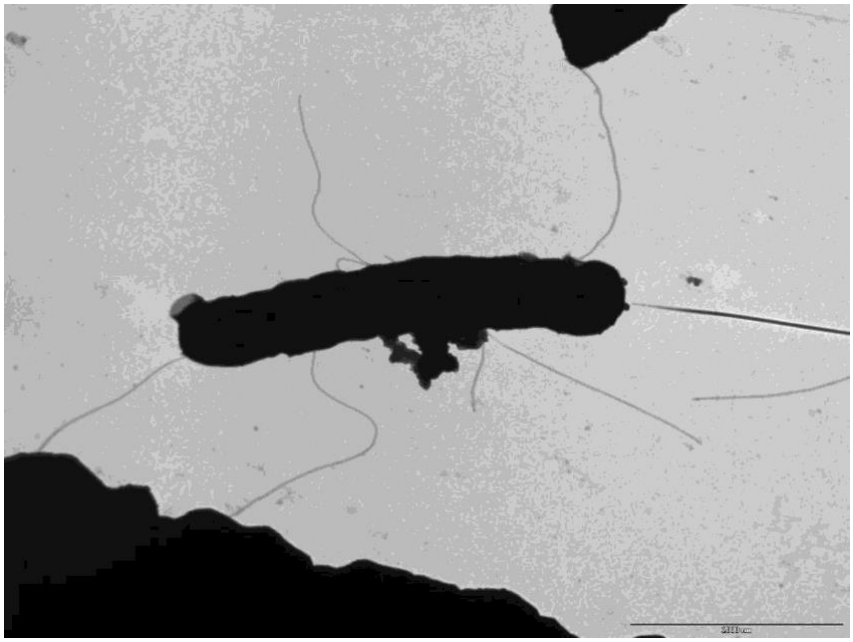


Figure 2 – Clostridium autoethanogenum, viewed under a Transmission Electron Microscope (TEM). Image taken by Synthetic Biology Research Centre, Nottingham.

In a study by Abrini *et al.*, the *C. autoethanogenum* bacterium, pictured in **Figure 2**, was isolated from rabbit faeces. Carbon monoxide was used as the source of energy for bacterial growth, which was then metabolised into ethanol, acetate and carbon dioxide, suggesting that the bacterium was acetogenic.[22] Before this, it was known that acetate synthesis occurred in the rabbit caecum,[27] a peritoneal pouch at the start of the large intestine, and the work by Abrini *et al.* confirmed this.[22] Other carbon sources that were utilised in this study leading to formation of the products of these pathways were carbon dioxide together with hydrogen,

pyruvate **4**, xylose, arabinose, fructose, rhamnose, and L-glutamate.[22]

Although genomic and physiological data for *C. autoethanogenum* is available and known, including the enzymatic activity that it has, very little is known about the specific structures of the enzymes themselves.[28, 29] Sequence data for the enzymes has been obtained, but the information on the active sites and activity of the enzymes themselves has only been determined by similarity to other proteins of the same family. Therefore, the way in which the enzymes interact with their substrate is unconfirmed. However, it is known that *C. autoethanogenum* and *C. ljungdahlii* behave in highly similar ways and are genomically similar.[21,22, 30] Analysis of the ribosomal DNA of the two strains has led to determination of the two sequences being phylogenetically indistinguishable.[31] This means that experiments conducted on one of these bacteria would be expected to give highly similar results in the other.

As mentioned previously, acetogenic bacteria, such as clostridia, produce acetate **1** and ethanol **8** as their major products. However, in the case of *C. autoethanogenum*, its

production of 2,3-butanediol **7**, is of interest because of its commercial value. In addition, through metabolic engineering, it may be possible to develop this organism to produce other compounds which have high commercial value. One example of this would be through the work of Liew *et al.* in which the organism was engineered to yield greater ethanol **8** as a major product.[32]

1.4.1 *C. autoethanogenum* metabolic products

Acetate, ethanol and 2,3-butanediol

As part of previous work by LanzaTech,[18, 21, 23] three strains of bacteria were grown in the presence of steel mill waste gas, utilising it as a feedstock, and the biomass that they produced was measured, as well as measuring the concentrations of four metabolic products. Two of these products are the major products of bacterial growth, ethanol **8** and acetate **1**; the other two products of interest were 2,3-butanediol **7** and lactate **11**.

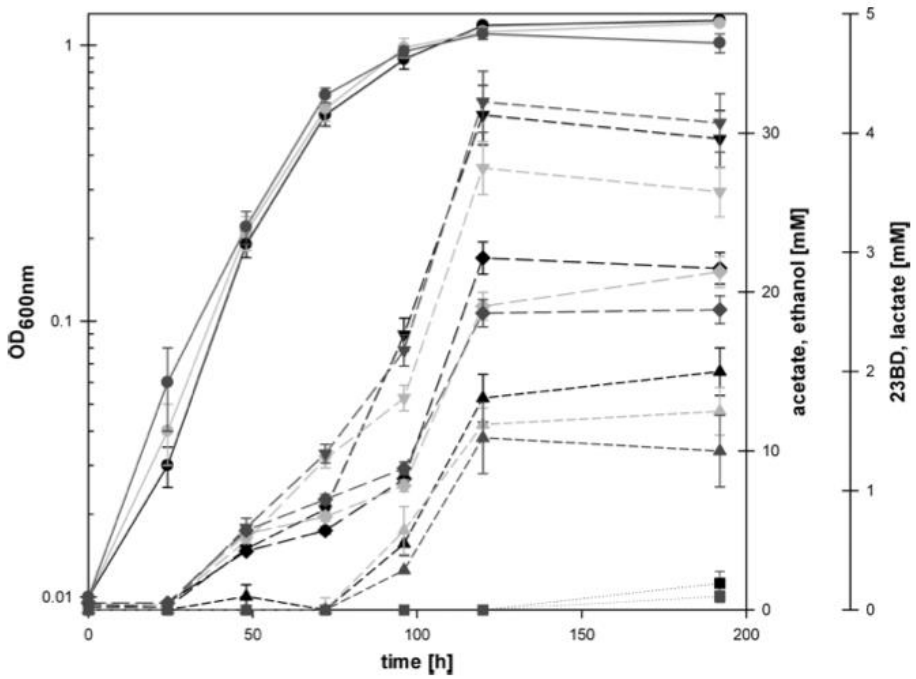
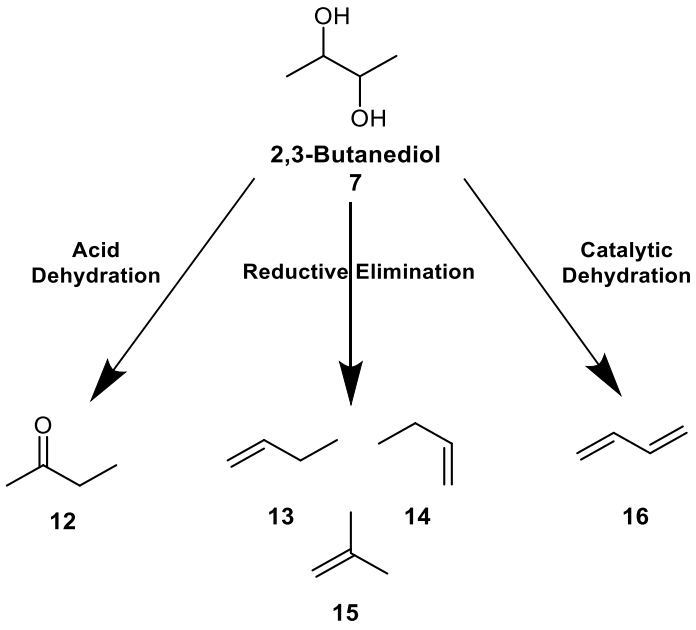


Figure 3 – Growth experiment from the Köpke et al. study showing the concentrations of solvents and acids made by C. autoethanogenum (black), C. ljungdahlii (light grey) and C. ragsdalei (dark grey). Circles, biomass; triangles down, acetate; triangles up, 2,3-BD; diamonds, ethanol; squares, lactate.21

As seen in **Figure 3**, the four metabolites being collected from the metabolic pathways in the bacteria were 2,3-butanediol **7**, ethanol **8**, acetate **1** and lactate **11**.[21] In *C. autoethanogenum*, it was observed that acetate **1** was the most abundant product from this process, with a concentration of approximately 30 mM, followed by ethanol **8** (~20 mM), 2,3-butanediol **7** (~2 mM) and a small concentration of lactate **11** (~0.3 mM).[21] Manipulation of the biosynthetic pathways in the bacteria could lead to a reduction of unwanted metabolites, such as acetate **1** and drive production of compounds such as 2,3-butanediol **7** or ethanol **8**.[32]

2,3-Butanediol **7** is of great commercial interest because it can generate large amounts of revenue, due to its versatility in making other compounds, namely butan-2-one **12**, 1-butylene **13**, 2-butylene **14**, *iso*-butylene **15** and 1,3-butadiene **16** (**Scheme 7**). Butanone **12** is commonly found in household products such as varnishes, lacquers and glues due to its solventogenic nature. Butylenes **13**, **14** and **15** are used in the production of synthetic rubber as well as the petrochemical industry. 1,3-Butadiene **16** is also used in the production of synthetic rubber. Ethanol **8** is also highly valuable due to its versatility, whether in alcoholic drinks, or a possible fuel source in vehicles; either as a sole fuel source *via* hydrogen fuel cells or in conjunction with diesel or petroleum fuels.[33,34]



Scheme 7 – The commercially viable downstream products of 2,3-butanediol 7, which can be synthesised via three simple pathways.

Using genomic information from Brown *et al.* and Humphreys *et al.* that is known about *C. autoethanogenum*, work is being conducted to attempt to knock out genes that play a role in the reductive CoA pathway in the Centre for Biomolecular Sciences, as part of a larger GASCHEM project, in collaboration with LanzaTech.[28, 29]

Other potential products in *C. autoethanogenum*

Due to previous work in clostridia by LanzaTech & others who have worked on increasing the scope of products to incorporate other high value chemicals and fuels, it has been suggested that *C. autoethanogenum* can be developed to

produce other metabolites by genetic modification.[18, 21, 23, 35] These products are summarised in **Figure 4**.

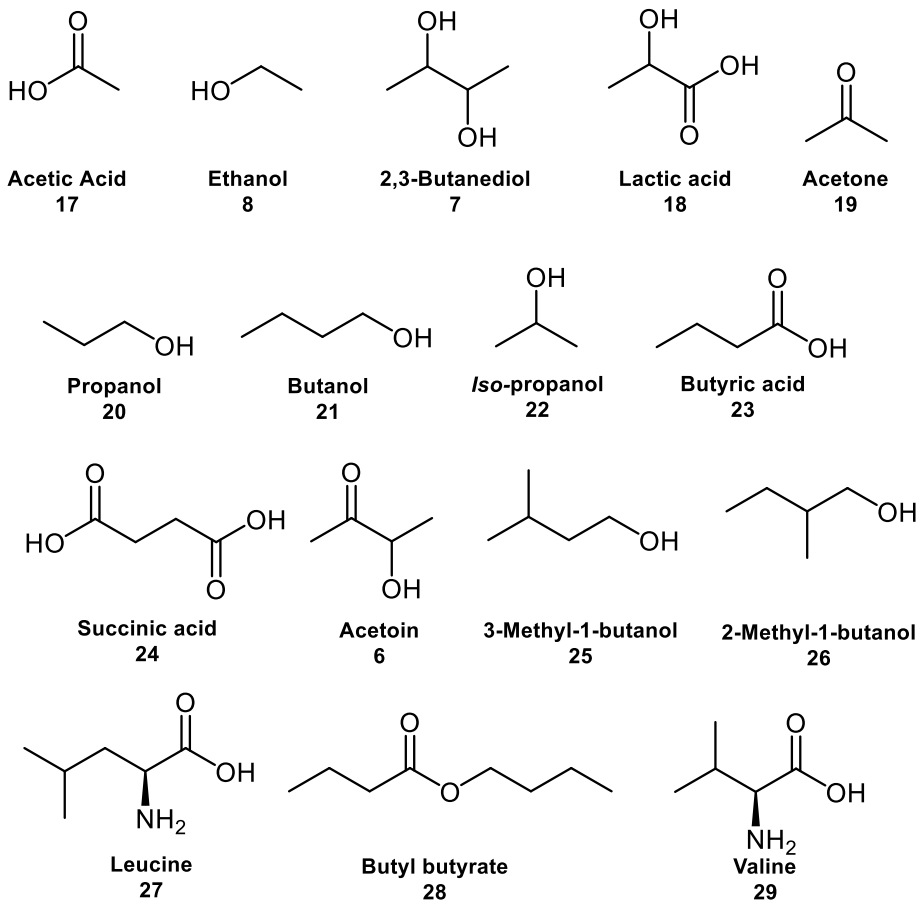


Figure 4 – A number of potential products from C. autoethanogenum based upon knowledge of its metabolic pathways.18, 21, 23, 35

In addition to those previously mentioned, other products of the metabolic network of *C. autoethanogenum* are wide-ranging and some of these can be seen in **Figure 4**. Of these, succinic acid is of particular interest because of its use in polyester and as a food additive. Within the Synthetic Biology Research Centre, Nottingham, investigations are being undertaken into genetically engineered pathways to drive

production of these high value compounds. Other compounds of interest are acetone **19**, butan-1-ol **21** and propanol **20** as these are used in many chemical processes as bench chemicals.

As discussed earlier, acetate **1** is one of the two major products produced in acetogenic bacterial growth, and is generated as a waste product of their anaerobic processes. However, the more desirable products of the reductive CoA pathway are solventogenic, such as ethanol **8** or 2,3-butanediol **7**. There are several ways to affect product formation in microbial fermentation, several of which have been investigated, including feedstock and pH variation and numerous genetic knockouts. Additionally, enzyme inhibition is of interest as it can offer a more temporary decrease in enzyme activity, which may be useful for affecting pathway flux.

1.5 Metabolic engineering in *C. autoethanogenum*

1.5.1 Feedstock and pH variability

One variable to consider for these pathways is feedstock composition and pH, as this could have a great effect upon the

bacterial culture. The feedstock of the bacteria can affect the solvent production in *Clostridia* species, as different feedstocks can allow acetogens to metabolise different products.

Previously, Girbal *et al.* investigated the way in which feedstock and pH variation affected acidogenesis and solventogenesis, as well as the flux between the two.[36] This suggests that different conditions are also important for the metabolic pathways involved in *C. acetobutylicum*, which may also be important in *C. autoethanogenum*. For example, when glycerol was given in addition to glucose, the solvent production increased and acid production decreased, suggesting glycerol as a better feedstock than purely glucose **132**.

Further work by Cotter *et al.* investigated the different growth and metabolic pathway variation in *C. ljungdahlii* and *C. autoethanogenum,* suggesting that ethanol **8** production correlates with growth and increases in conjunction with increased cell density. This lead to a plateau when the cells begin slow growth, whereas acetate **1**, which appears to be a principle metabolite, continues to increase despite the rate of growth decreasing.[37] This suggests that ethanol **8** is a growth-

associated metabolite and acetate **1** is a non-growth-associated metabolite in these bacteria. This study also showed that the overall growth of the bacteria is lower at a lower pH, with a reduction in ethanol **8** production. However, acetate **1** production was only slightly decreased at a lower pH.

Conversely, Girbal *et al.* showed that overall biomass growth and solvent production were increased at lower pH levels, suggesting that growth is required for solvent product formation.[36] It is worth noting that the differences observed here could be caused by feedstock in addition to the pH shift; Cotter *et al.* grew the two strains of Clostridia using carbon monoxide whereas Girbal and co-workers used sugars as a feedstock. This suggests a potential disparity between results in the two feedstocks, which may require process optimisation if correct.[36, 37] Another point to note is that these studies were carried out in different organisms, so comparisons between the two must be drawn cautiously.

1.5.2 Gas stripping

A different type of conducted study by Ezeji *et al.* utilised gas stripping of solvents in *C. beijerinckii*, to ensure that the fermentation process had gone to "completion", where only

solvent was produced at the time where the experiment was ceased.[38] Gas stripping involves removal of the products from the reaction vessel as they are generated. This meant that all of the products that were formed at completion of the experiment were acetone **19**, butan-1-ol **21** or ethanol **8**. This suggests that gas stripping is a method to obtain complete conversion of starting materials to feedstocks. However, the drawback in this method is that its sole suitability is batch fermentation, where a set amount of feedstock was used to grow the bacterial culture and converted to solvent; whereas for industrial applications, continuous fermentation would be preferred, as this provides a constant stream of products from the bacterial culture.

1.5.3 Genetic knockouts

Investigations by Cooksley *et al.* undertook investigations into other genetic knockouts in *C. acetobutylicum* and explored the way in which acid and solvent production was affected. Of interest in this study was the knockout of the *ack* gene, which encodes production of the acetate kinase (ACK) enzyme of the reductive acetyl-CoA pathway. Acetate kinase is the second enzyme in the pathway from acetyl-CoA **3** to acetate **1** (**Scheme 6**) and is of interest for removal in *Clostridia* to

attempt promotion of the other pathways in this metabolic cycle.[39]

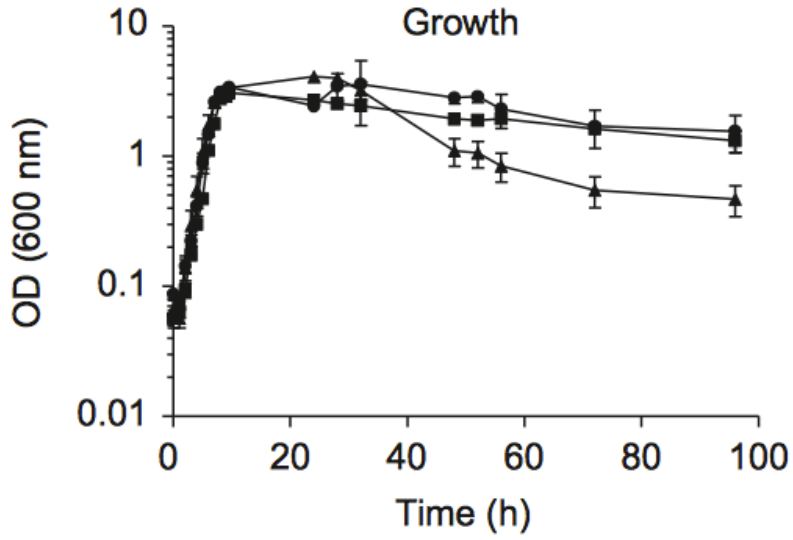


Figure 5 – A graph showing optical density plotted against time, which shows the growth of the mutants in the study by Cooksley et al.[39] Black triangles denote the ack genetic knockout mutant.

The results in **Figure 5** show that deletion of the *ack* gene initially has little effect on growth as the OD_{600} of the mutant increases in line with the wild-type, however, the OD_{600} decreases significantly after ~30 h. This suggests that the *ack* knockout mutation is harmful to *C. autoethanogenum* as initial growth was seen before observing that the bacterium could not survive without the *ack* gene. However, these samples were grown on glucose, which was used more quickly in the *ack* mutant, suggesting that the feedstock was used purely as an initial energy source. Following consumption of the glucose,

the organism struggles to re-generate ATP **131**, causing a decline in the growth, and ultimately leading to bacterial culture death. In addition, unlike *C. autoethanogenum*, *C. acetobutylicum* contains a pathway to produce butyrate and butanol **21**. The native metabolic pathway to butyrate is similar to that of acetate **1** in that butyrylphosphate is an intermediate in the pathway that can regenerate ATP **131**.[39] There may be a difference of results in an organism that does not contain this pathway as a major route to products.

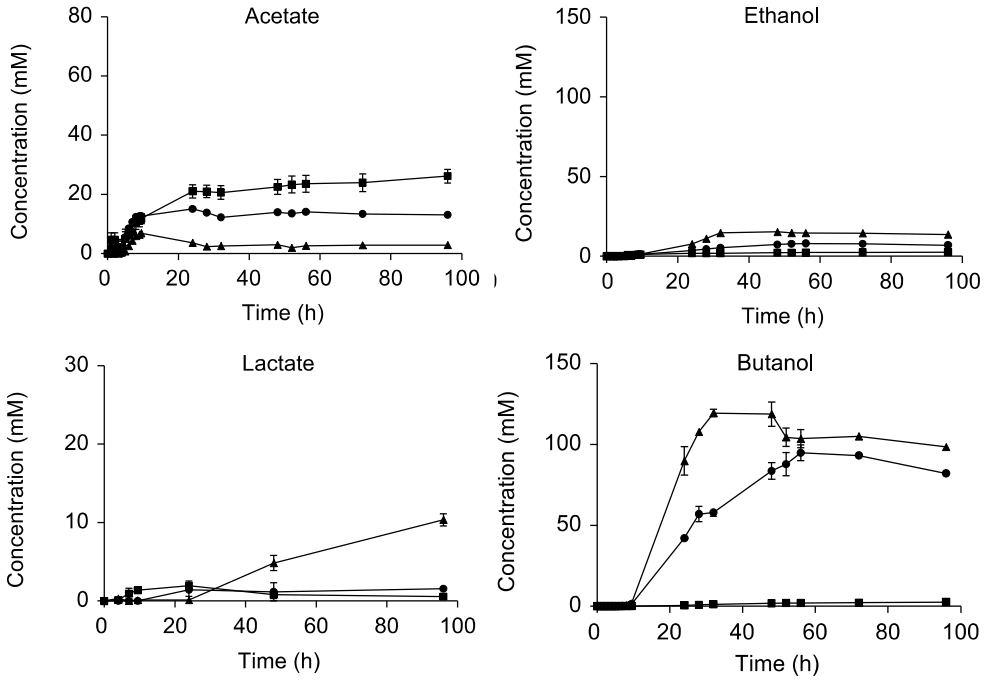


*Figure 6 – The production of acetate **1**, ethanol **8**, lactate **11** and butan-1-ol **21** in the ack genetic knockout mutant in the study by Cooksley et al.[39] Black triangles denote the ack genetic knockout mutant.*

From the graphs in **Figure 6** taken from Cooksley *et al.*, it can be seen that the *ack* mutant also has a different product output than the wild-type and other genetic knockout mutants

in the study.[39] It can be seen that the amount of acetate **1** at the end of the timecourse is close to 0 mM, while outputs of butan-1-ol **21** and ethanol **8** have also had small increases. This suggests a shift towards solventogenic products. However, as discussed, following initial growth, the cultures OD_{600} dropped. This correlates with an increase in lactate **11** concentration, which forms from pyruvate **4** as seen in **Scheme 6**. It can be concluded that the genetic knockout of *ack*, and its subsequent lack of an ACK enzyme leads towards increased solvent products, which is desired, but with a much less sustainable survival during stationary phase of *C. acetobutylicum*.

The results in the study by Cooksley *et al.* are supported by work by Winzer *et al.* which investigated the *ack* gene and acetate kinase enzyme in a study working on the transcription of the gene and the translation of the enzyme in acidogenic and solventogenic conditions.[40] The results of this study showed that both processes occurred in each condition, which supports the above result, where solvent production was not driven by deletion of the gene, rather survival rate of the cells and overall metabolite production decreased.

The enzymes which are affected by the deletions of the *pta, ack* and *adh* genes are of interest to the reductive acetyl-CoA pathway, as they are the genes involved in the acetate **1** and ethanol **8** branches of the pathways. Additionally, the enzymes that they encode, phosphotransacetylase (PTA), acetate kinase (ACK) and alcohol dehydrogenase (ADH) could also be affected by enzyme inhibition. Their substrates, acetyl-CoA **3**, acetyl phosphate **2** and acetaldehyde **9** are intermediates in the reductive acetyl-CoA pathway, therefore if these enzymes were not available in the host cell, the metabolic pathway would be switched off.

1.6 Isotopic labelling and metabolic flux analysis

Isotopic labelling is a technique used to trace the path that a labelled compound takes through a metabolic pathway or reaction. Labelled compounds contain one or more atoms that have been isotopically labelled, meaning that they contain an atom which does not have the atomic mass of the most commonly abundant isotope. For example, ^{2}H, ^{13}C or ^{15}N are isotopic labels of hydrogen, carbon and nitrogen respectively. Isotopic labels can either be traced by NMR, if they are spin active, or mass spectrometry, either by GC-MS or LC-MS.

Isotopic labelling can be used in flux analysis of metabolic pathways. By labelling compounds using stable isotopes and feeding them to bacterial cells, the compounds can then be taken up by metabolic pathways and undergo the reactions in that pathway. They can provide information on the formation of products from reactants, which shows how the metabolite passes through both the forward and reverse reactions in the pathway, because a number of enzymes have reversible activity.[41, 42]

In addition to the study by Köpke *et al.* mentioned previously focusing on 2,3-butanediol **7** production, which utilised HPLC as an analytical method for detecting products and determining their concentration, work by Yun *et al.* investigated *C. ljungdahlii* to trace isotopically labelled carbon monoxide uptake.[18, 43] This utilised GC-MS to trace the ^{13}C-labelled carbon monoxide through the metabolic pathways in bacteria towards ethanol **8** and acetate **1** formation.

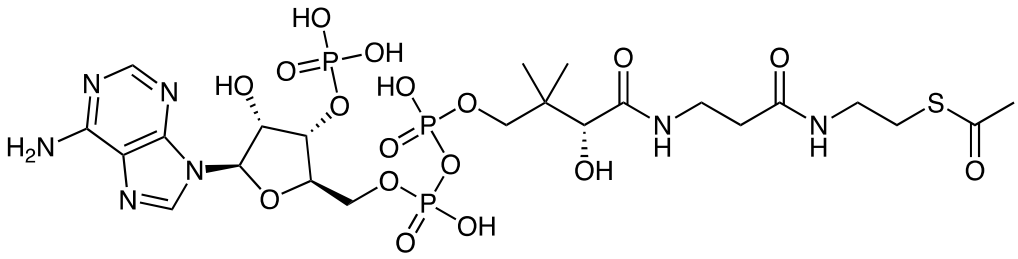
The research by Köpke *et al.* also showed that the bacteria assimilate a very low level of carbon monoxide from the air, which could be seen due to the ^{13}C-enrichment of the feedstock. This carbon monoxide is fermented by the bacteria

and forms both solvent, such as ethanol **8**, and acidic products, such as acetic acid **17**.[43]

Isotopic labels can also be observed in NMR spectroscopy, as they are usually spin-active nuclei. The most commonly used spin-active nuclei used for isotopic labelling experiments are ^{2}H, ^{13}C and ^{19}F. ^{13}C has 1 % natural abundance, therefore by labelling a molecule with a ^{13}C label, the abundance increases 100-fold, which allows metabolite tracing through metabolic networks by ^{13}C NMR, for example.

It may be useful to label the intermediates of the reductive acetyl-CoA pathway, as this would allow tracing through the enzymatic pathways. This would be particularly useful in conjunction with genetic knockouts/enzyme inhibition or strains containing additional genes to develop new products, as the flux of labelled feedstocks could be followed.

1.6.1 Acetyl-CoA



3

*Figure 7 – Acetyl-CoA **3**, which is a major branch point in the network of the metabolic pathways in C. autoethanogenum.*

Coenzyme-A is an important co-factor in all living organisms, as it forms the basis of metabolic pathways and energy networks. It is responsible for a large number of biochemical transformations, and is thought to be involved in over 100 intracellular reactions.[44, 45] Acetyl-CoA **3**, shown in **Figure 7**, is synthesised at the end of the Wood-Ljungdahl Pathway in *C. autoethanogenum*, which can be seen in **Scheme 5**.

Acetyl-CoA **3** is synthesised initially from pantothenate *via* several steps through reaction with ATP **131**, cysteine and tetrahydrofolates *via* a cascade of enzyme-catalysed reactions. Following biosynthesis in *C. autoethanogenum*, it undergoes enzymatic transformations to form several metabolites, depending upon which enzyme it interacts with. This means that acetyl-CoA **3** may be of interest for isotopic labelling

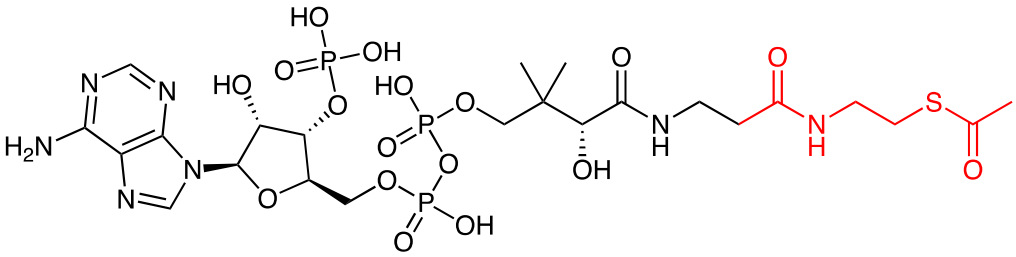
studies as it may give a platform to investigate these pathways.

SNAC thioester mimics

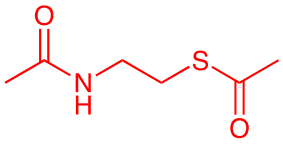
Acetyl-CoA **3** is a large polar molecule with a cLogP of -6.04, shown in **Figure 7**, and due to this polarity and its hydrophilicity is unlikely to permeate cell membranes in bacterial culture growth without the presence of an active transport system. Mimics of the parent molecule may be useful for analysis of metabolic flux using isotopic labels. One such class of molecule of interest in this area are *S,N*-acetylcysteamine (SNAC) thioesters, which mimic the tail region of acetyl-CoA **3**. Previous work by several research groups, initially pioneered by Gilbert *et al.* showed that SNAC thioesters can be utilised as part of polyketide biosynthesis, to build larger chain compounds.[46-48] This suggests that the enzymes responsible in this pathway identify the SNAC thioester as a mimic for the parent CoA for the purposes of biosynthesis and utilise them.

As mentioned previously, SNAC thioesters encompass the tail region of the CoA molecule as shown in **Figure 8**, and they

have an increased likelihood of being cell permeable versus their parent compound **3** due to their increased hydrophobicity and decreased size, which is potentially important in cell culture studies to test these compounds.



3

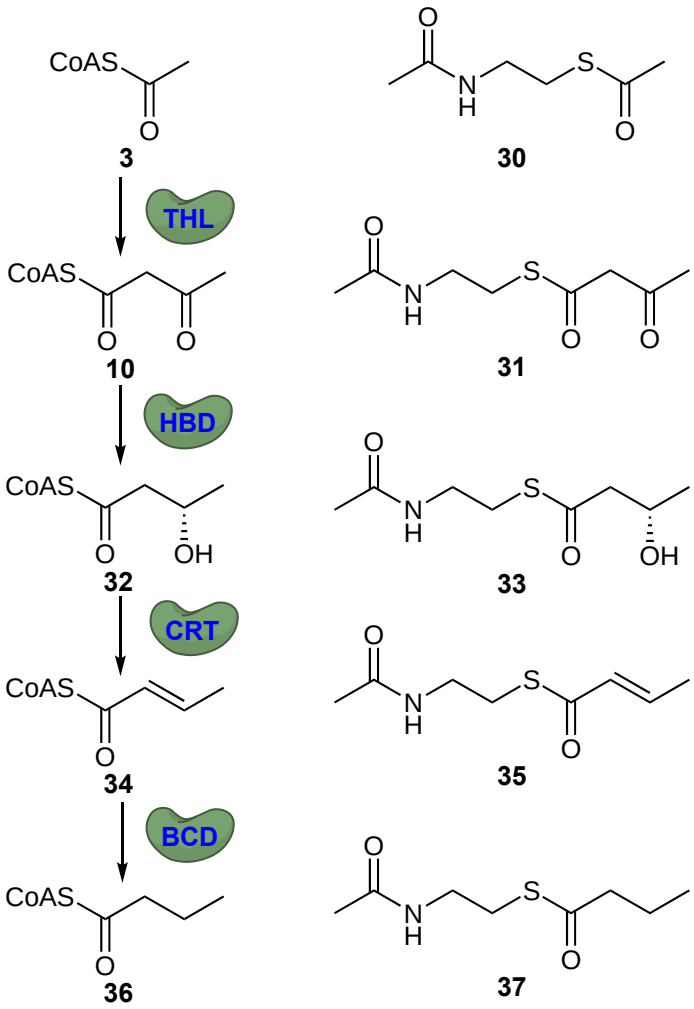


30

*Figure 8 - The parent acetyl-CoA molecule **3** compared to its potential acetyl-SNAC thioester **30** mimic.*

Hydrophobicity is a good indicator of uptake in biological systems and is one of the determining factors of Lipinski's rules, which refer to bioavailability of molecules.[49] Compounds that are too hydrophilic may be less effective than those with a marginally increased hydrophobicity. A further study by Ghose *et al.* suggests that there is a lower limit to the clogP of compounds which also affects their bioavailability. This was determined to be -0.4.[50] Therefore, compounds which are too hydrophilic may have difficulty being uptaken by cells also. This is important when looking at a molecule such as acetyl-

CoA as a target for isotopic labelling, because it is unlikely to be taken up by cells due to it polar, hydrophilic nature. The relative cLogP values for acetyl-CoA **3** and its corresponding SNAC thioester **30** are -6.04 versus -0.03, meaning that the SNAC thioester is more hydrophobic than its acetyl-CoA **3** counterpart. Additionally, they are smaller compounds which may be more easily transported across cell membranes than acetyl-CoA **3**, which is significantly larger. Therefore, isotopically labelled SNAC thioesters may be a suitable mimic to be incorporated into metabolic networks.

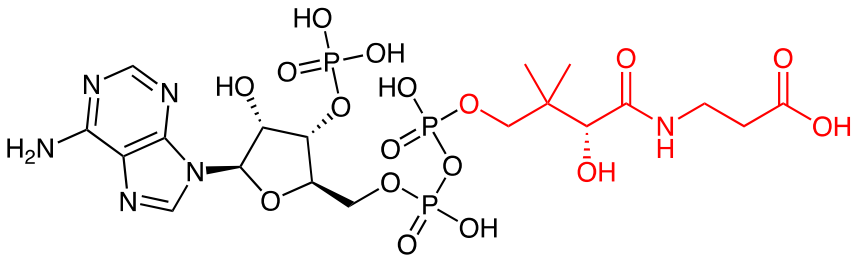


Scheme 8 – The thiolase pathway, which may be present in clostridia and their counterpart SNAC thioesters that were synthesized as part of the previous work by Jenner et al.[48] THL – thiolase; HBD – hydroxybutyryl-CoA dehydrogenase; CRT – crotonyl-CoA transferase; BCD – butyryl-CoA dehydrogenase.

Previously, a number of SNAC thioesters were synthesised by Jenner *et al*. These were acetyl-SNAC **30**, butanyl-SNAC **37**, crotonyl-SNAC **35**, acetoacetyl-SNAC **31** and *S*-3-hydroxybutanyl-SNAC **33** thioesters.[47, 48] All of these SNAC thioesters mimic different CoA molecules that are present in the reductive acetyl-CoA pathway and thiolase pathways of interest (**Scheme 8**). A number of genes encoding for the

enzymes of the thiolase pathway have potentially been identified in genome sequencing in *C. autoethanogenum*, however, thiolase has not been confirmed, therefore the pathway shown in **Scheme 8** may not be active.[28, 29] Although it is not necessarily present at the time of writing, it may be possible to genetically engineer *C. autoethanogenum* in the future to contain the pathway, increasing the impact of this work should it be successful.

Pantothenate mimic

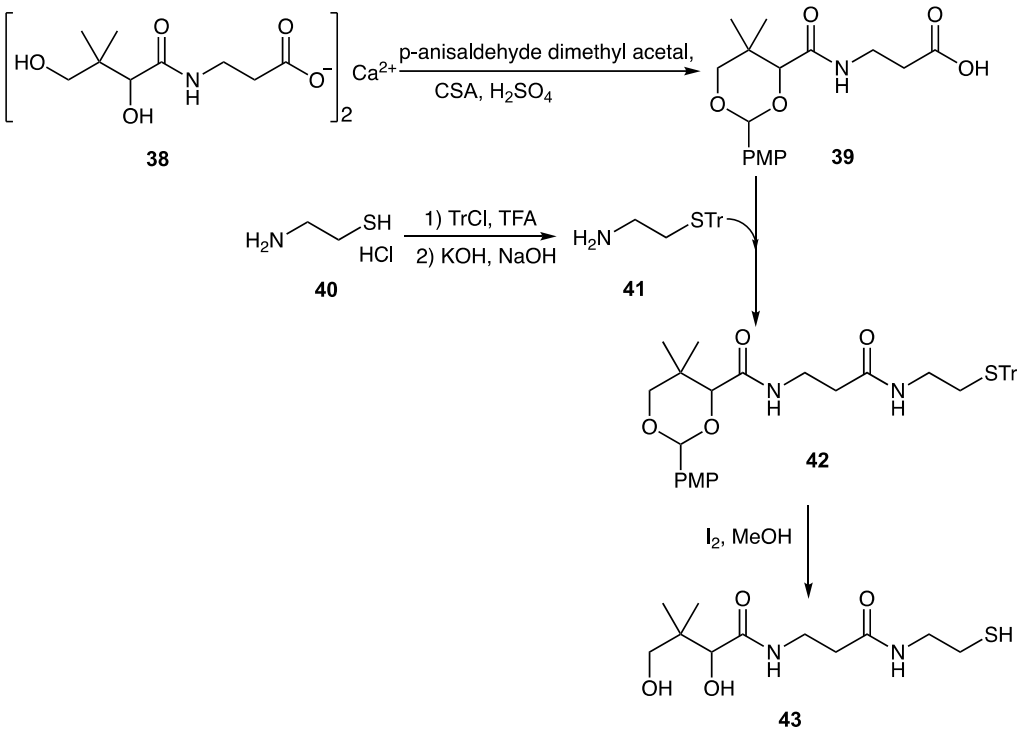


3a

*Figure 9 – CoA **3a**, with the structure of pantothenate highlighted in red.*

From work by Humphreys *et al.* and Annan *et al.* on the genome sequence and on pantothenate as a limiting factor to growth, it is known that the genes *panB, panC and panD* are not present in *C. autoethanogenum*, therefore it lacks the enzymes to be able to naturally biosynthesise pantothenate.[29, 51] This means that pantothenate must be added to the growth

media for the organism to grow, such that it can biosynthesise CoA **3a** *via* its natural biosynthetic pathways, involving enzymatic reactions with cysteine and ATP to form CoA **3a**, shown in **Figure 9**.[44, 45]



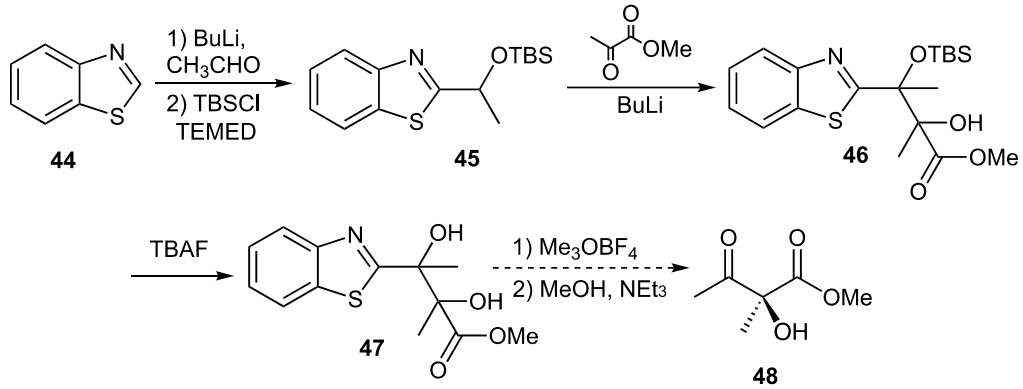
*Scheme 9 – The route followed by Mandel et al. to synthesise pantetheine **43**.*[52]

Previous work by Mandel *et al.* utilised calclum pantothenate **38**, and through a series of transformations as seen in **Scheme 9**, yielded pantetheine **43**, which following acetylation, would form the tail region of acetyl-CoA **3**. Acetyl-pantetheine could be tested as a substitute in the growth media for calcium pantothenate **38** as it may be recognised by the cell and incorporated directly into the biosynthesis of

acetyl-CoA **3** as one of its intermediates. The aim of this would be to track it through the metabolic network *via* mass spectrometry and NMR spectroscopy.

1.6.2 Metabolites of the reductive acetyl-CoA pathway

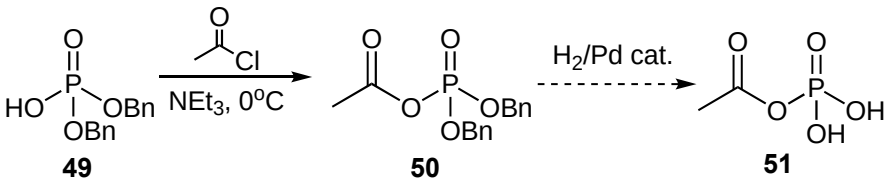
Previous work exploring downstream metabolites of acetyl-CoA has been completed previously, although there are no examples of metabolite research in *C. autoethanogenum*. Therefore, it would be of interest to examine isotopic labelling of the downstream metabolites.



*Scheme 10 – The previously explored route to methyl acetolactate **48** synthesis.*[53]

Methyl acetolactate **48** is reported to have been successfully synthesised, *via* a thiazolium-mediated route by David *et al.*, using acetaldehyde **9** to perform an aldol reaction with the thiazolium scaffold *via* an intermediate **46** with methyl pyruvate. This was then deprotected and removed from the scaffold using trimethoxytetrafluoroborate to yield the desired

product **48**. This differs from acetolactate **5** by a single methyl ester, which has the potential for removal *via* base hydrolysis.[53] This work contained no experimental detail so would require design of a synthetic route to acquire the product if it were to be repeated.



*Scheme 11 – A possible route to synthesis of acetyl dihydrogen phosphate **51**, via acetylation of dibenzyl phosphate **49**, followed by hydrogenolysis to form the desired product **51**.*

Acetyl dibenzyl phosphate **50** has previously been synthesised *via* reaction of acetyl chloride with dibenzyl phosphate **49** by Cope *et al.* in good yields,[54] which could potentially undergo a hydrogenolysis reaction using palladium on carbon under a hydrogen atmosphere to form acetyl dihydrogen phosphate **51**, the protonated form of acetate **1**.

1.7 Enzyme inhibition

The inhibition by a ligand prevents the binding of the natural substrate of the enzyme *via* competitive inhibition or by binding to an allosteric site in non-competetive inhibition on the enzyme, which can then alter the conformation of the active site, blocking substrate binding.[55] Both of these types of

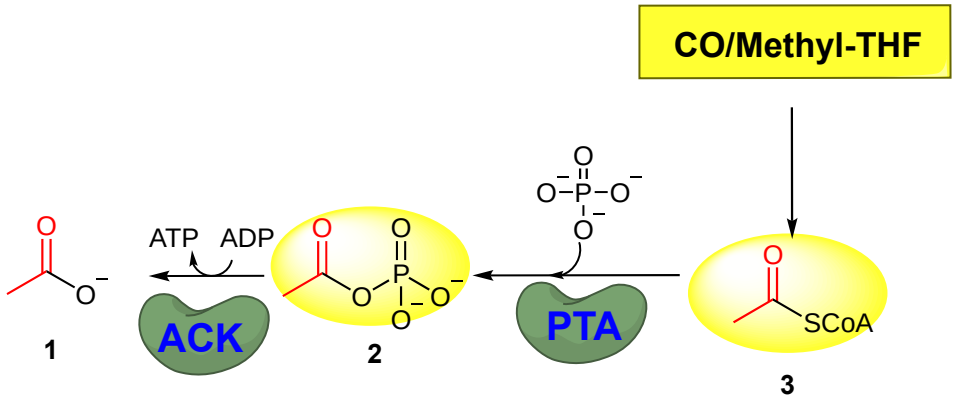
inhibition can be reversible or non-reversible. Reversible inhibition means that the substrate can overcome the inhibitor and the enzyme can turnover the substrate to product, albeit at a slower rate.[55]

Enzyme inhibition is an alternative approach for manipulation of organisms to genetic knockouts because of this reversibility. It has the potential to reduce the activity of the enzyme, rather than remove it in the way that a genetic knockout would. Of interest in *C. autoethanogenum* is the enzyme acetate kinase, which catalyses the hydrolysis of acetyl phosphate **2** to regenerate ATP, as the acetate **1** produced is a lower value byproduct than the solventogenic fuels such as ethanol **8** and 2,3-butanediol **7**.

1.7.1 Acetate kinase

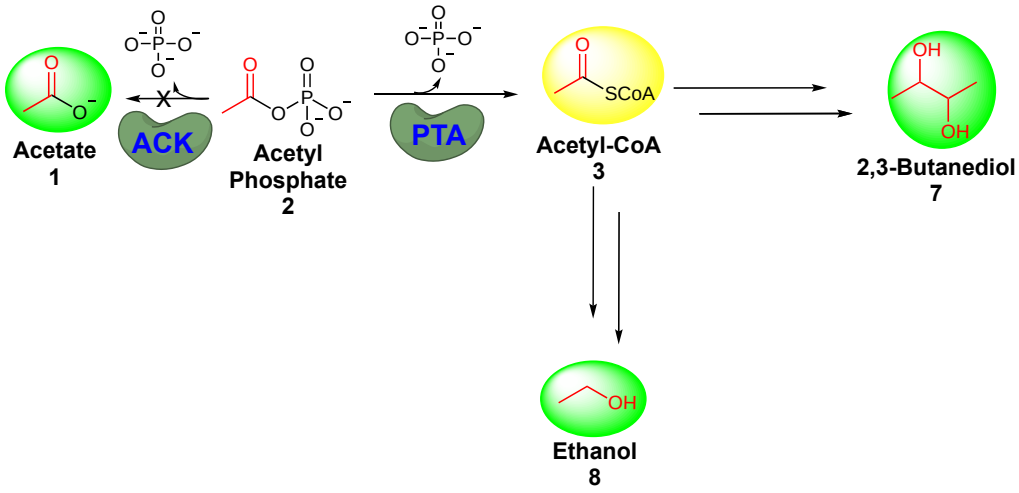
Acetate kinase (ACK), encoded by the *ack* gene, has been shown in previous research by Cooksley *et al.* that gene knockout had a dramatic reduction in survivability of *C. acetobutylicum* during culture, whilst only a mild shift to solventogenic products was seen.[39] Therefore enzyme inhibition, a transient knockout using a small molecule

inhibitor, may be a better method for regulation of ACK function as part of *C. autoethanogenum* growth.



*Scheme 12 – The native pathway of C. autoethanogenum, showing acetate **1** formation via acetyl phosphate **2** from acetyl-CoA **3**.*

In *C. autoethanogenum*, acetate kinase is responsible for the regeneration of ATP by turnover of acetyl phosphate **2** to to produce acetate **1** as a product as seen in **Scheme 12**. The organism uses this as an essential pathway for energy recovery and growth.[21, 39] This may explain the reduction in overall growth that is observed in previous work detailing genetic knockouts of the *ack* gene.[39]



*Scheme 13 – The potential result of enzyme inhibition of ACK, which may cause the flux of PTA to reverse, leading to formation of 2,3-butanediol **7** or ethanol **8**.*

Enzyme inhibition would potentially provide a more controlled regulation of acetate formation, whereby the level of inhibition is controlled by the concentration of inhibitor present which can act upon the ACK enzyme. This would mean a potential regulation of the output of acetate **1** from *C. autoethanogenum*, because if a percentage of acetate kinase is inhibited, acetyl phosphate **2** may be recycled to acetyl-CoA **3**, as shown in **Scheme 13**. This regulated flux to acetate **1** may drive product formation towards more high value compounds such as ethanol **8** or 2,3-butanediol **7**, without completely abating culture growth.

As mentioned earlier, enzyme inhibition can be transient and reversible, meaning that it can be overcome by the substrate over time in sufficient concentration.[55] Reversible inhibition

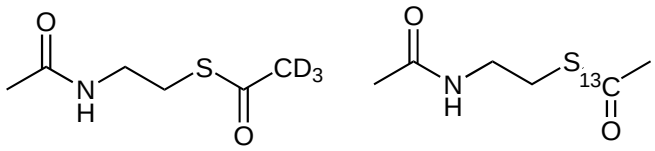
would be desired in *C. autoethanogenum*, as if complete inhibition was achieved and the inhibitor permanently outcompetes the substrate, the organism would mirror the genetic knockout of *ack*. This would lead to no energy regeneration and therefore would not grow sufficiently to drive desired product formation as found by Cooksley *et al.* in *C. acetobutylicum.*[39]

1.8 Project Aims

The overall aim of this work was to develop methods for probing and affecting the native biosynthetic pathways in *C. autoethanogenum*. This would be approached in two ways, isotopically labelled compound synthesis for metabolite tracing and inhibition of the native enzymes in an attempt to direct product formation.

As mentioned previously, to investigate the substrate flux of native pathways, isotopically labelled compounds incorporating either ^{2}H or ^{13}C atoms can be used, as these can be traced by mass spectrometry or NMR. In this case, to find products of the pathway of interest, the reductive acetyl-CoA pathway, as shown in **Scheme 6**. Synthesis of isotopically labelled compounds involved in this pathway was of particular interest.

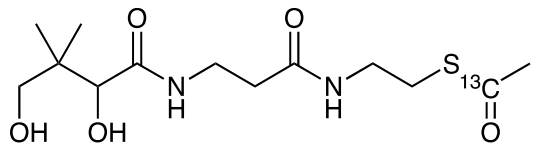
Previous work by Gilbert *et al.* and more recently Gao *et al.* and Jenner *et al.* have shown that the enzymes involved in polyketide biosynthesis can recognise SNAC thioesters as natural substrates in place of acetyl-CoA **3**.[46-48] As acetyl-CoA **3** is the branching point of the reductive acetyl-CoA pathway, it would be of interest to synthesise mimics which *C. autoethanogenum* could incorporate into its pathways to observe downstream metabolite and product formation.



52 **53**

Figure 10 – Initial SNAC thioester targets for use in whole cell growth experiments to test metabolic flux.

As shown in **Figure 10**, primary targets for synthesis for SNAC thioesters would be 2-$^{2}H_3$-acetyl-SNAC **52** and 1-^{13}C-acetyl-SNAC **53**. These could then be tested in a whole cell culture growth of *C. autoethanogenum* and the products of the growth analysed by NMR and mass spectrometry. Further work could go towards the synthesis of the other SNAC thioesters of the thiolase pathway, to test its activity in *C. autoethanogenum*.



54

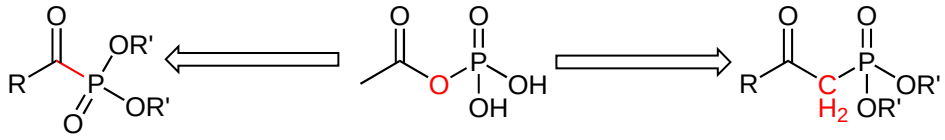
*Figure 11 – The desired mimic to replace pantothenate in whole cell culture growth experiments, in the hope that it will be incorporated into acetyl-CoA **3** biosynthesis.*

Further to more recent research regarding the pantothenate biosynthesis pathway, synthesis of a mimic based on incorporation of calcium D-pantothenate **38**, shown in **Figure 11**, may also be of interest in an attempt to incorporate a ^{2}H or ^{13}C label *via* this route.[29, 51] The mimic chosen for the work is 1-^{13}C-acetyl pantetheine **54** which may be utilised in the biosynthesis of acetyl-CoA **3**, which can then undergo whole cell culture growth and testing *via* NMR and mass spectrometry.

The second aspect of this project was cell permeable inhibitor design to drive product formation towards more high value compounds in *C. autoethanogenum*, such as the solventogenic products ethanol **8** and 2,3-butanediol **7**. To achieve this, acetate kinase was the enzyme chosen for inhibition as it is the enzyme catalysing the reaction of acetyl phosphate **2** to acetate **1**, to regenerate ATP in the organism.

In conjunction with this project investigating enzymatic inhibition, studies into knockouts of all of the genes in *C. autoethanogenum* have been undertaken as part of the GASCHEM project at the Centre for Biomolecular Sciences, University of Nottingham. It was therefore hoped that comparisons between an *ack* genetic knockout and ACK enzyme inhibition would be possible.

When designing inhibitors for an enzyme, if a crystal structure of the enzyme is available, it is possible to do *in silico* screening of compounds, to try to find a 'hit' for binding to the enzyme. However, in the case of *C. autoethanogenum*, no enzymes from the organism have been purified and crystallised, and similarity to the nearest determined structure is 63 %.[56] This similarity is to acetate kinase from *M. thermophila*, a thermophilic methanogenic archaeon.[57] Because of this lack of similarity, a compound docking approach was not taken to inhibitor design.

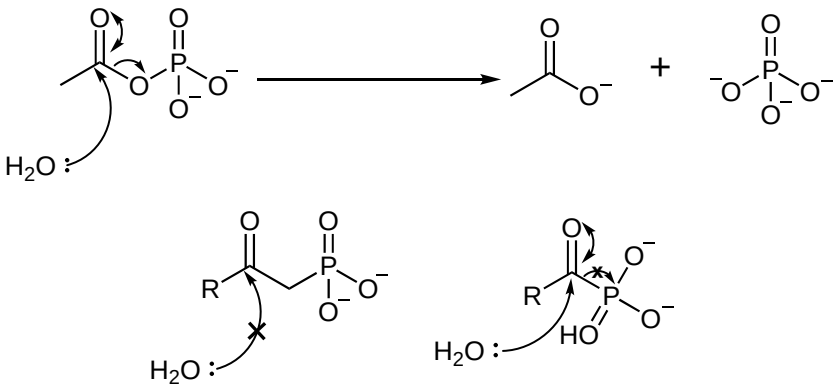


2

R' = H/Me

Scheme 14 – The two proposed classes of compound for inhibition of acetate kinase in C. autoethanogenum.

Instead, a ligand-based structure design was adopted for synthesis of the inhibitors. This involved design of the inhibitors based upon the theoretical mechanism of action of the enzyme, in an attempt to block it. In this case, inhibitors were modelled upon acetyl phosphate **2**, as seen in **Scheme 14**.



*Scheme 15 – The hydrolytic cleavage of acetyl phosphate **2** to acetate **1** and free phosphate, and its blockage by the two potential classes of inhibitor.*

Two families of compound were identified as a starting point to build libraries of compounds. The first class of compounds, α-carbonylphosphonates differ from acetyl phosphate by removal of the oxygen atom linking the carbonyl and phosphate

moieties, whereas the second utilises an isostere of the oxygen phosphate bond being incorporated, a methylene, to give β-ketophosphonate inhibitors. Both of these classes of compound remove the hydrolytically labile C-O bond which can normally be broken to regenerate ATP *via* phosphate. The two classes of inhibitor compounds can be seen in **Scheme 14** and their mode of blocking the hydrolytic attack can be seen in **Scheme 15**. It is worth noting that the α-carbonylphosphonate inhibitors can theoretically undergo nucleophilic attack by water in the same way as acetyl phosphate **2**; however, this is unlikely to occur as phosphonates are poor leaving groups.

2.0 Results and Discussion

Chapter I – Metabolic flux in *C. autoethanogenum*

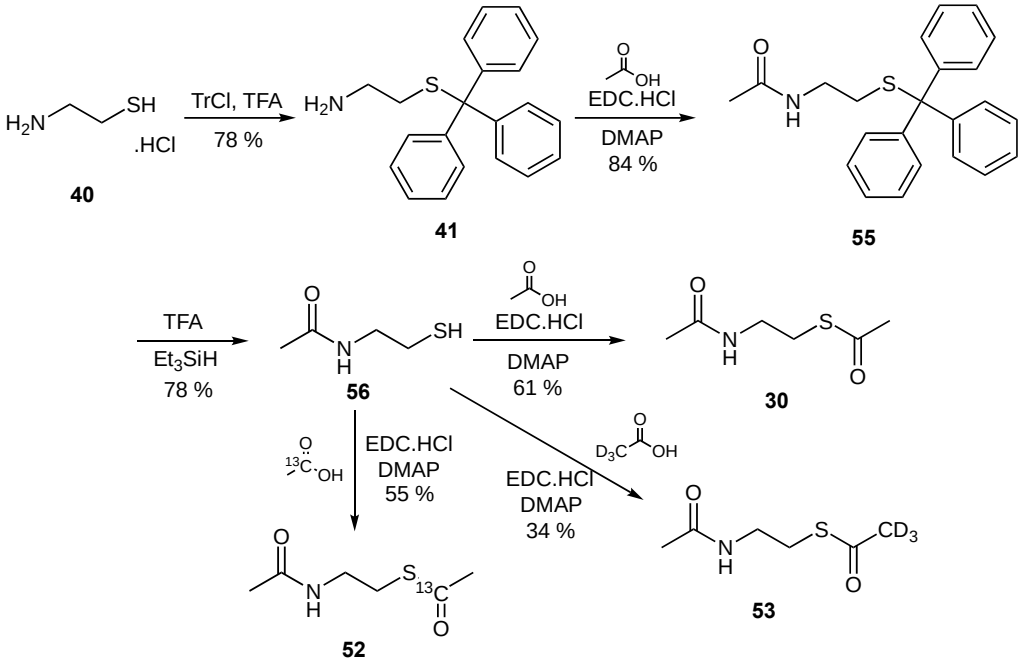
The initial aim of this project was to investigate the metabolic pathways in *C. autoethanogenum*, to examine the metabolic flux through the different pathways under differing conditions and track product formation, with potential to be used in conjunction with genetic knockouts to examine the resulting differences. This was to be performed by synthesis of the natural pathway metabolites incorporating ^{13}C or ^{2}H labels. The fate of the isotopic label could then be followed by NMR and/or mass spectroscopy.

Additionally, in collaboration with Ronja Breitkopf (RB), NMR analyses of *C. autoethanogenum* were performed to assess the metabolism of an alternate feedstock, fumaric acid **65**. Growth experiments were performed by RB, before handing over for NMR analysis, supported by Dr Huw Williams.

SNAC thioester synthesis

Initial work on the SNAC thioesters was to synthesise the putative mimic of acetyl-CoA **3**, acetyl-SNAC **30**, with an aim to generate a mimic which had improved cellular uptake over

the parent compound. The synthetic route undertaken can be seen in **Scheme 16**.



Scheme 16 – The synthesis of acetyl-SNAC thioesters incorporating both non-labelled and isotopically labelled material.

This first step of the synthesis followed work reported by Mandel *et al.*, in which cysteamine hydrochloride **40** underwent trityl protection with trityl chloride in trifluoroacetic acid/dichloromethane, to afford *S*-tritylcysteamine **41** in a 78 % yield (lit. 63 %) (this was later superceded by a different method, detailed in experimental section).[52] Subsequent EDC coupling with acetic acid provided *N*-(2-(tritylthio)ethyl)acetamide **55** in an 84 % yield. Trityl deprotection was achieved using trifluoroacetic acid with

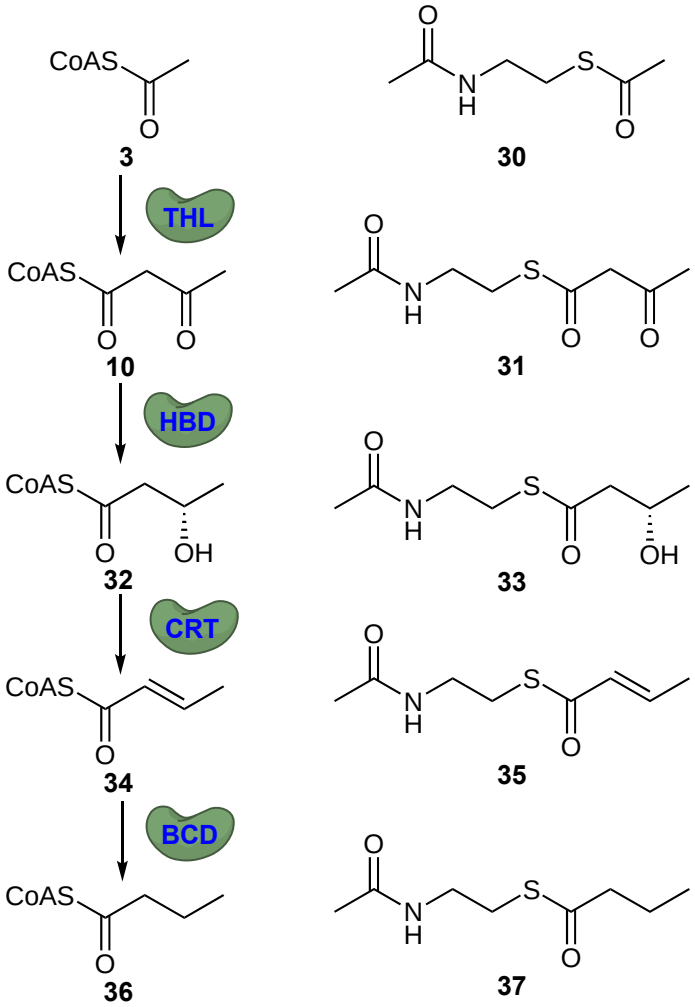
triethylsilane to sequester the trityl cation generated to afford *N*-acetylcysteamine **56** in a 78 % yield.

To obtain the acetyl-SNAC thioester **30**, EDC coupling of *N*-acetylcysteamine **56** with acetic acid gave the desired product in a yield of 61 %. However, synthesis of the non-isotopically labelled material was performed as a trial before incorporating the isotopic label. The material required for whole cell studies required a ^{13}C or ^{2}H isotopic label to be able to be observed by NMR and mass spectrometry as part of native metabolic pathways.

EDC coupling with the d_4-acetic acid and 1-^{13}C-acetic acid afforded the corresponding isotopically labelled SNAC thioesters, 2-$^{2}H_3$-acetyl-SNAC **52** and 1-^{13}C-acetyl-SNAC **53**, in 34 % and 55 % yields respectively. Although these yields were lower than the synthesis of the non-isotopically labelled compound, sufficient material was synthesised for testing in *C. autoethanogenum.*

The successful synthesis of isotopically labelled SNAC thioesters allowed whole cell growth experiments using *C.*

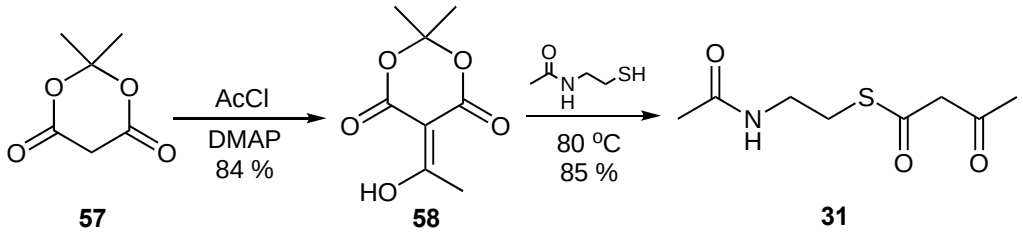
autoethanogenum to be performed, and analysis by NMR and mass spectrometry were completed, which is discussed later**.**



Scheme 17 – The thiolase pathway (left) showing the parent CoA molecules next to their SNAC thioester counterpart (right). THL = thiolase; HBD = hydroxybutyric-CoA dehydrogenase; CRT = crotonyl-CoA reductase; BCD = butyryl-CoA dehydrogenase.

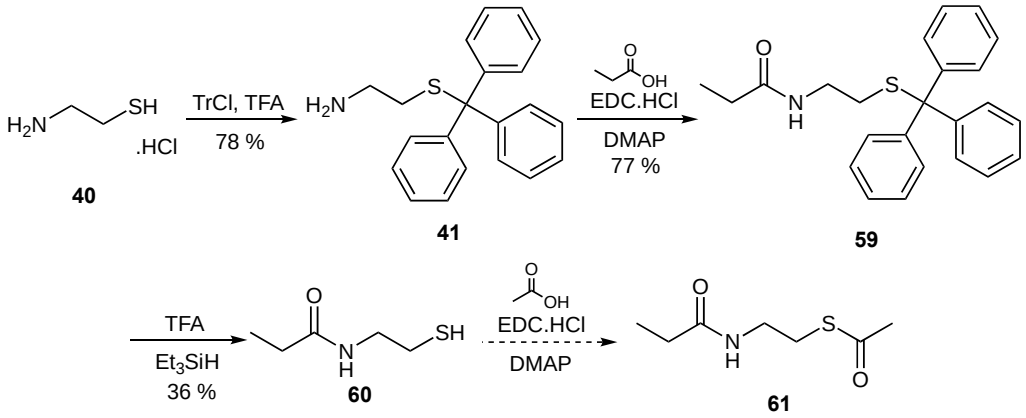
Following on from the synthesis of acetyl-SNAC **30**, further potential mimics of the potential thiolase pathway were identified, depicted in **Scheme 17**. Synthesis of crotonyl-SNAC **35** and butyryl-SNAC **37** thioesters was successful *via* EDC coupling in 66 % and 67 % yields respectively. As seen in **Scheme 17**, these SNAC thioesters may be comparable to

their parent CoA molecules if used as part of a whole cell experiment in *C. autoethanogenum.*



*Scheme 18 – Synthetic route to achieve the synthesis of acetoacetyl-SNAC **31**.*

As shown in **Scheme 18**, the synthesis of the acetoacetyl-SNAC **31** thioester was achieved by a different route based upon the work of Gao *et al*. and Gilbert *et al*.[47, 58] Initially, Meldrum's acid **57** was acetylated at the α-position with respect to the protected acids, to yield intermediate compound **58** in an 84 % yield. This allowed for the subsequent formation of acetoacetyl-SNAC **31** using *N*-acetylcysteamine **56** with heating to decarboxylate the intermediate **58** to form the desired product, acetoacetyl-SNAC **31** in an 85 % yield.



*Scheme 19 - The route to obtain an N-propyl derivative of the acetyl-SNAC thioester **61**, which was later superceded by work on pantothenate mimics.*

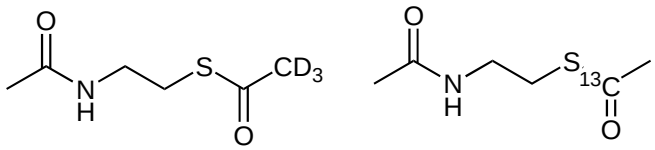
In an attempt to further increase the hydrophobicity of the SNAC thioester, work towards an *N*-propyl derivative of acetyl-SNAC **61** was commenced. The cLogP value for the *N*-propyl derivative **61** is 0.50, versus values of -6.04 for acetyl-CoA **3** and -0.03 for acetyl-SNAC **30**. These values were computationally obtained using ChemDraw Professional 16.0.

An *N*-propyl-*S*-acetylcysteamine thioester **61** can be seen in **Scheme 19** which utilised a similar synthetic route to the synthesis of acetyl-SNAC **30**. Upon obtaining *S*-trityl cysteamine **41**, EDC coupling with propionic acid afforded *N*-(2-(tritylthio)ethyl)propionamide **59** in 77 % yield. Further synthesis was not undertaken following unsuccessful whole cell growth experiments in the presence of the 2-^{13}C-acetyl-SNAC thioester **53** in *C. autoethanogenum.* This was due to a lack of uptake of the SNAC thioesters in the growth experiment with 2-^{13}C-acetyl-SNAC thioester **53**, discussed below.

Whole cell growth experiments using acetyl-SNAC thioesters

With the acetyl-SNAC thioesters incorporating ^{13}C and ^{2}H labels successfully synthesised, depicted in **Figure 12**, they were tested as part of a *C. autoethanogenum* whole cell

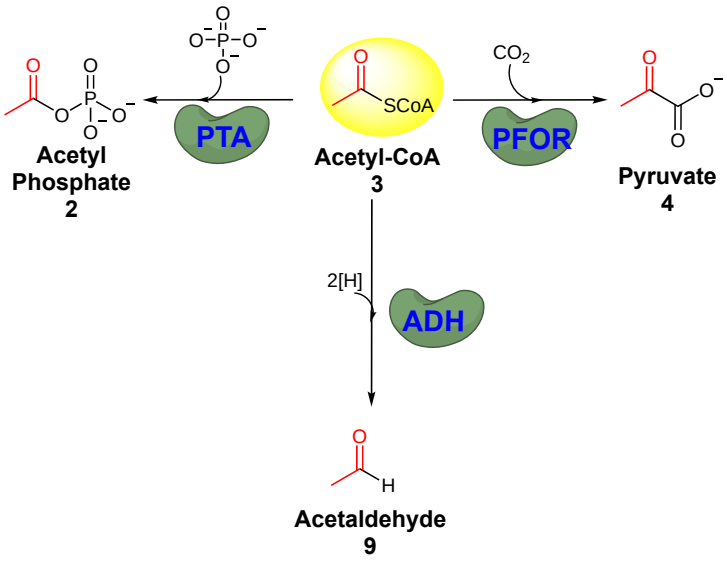
growth experiment, where they were added to the PETC-MES media used for bacterial culture growth.



52 **53**

Figure 12 – The two successfully synthesised isotopically labelled acetyl-SNAC thioesters.

SNAC thioesters were added to sterile medium before inoculation with *C. autoethanogenum*. The bacterium was grown, samples taken and analysed to measure uptake and utilisation of the thioesters. The SNAC thioesters used as part of this work can be seen in **Figure 12.** As observed in polyketide biosynthesis work by Gilbert *et al.,* where SNAC thioesters were recognised by the enzymes involved *in vitro*, investigations incorporating the isotopically labelled SNAC thioesters were undertaken in whole cell growth experiments. These were to determine the uptake and recognition by the active sites of the enzymes utilising acetyl-CoA **3** to assess the incorporated SNAC thioester and any metabolism towards downstream products.[46-48] In this case, the enzymes of interest were pyruvate ferredoxin oxidoreductase (PFOR) yielding pyruvate **4**, phosphotransacetylase (PTA) yielding acetyl phosphate **2** and alcohol dehydrogenase (ADH) yielding acetaldehyde **9**. These pathways can be seen in **Scheme 20**.



Scheme 20 – The metabolic pathways of interest branching from acetyl-CoA 3. PTA = phosphotransacetylase; PFOR = pyruvate ferredoin oxidoreductase; ADH = alcohol dehydrogenase.

If the SNAC thioesters were transported across the cell membrane and incorporated, it was envisaged that they would enter the pathways shown in **Scheme 20**, and the pathway end products incorporating an isotopic label could be detected by NMR or mass spectroscopy.

C. autoethanogenum underwent anaerobic growth on fructose in PETC-MES medium, which is detailed in the experimental section. The exception to this was that two of the sets of triplicates included isotopically labelled material, one with the ^{13}C-acetyl-SNAC and one with ^{2}H-acetyl-SNAC. Samples were taken from each of the experiments at selected timepoints to check the OD_{600}, before centrifugation, separation into pellet

and supernatant, and frozen storage at -80 ^{o}C. The timepoints taken for analysis were: 0 h, 12 h, 24 h, 36 h, 48 h, 72 h, 83 h, 103 h and 120 h.

It was observed that all of the cultures grew to a comparative OD_{600} to that of the control of ~1.8-2.0, suggesting that the SNAC thioesters were not toxic as part of the culture media, as *C. autoethanogenum* could grow in their presence. However, comparison of the OD_{600} showed no increased growth of cell culture as may be expected if SNAC thioesters were being incorporated into the biosynthetic pathways as an additional feedstock to the fructose in the media. This result suggested that the SNAC thioesters were not being incorporated and degraded for use as a source of carbon or energy in *C. autoethanogenum.*

For NMR spectroscopy, the growth experiment which included the 1-^{13}C-acetyl SNAC **53** was analysed. The ^{13}C growth experiment was selected for primary analysis because of the simplicity of visualisation of the ^{13}C-enriched peaks in the ^{13}C NMR experiment. This was due to a 100-fold increase in the intensity of the ^{13}C signal due to isotopic enrichment of the sample, which would allow easier visualisation in the NMR

spectra. The supernatant and cell pellet samples were analysed to determine presence of intracellular and extracellular concentration of 1-^{13}C-acetyl-SNAC **53** and any downstream metabolites incorporating a ^{13}C isotopic label which may have been biosynthesised during the growth experiment.

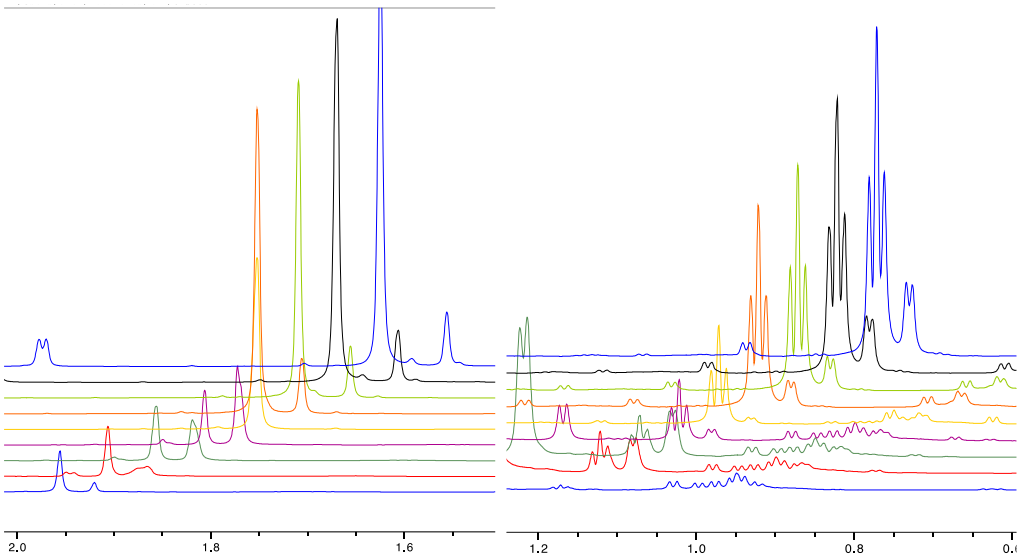


*Figure 13 – Zoomed ^{1}H NMR spectra taken from supernatants of growth experiments of ^{13}C-enriched feeding experiments. The increase of acetate **1** (left) and ethanol **8** (right) signals which denotes growth and metabolism of C. autoethanogenum. Timepoints were taken at 0 h, 12 h, 24 h, 36 h, 48 h, 72 h, 83 h, 103 h and 120 h.*

In addition to ^{13}C experiments, an excitation sculpted ^{1}H experiment incorporating a known concentration of DSS as an internal standard was performed, to allow for concentrations of any metabolites found to be determined. This showed an increase of the acetate and ethanol signals, depicted in **Figure 13**, which is a simple method for assessment of growth as the major product in *C. autoethanogenum*. The acetate peak can

also be seen, but not quantified, as part of the ^{13}C NMR experiment and is depicted in **Figure 14**.

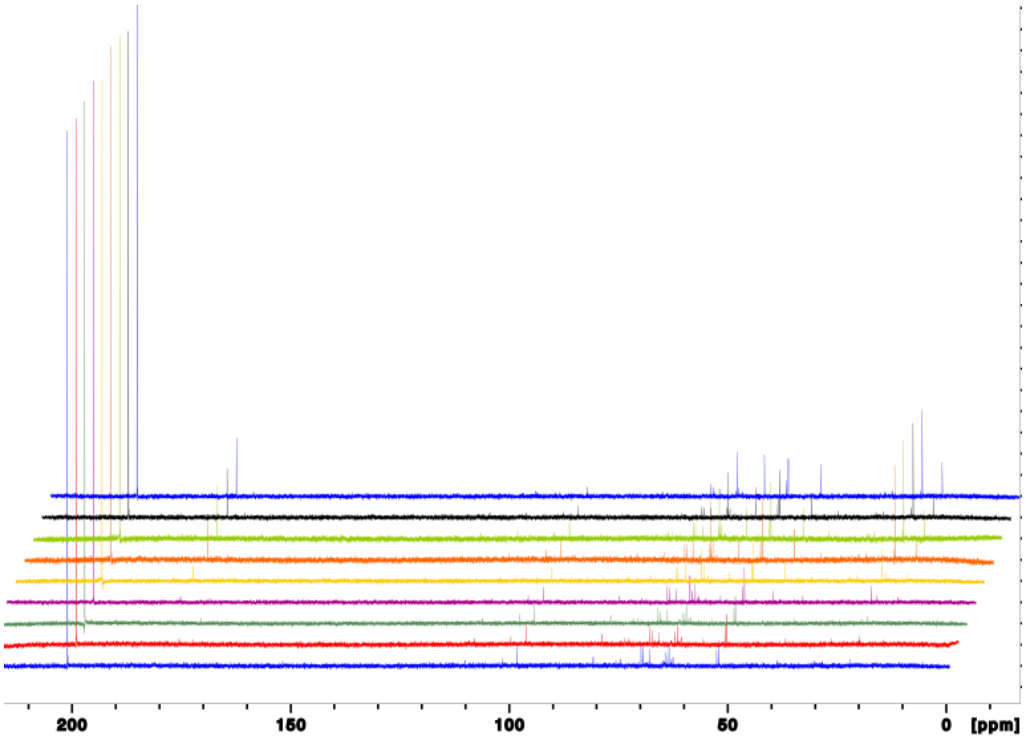


*Figure 14 – ^{13}C NMR spectra taken from timecourse of the 1-^{13}C-acetyl-SNAC **53** thioester growth experiment, the ^{13}C-enriched peak of which can be seen at 202.0 ppm. Timepoints were taken at 0 h, 12 h, 24 h, 36 h, 48 h, 72 h, 83 h, 103 h and 120 h.*

A ^{1}H,^{13}C-HSQC allowed for signals to be correlated between the two experiments. However, upon analysis of the supernatant samples, it was observed that the peak of the ^{13}C-enriched carbonyl of the SNAC thioester at 202 ppm in the ^{13}C NMR experiment did not change. This can be seen in **Figure 14**, where the intensity of the signal remains constant, and can be correlated with the spectra seen in **Figure 15** which shows no observable change in the concentration of the corresponding α-methyl of the SNAC thioester in the ^{1}H

experiment, which was determined by reference to the DSS internal standard of known concentration. This further supports the previous hypothesis that the SNAC thioesters were not transported across the cell membrane and incorporated into the natural metabolic pathways during the whole cell growth experiment.

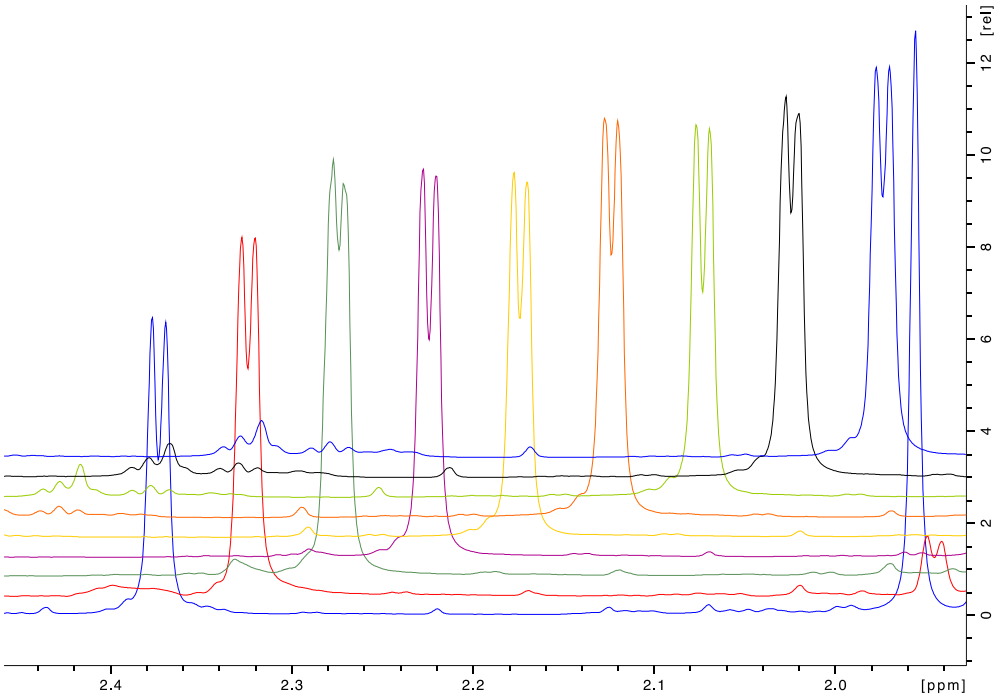


*Figure 15 – A zoomed ^{1}H NMR spectra showing the α-methyl peak to the isotopically labelled carbonyl peak within 1-^{13}C-acetyl-SNAC **53**, which remained constant throughout the timecourse experiment.*

Cell pellet samples were extracted according to the procedure detailed in the experimental section. Following cell lysis, samples were analysed by the same techniques as with the supernatant samples. Analysis of the pellet sample yielded poor quality spectra which did not contain useable data. This is likely due to the low concentration of internal cell metabolites

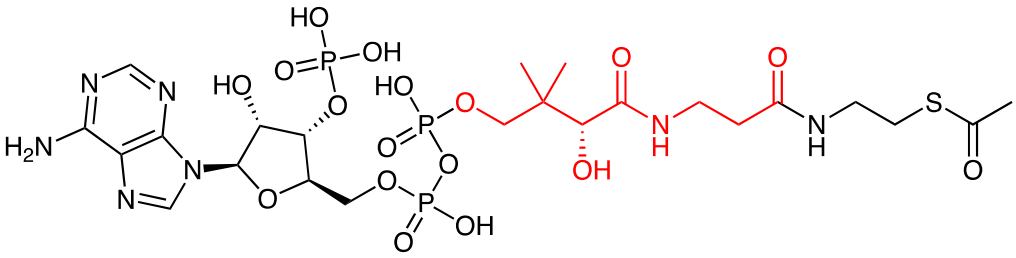
from the small pellet sample obtained being below the detection limits of the 800 MHz NMR instrument using the selected parameters. However, from analysis of the supernatant samples, it can be determined that the *C. autoethanogenum* appear to not uptake SNAC thioesters from growth media. This suggests that cells may not be permeable to SNAC thioesters as envisaged or that they are unable to use them as part of their native pathways, so further analysis optimisation of the pellet samples was not performed.

Due to the lack of uptake of SNAC thioesters, suggesting that *C. autoethanogenum* does not recognise them as a mimic for acetyl-CoA **3,** a new synthetic target was determined. This utilised a potential pantothenate mimic, 2-^{13}C-acetyl pantetheine **54** to be incorporated into acetyl-CoA **3** biosynthesis by *C. autoethanogenum* as part of its metabolic network.

^{13}C-Acetyl panthetheine synthesis

Following on from the unsuccessful incorporation of SNAC thioesters into metabolic pathways in whole cell *C. autoethanogenum* growth experiments, and based upon work within the group by Annan *et al.,* a new potential target was

devised.[51] Genome annotations by Brown *et al.* and Humphreys *et al.*, in conjunction with experimental work by Annan *et al.* have shown that the genes *panB, panC and panD* are not present in *C. autoethanogenum*, therefore it lacks the enzymes to be able to biosynthesise pantothenate.[28, 29, 51] This is important because pantothenate is a building block for the synthesis of acetyl-CoA **3**, shown in **Figure 16**, which as shown previously in **Scheme 6** is the branch point in the native pathways of *C. autoethanogenum*.

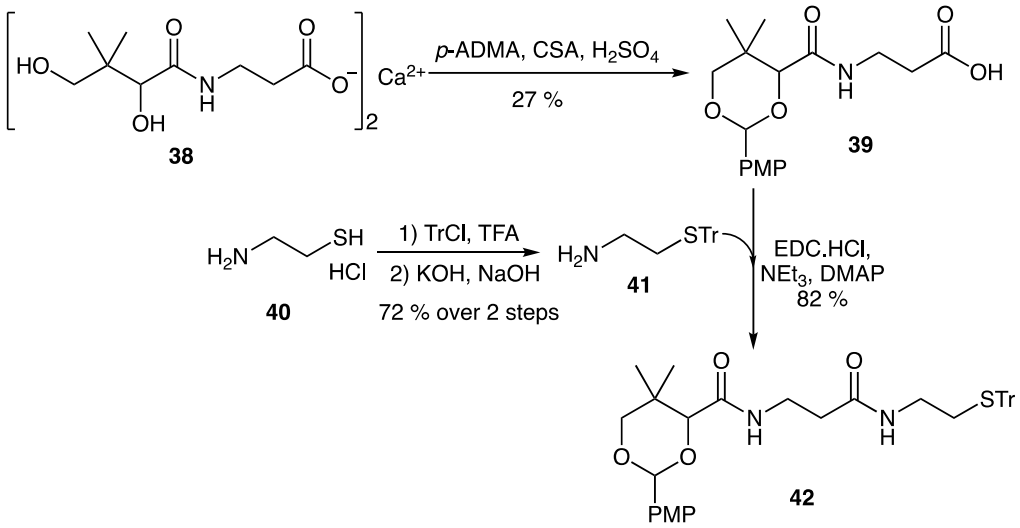


3

*Figure 16 – Acetyl-CoA **3**, with pantothenate highlighted in red, where it is incorporated into the CoA molecule.*

As part of the work on pantothenate feeding and limitation undertaken by Annan *et al.*, it was shown that pantothenic acid is essential in the growth media for *C. autoethanogenum*.[51] Pantothenate is added to the media as part of a vitamin solution known as Wolfe's vitamin solution. This provides an opportunity to synthesise 1-^{13}C-acetyl panthetheine **54** to replace pantothenate in the growth media in an attempt to incorporate a ^{13}C label into acetyl-CoA **3**.

Based upon this, attempts to synthesise 2-^{13}C-acetyl panthetheine **54** were undertaken.



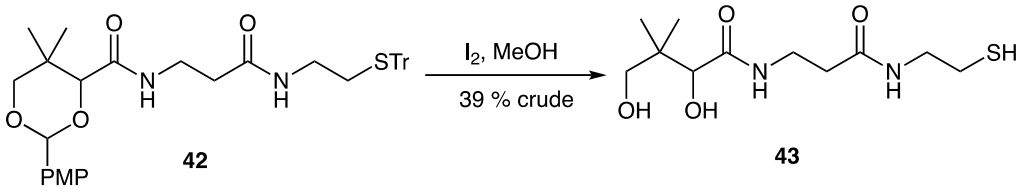
*Scheme 21 – The route to synthesise globally protected pantetheine **42**.*

A more efficient route than used previously in the SNAC thioester synthesis to obtain *S*-tritylcysteamine **41** was used. As shown in **Scheme 21**, following the work of Jullien *et al.*, *S*-tritylcysteamine **41** was successfully synthesised in two steps starting from cysteamine hydrochloride **40**, which was trityl-protected to give *S*-tritylcysteamine hydrochloride in a 90 % yield.[59] This underwent treatment with sodium hydroxide (9.6 M) in diethyl ether and potassium hydroxide to yield the free amine of *S*-tritylcysteamine **41** in 80 % yield.[59]

Concurrently, following a literature procedure from Patrone *et al.*, calcium pantothenate **38** was protected using *p*-anisaldehyde dimethyl acetal (*p*-ADMA) in the presence of

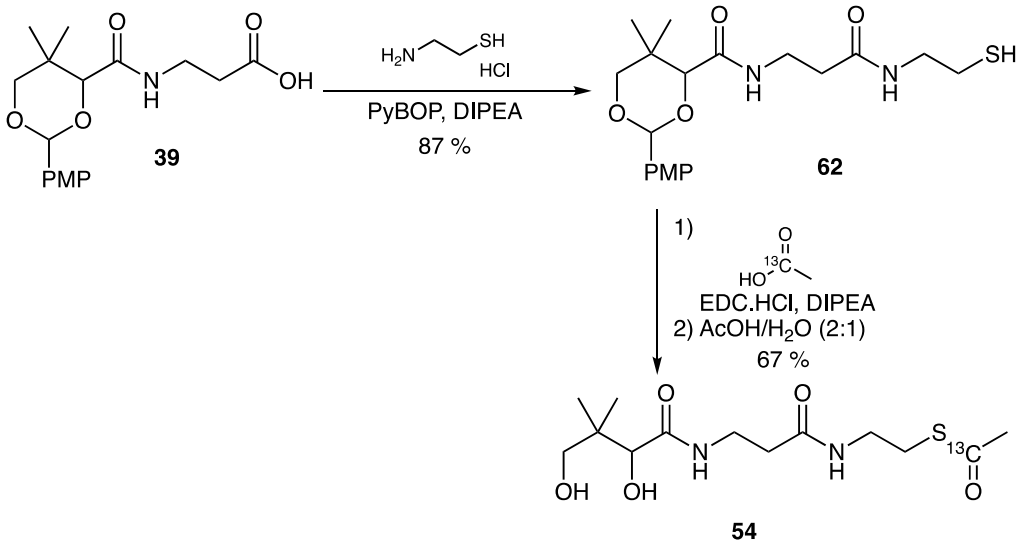
sulfuric acid and catalytic amounts of 10-camphorsulfonic acid (CSA).[60] This gave the PMP-protected pantothenate **39** in a 27 % yield, which although low, afforded sufficient material for further synthesis. One reason to account for the lower than reported yield (lit. 86 %) could be the age of the starting materials, which could have led to degradation. In addition, not previously observed in the literature was that the mixture turned to gel upon acidification with sulfuric acid, which was overcome by the addition of further solvent and sonication.

The successful PMP-protection allowed for coupling to *S*-tritylcysteamine **41**. This was initially attempted by synthesis of an NHS ester intermediate with DCC as used by Patrone *et al*.[60] Upon trialling the direct coupling reaction without NHS, it was shown to be unnecessary so was not included in further syntheses. The reaction proceeded successfully and product **42** was obtained in a 78 % yield. This was later optimised, with use of EDC.HCl as a replacement for DCC, which allowed a simpler purification method that did not require flash column chromatography, and removed any potential DCU byproduct, as well as a slightly improved yield of the coupled product **42** of 82 %.



*Scheme 22 – The global deprotection of compound **42** to yield impure pantetheine **43**.*

Following successful synthesis of the globally protected pantetheine, work by Mandel *et al.* which utilised an iodine solution in methanol to afford global deprotection of compound **43** was followed, as depicted in **Scheme 22**.[52] Upon attempting synthesis on a low milligram scale, deprotection was achieved, however, following purification by flash column chromatography, an impure product was obtained.



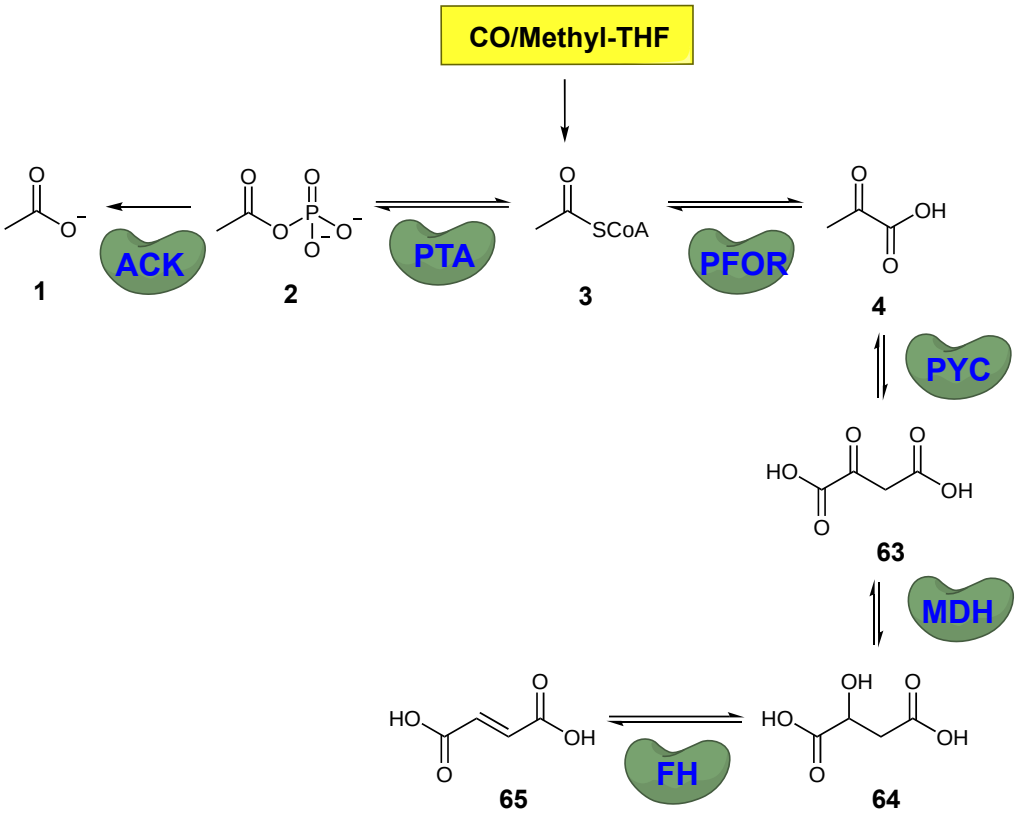
*Scheme 23 - The modified synthetic route based upon Agarwal et al. to yield 1-^{13}C-acetyl pantetheine **54**.[61]*

A new, more expedient synthetic route modified from Agarwal *et al.* was attempted to synthesise 1-^{13}C-acetyl pantetheine

54, shown in **Scheme 23**.[61] Coupling of cysteamine hydrochloride to PMP-protected pantothenate **39** using PyBOP was achieved in an 87 % yield. This differs from the published synthesis by use of PMP as the protecting group on the 1,3-diol, rather than an acetonide. Subsequent EDC coupling of 1-^{13}C-acetic acid with PMP protected pantetheine **39** was shown to be successful by mass spectrometry and NMR. The crude product was taken forward and deprotected using acetic acid:water to give the completed mimic, 2-^{13}C-acetyl pantetheine **54** in a 67 % yield over two steps.

Successful synthesis of the mimic **54** would have allowed for whole cell growth experiments in *C. autoethanogenum* to be performed, to investigate incorporation of 2-^{13}C-acetyl pantetheine into acetyl-CoA **3** biosynthesis. Unfortunately, there was insufficient time for whole cell growth experiments and analyses during this PhD.

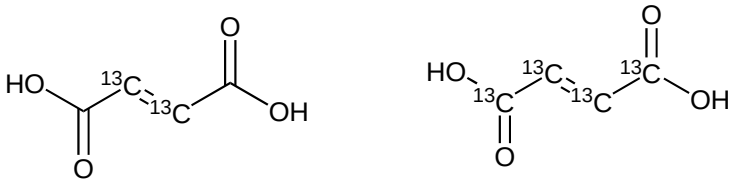
Fumarate growth experiments



*Scheme 24 – The proposed fumarate fermentation pathyway in C. autoethanogenum, showing the important interconversions between the feedstock of these metabolomics experiments, fumaric acid **65**, and the major product, acetate **1**. ACK = acetate kinase; PTA = phosphotransacetylase; PFOR = pyruvate ferredoxin oxidoreductase; PYC = pyruvate carboxylase; MDH = malate dehydrogenase; FH = fumarate hydratase.*

Whole cell growth experiments using fumaric acid **65** as a feedstock were undertaken to investigate *C. autoethanogenum* metabolism. For isotopic label detection through metabolic pathways, ^{13}C was used for labelling the fumaric acid **65** feedstock, with growth experiments being undertaken and samples collected at various timepoints for analysis. The initial aim of this work was to determine that the pathway proposed in **Scheme 24** is the pathway responsible for growth in

fumarate metabolism, generating acetate **1** as its major product. At each timepoint, 1 mL samples were taken, centrifuged to separate the pellet from the supernatant and both were frozen for analysis.



66 **67**

*Figure 17 – The two isotopically labelled fumaric acid compounds used as part of this work, 2,3-^{13}C2-fumaric acid **66** or ^{13}C4-fumaric acid **67**.*

Either 2,3-$^{13}C_2$-fumaric acid **66 or** $^{13}C_4$-fumaric acid **67**, shown in **Figure 17**, were used for growth experiments and approximately 30 - 40 mM was added to the cultures before inoculation from pre-cultures of *C. autoethanogenum*. All growth experiments and sampling were performed by Ronja Breitkopf (RB) and samples handed over for analysis.

Both growth experiments involved NMR analysis of supernantant samples only, as experimental acquisition data for the pellet samples showed insufficient concentration of metabolites to be detected by the 800 MHz instrument according to the experimental parameters used.

Thus, for each supernatant sample, three analysis experiments were performed: a ^{1}H NMR for identification and concentration of metabolites; a $^{1}H,^{13}C$-decoupled NMR for determination of ^{13}C peaks; and a $^{1}H,^{13}C$-HSQC NMR experiment for correlation to ^{13}C spectra. These experiments were performed quantitatively and detailed in the experimental section. ^{13}C NMR experiments were performed for some timepoints for reference of the ^{13}C peaks of any detectable metabolites incorporating a ^{13}C-label, using data from the human metabolome database (HMDB) for identification purposes.[62-64] Samples were made up in H_2O:D_2O (90 % : 10 % v/v) with a known concentration of DSS as an internal standard for calculating the concentrations of the metabolites in the samples.

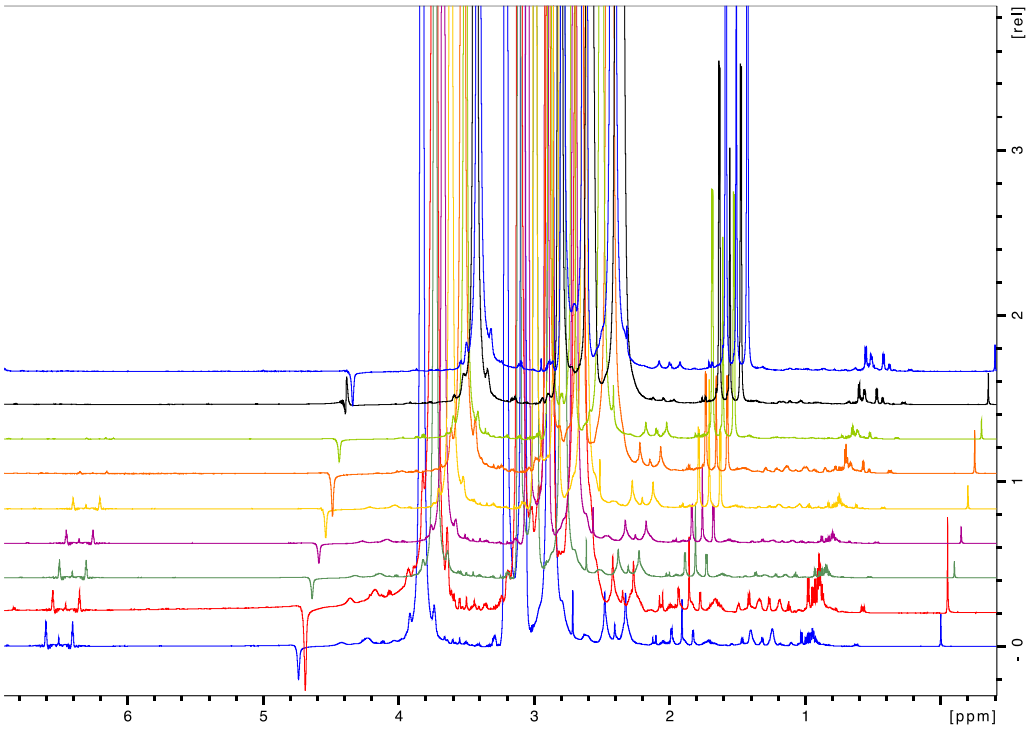


Figure 18 – A stacked ^{1}H spectrum taken from timepoints of a batch growth from a $^{13}C_4$-fumaric acid growth experiment, showing fumaric acid at 6.53 ppm, and acetate at 1.92 ppm.

As shown in **Figure 18**, $^{13}C_4$-fumaric acid **67** can be seen at 6.52 ppm, as well as acetate **1** at 1.92 ppm. The three large signals on the spectra between 3 and 4 ppm are the MES buffer system used for the whole cell growth experiments. Signals for metabolites in these regions are more difficult to visualise in this experiment because of the overlap of these intense signals. Other metabolites detectable in micromolar concentrations incorporating a ^{13}C label have been determined by reference to HMDB to be alanine and ethanol **8**.[62-64] Ethanol **8** concentration is significantly lower when compared to samples grown on fructose or gaseous feedstocks. The ethanol

8 and alanine detected in these growth experiments were re-metabolised, as their corresponding peaks disappeared after approximately 50 hours. The detection of alanine suggests incorporation of a metabolite from the pathway in **Scheme 24** into an amino acid pathway, most likely pyruvate **4** converting to alanine *via* alanine transaminase; however this was not investigated further in this project.

The batch growth experiment shows that as the concentration of $^{13}C_4$-fumaric acid **67** decreases, the ^{13}C-acetate concentration increases. The increased signal for ^{13}C-acetate suggested conversion of $^{13}C_4$-fumaric acid **67** *via* the pathway in **Scheme 24**, however, none of the intermediate metabolites were detectable.

One drawback to this method was sampling of timepoints didn't allow for concentrations of the metabolites to be observed in real time, and there was concern that smaller concentrations of intermediate compounds in the supernatant might be missed. It was therefore decided to move towards whole cell growth in an anaerobic NMR tube with continuous analysis.

The in-tube timecourse experiments were carried out over a period of 120 hours, where samples were agitated every 8 hours, ensuring good mixing of the cells within the media. Every hour, a ^{1}H NMR, a ^{1}H,^{13}C-decoupled NMR and a ^{1}H,^{13}C-HSQC NMR experiment were performed to explore the metabolism of the cells and follow the progress of the metabolic pathways using the isotopic labels. As previously, both $^{13}C_4$-fumaric acid or 2,3-$^{13}C_2$-fumaric acid were tested and any differences noted.

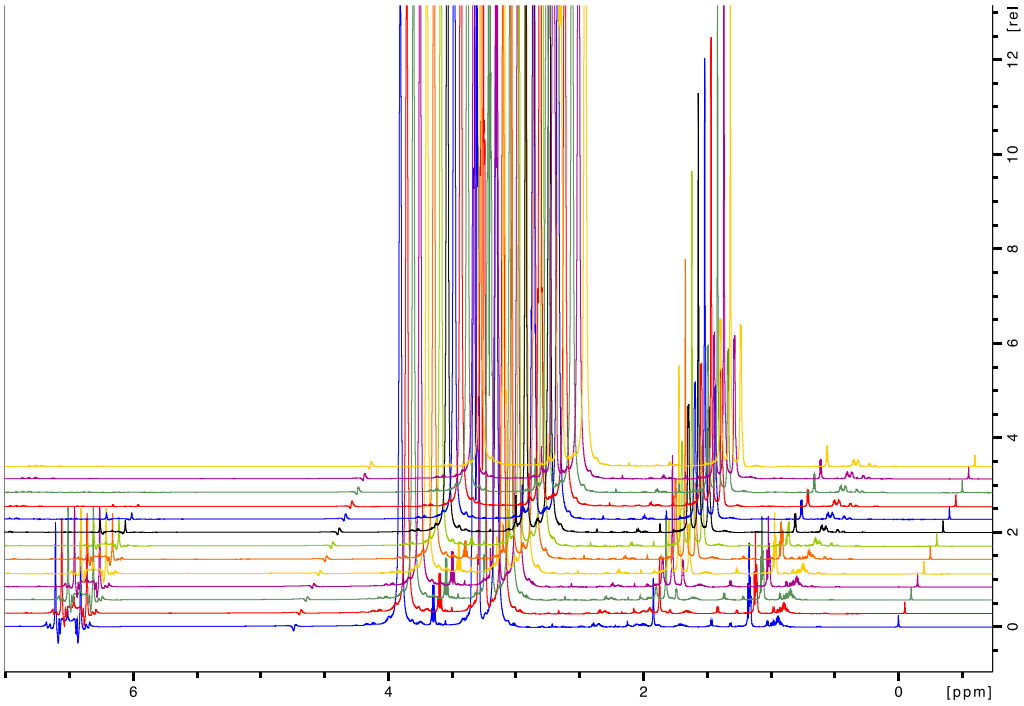


Figure 19 – A stacked plot of the ^{1}H experiments from whole cell growth in an NMR tube, showing a decrease in the concetration of 2,3-$^{13}C_2$-fumaric acid 66 in the sample, as well as an increase in ^{12}C and ^{13}C acetate in the sample. The timepoints shown are 0 h, 8 h, 16 h, 24 h, 32 h, 40 h, 48 h, 56 h, 64 h, 72 h, 80 h, 88 h and 96 h.

Initially, 2,3-$^{13}C_2$-fumaric acid **66** was used as the feedstock for the bacterial culture, which showed an increase in both the

^{12}C and ^{13}C acetate signals for the duration of the experiment, as well as the 2,3-$^{13}C_2$-fumaric acid **66** being consumed. The other compounds to have a detectable change in concentration as part of the growth were alanine and ethanol **8**, however, these were in low micromolar (<50 μM) concentrations and were re-metabolised by the organism. A stacked plot of the overall spectra can be seen in **Figure 19**, and more detailed images of fumaric acid and the acetate over the timecourse can be seen in **Figure 20**.

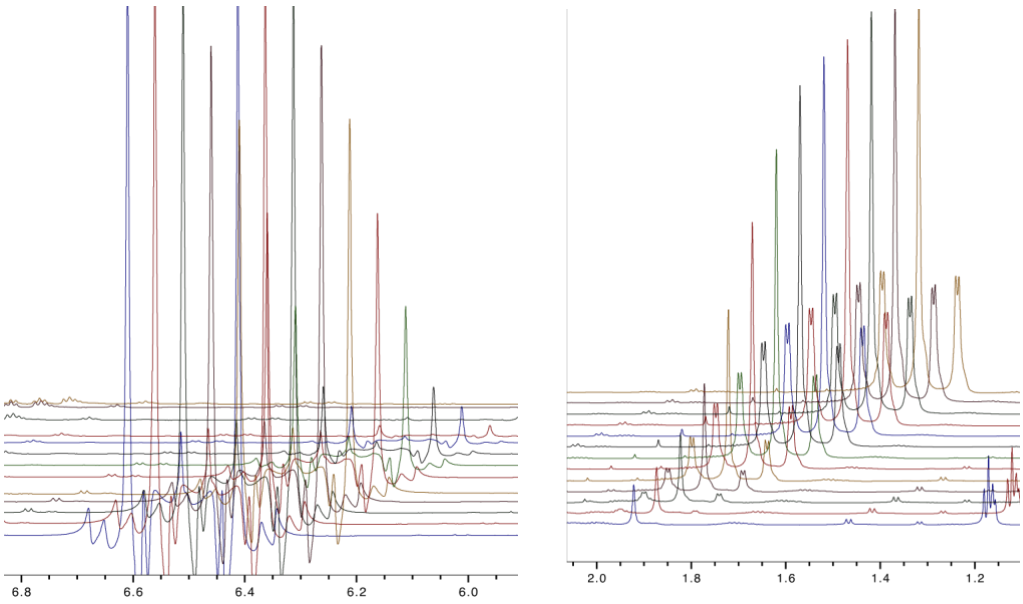


Figure 20 – Zoomed ^{1}H NMR spectra showing stacked plots of the fumaric acid (left) decreasing as acetate (right) signals increase. It can be seen that both ^{12}C and ^{13}C peaks can be seen to decrease and increase respectively.

The stacked plots in **Figure 20** show the signal for 2,3-$^{13}C_2$-fumaric acid **66** decreasing over a period of 72 hours, and the ^{12}C and ^{13}C acetate **1** signals increasing. Interestingly, this

experiment shows a large increase in ^{12}C-acetate **1** over the course of the growth experiment, which suggests that *C. autoethanogenum* has a ^{12}C feedstock present, shown in **Figure 21**. These were calculated by reference of the integration of each peak to the integration of the DSS peak, which was of a known concentration, allowing the concentration of metabolites to be determined.

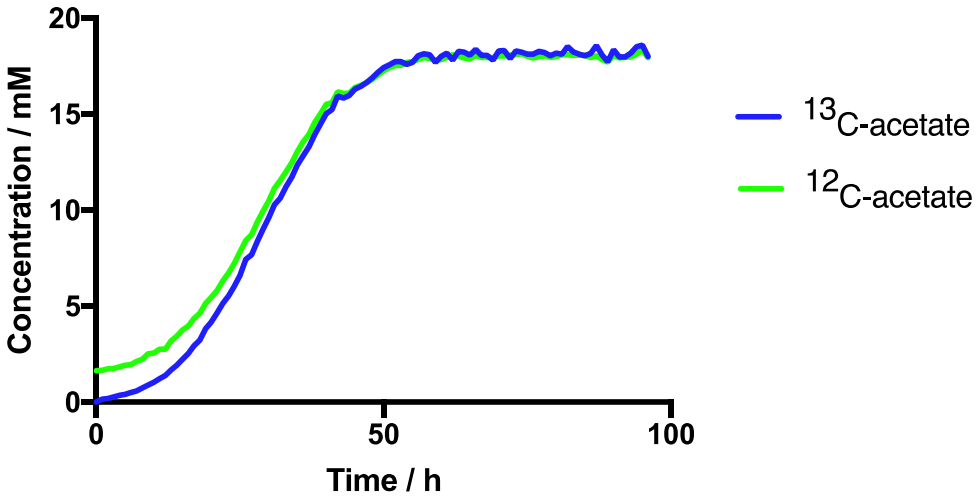


*Figure 21 – The formation of acetate **1** in the 2,3-$^{13}C_2$-fumaric acid **66** growth experiment, showing equal formation of ^{13}C-acetate and ^{12}C-acetate **1**.*

The graph in **Figure 21** shows the increase in both ^{13}C-acetic acid and ^{12}C-acetate **1** to concentrations of 18 mM, supporting evidence of presence of a ^{12}C feedstock throughout the timecourse experiment.

These data, in conjunction with knowledge of the chemical pathway, suggest that the decarboxylation of oxaloacetic acid

63 to pyruvate **4** and of pyruvate **4** to acetyl-CoA **3** provides carbon dioxide that can be utilised by the organism as a feedstock. This supports a hypothesis that the Wood-Ljungdahl pathway, shown in **Scheme 5**, is active in the organism during fumarate metabolism, which would utilise the carbon dioxide from the decarboxylation to be incorporated into acetate **1** production.

One issue with the methodology was quantification of the isotopically labelled fumaric acid, which appears to be a symmetrical molecule. However, these NMR experiments showed magnetic inequivalence between the two α-protons of the alkene. This results in a complex multiplet, caused by second order effects, which spans a large section of the baseline ~6.5 ppm.

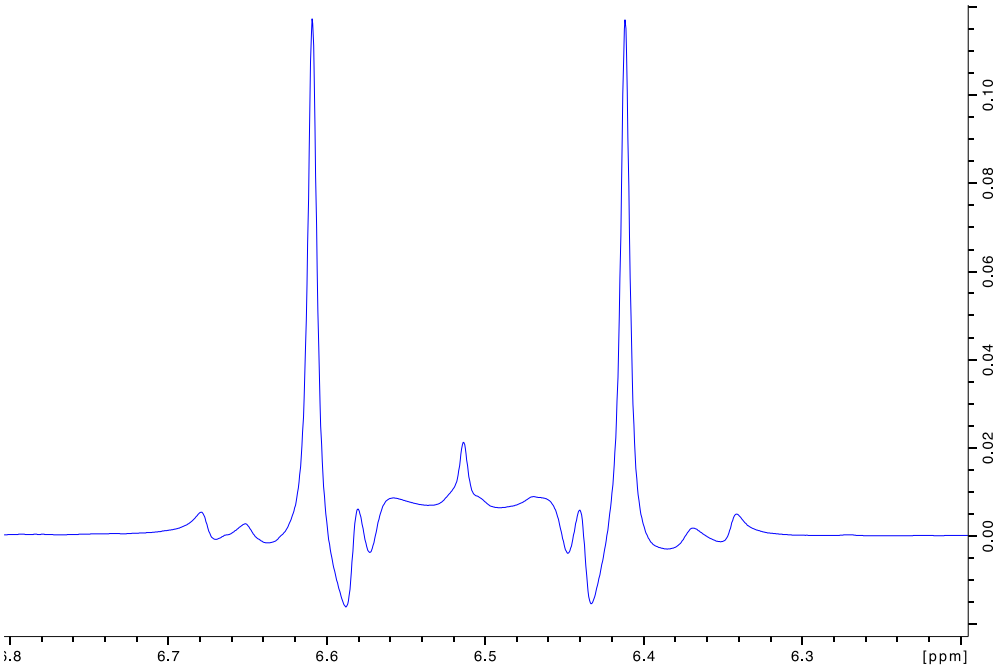


*Figure 22 – The ^{1}H NMR signal for 2,3-$^{13}C_2$-fumaric acid **66**, which has undergone undesired signal suppression.*

This multiplet also undergoes suppression as part of the water suppression parameters of the ^{1}H experiment used. Because of the unwanted suppression of the 2,3-$^{13}C_2$-fumaric acid **66** signal, the peak has an anti-phase component seen in **Figure 22**, which leads to an incorrect integration of the peak and it is unable to be properly quantified. While product quantification was still possible, an accurate concentration curve for the rate of consumption of 2,3-$^{13}C_2$-fumaric acid **66** was not possible. Hence, the graphs in **Figure 21** show product formation, but the decrease in concentration of the 2,3-$^{13}C_2$-fumaric acid **66** is not shown. This solvent suppression method, known as WATERGATE, has been studied

by Morris *et al.* and by variation of the pulse program, it is possible to obtain higher quality spectra which in this case, does not affect the isotopically labelled fumaric acid.[65, 66] However, during method development for the perfect echo sequence used by Morris *et al.,* it was discovered that it was possible to make a different modification for these specific experiments, which utilised a NOESY 1D experiment including solvent suppression, which did not affect the isotopically labelled fumaric acid signal.

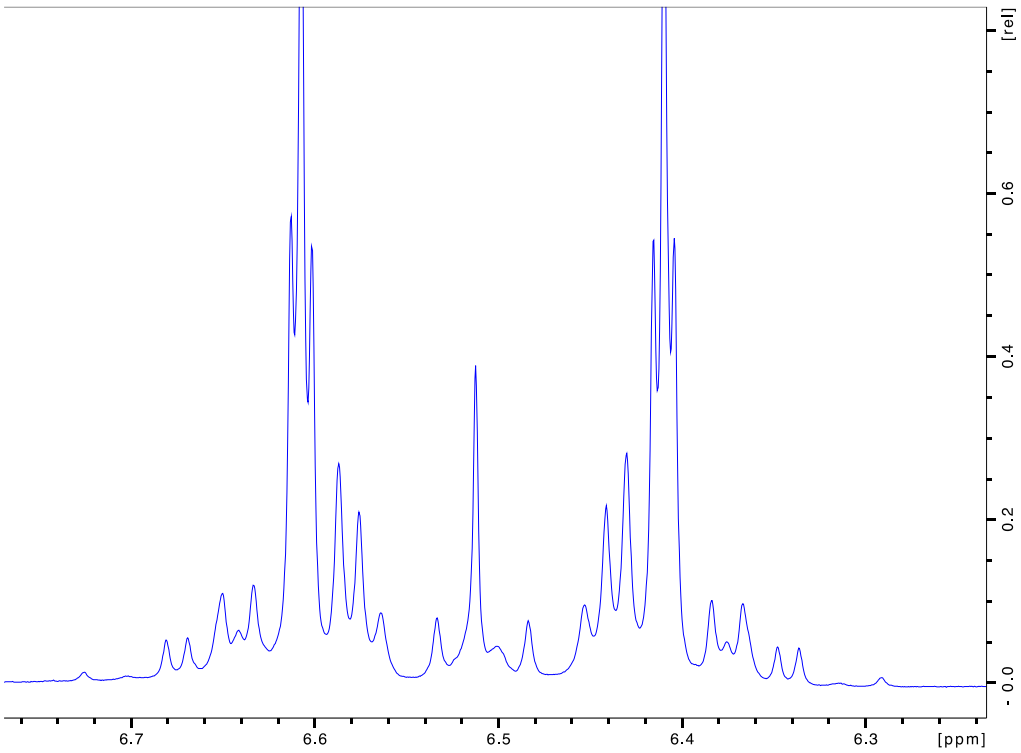


*Figure 23 – The $^{13}C_4$-fumaric acid **67** signal, which was obtained using a modified ^{1}H NMR pulse program to show all of the multiplicity and lead to a quantifiable integration value.*

Following on from the 2,3-$^{13}C_2$-fumaric acid **66** experiment, $^{13}C_4$-fumaric acid **67** growth experiments were carried out in an NMR tube. With the modification to the ^{1}H program used, a multiplet for fumarate was obtained, where all of the peaks were in-phase and quantifiable. The peak multiplicity can be seen in **Figure 23**, which has additional complexity when compared to the peak in the 2,3-$^{13}C_2$-fumaric acid **66** experiment due to the additionally labelled 1,4-carbons interacting with long range coupling.

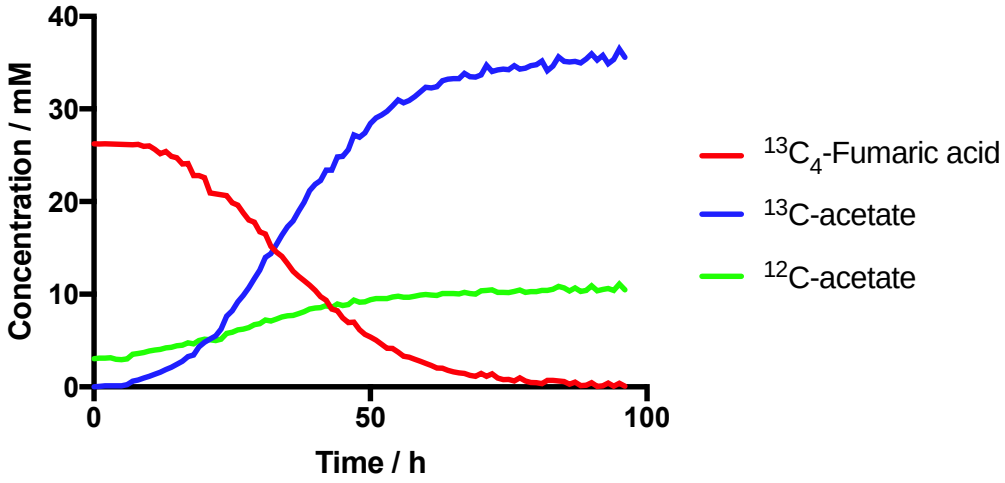


*Figure 24 – The concentration of the feedstock and major product of the $^{13}C_4$-fumaric acid **67** in-tube growth experiment.*

As in previous experiments, ^{1}H NMR, a ^{1}H,^{13}C-decoupled NMR and ^{1}H,^{13}C-HSQC NMR experiments were carried out and analysed. With a quantifiable multiplet for $^{13}C_4$-fumaric acid **67** now obtainable, a concentration curve of feedstock and

product could now be plotted, showing the quantitative conversion of $^{13}C_4$-fumaric acid **67** to ^{13}C-acetic acid. The graph in **Figure 24** shows an increase in concentration of ^{13}C-acetate to a concentration of 36 mM, from a starting concentration of 26 mM of $^{13}C_4$-fumaric acid **67**. The conversion is consistent with the data from the other experiments which show complete utilisation of the isotopically labelled feedstock to generate isotopically labelled ^{13}C-acetate.

This correlates with the previous data and hypothesis that the Wood-Ljungdahl pathway is active, and is metabolising carbon dioxide that is a product of the decarboxylation of oxaloacetic acid **63** in its reaction to form pyruvate **4** and pyruvate **4** to acetyl-CoA **3**, seen in **Scheme 24**. The concentration of ^{12}C-acetate **1** present can be explained by non-isotopically labelled fumaric acid **65** which remained from bacterial pre-cultures.

These isotopically labelled NMR experiments, initially batch growth followed by an in-tube growth timecourse experiment have shown that fumarate metabolism is possible in *C. autoethanogenum* with acetate **1** as the major product. It has also provided evidence for the pathway in **Scheme 24** being active in the organism to metabolise fumaric acid **65**, as well

as support for the hypothesis that the Wood-Ljungdahl pathway, shown in **Scheme 5**, is active during fumarate metabolism.

Summary

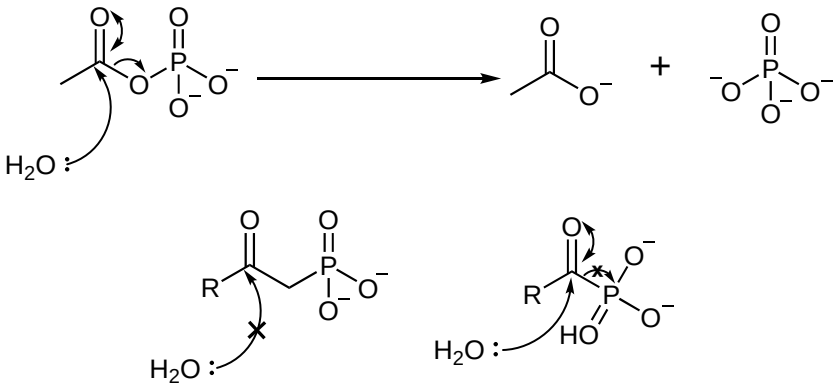
Acetyl-SNAC **30**, 2-$^{2}H_3$-acetyl-SNAC **52**, and 1-^{13}C-acetyl-SNAC **53** were successfully synthesised and tested in whole cell growth experiments in *C. autoethanogenum*. These were shown not to be incorporated into the bacterial metabolic network, due to lack of cellular uptake. This was demonstrated by ^{13}C NMR experiments, ^{1}H NMR experiments and ^{1}H,^{13}C-HSQC experiments over a timecourse were extracted samples were analysed.

Following on from this, attempts were made towards the synthesis of 1-^{13}C-acetyl pantetheine **54** with an aim for this to be incorporated into the biosynthetic pathways of *C. autoethanogenum* as a substitute for pantothenate as part of a whole cell growth experiment. The synthesis was successful following a method based upon the work of Agarwal *et al.* to give 1-^{13}C-acetyl pantetheine **54** for use in whole cell growth experiments, which were not completed during the project.[61]

Fumarate growth experiments were undertaken in collaboration with Ronja Breitkopf, where analysis of samples and timecourse experiments revealed that fumaric acid **65** is metabolised to acetate **1** *via* two pathways. One of these pathways is depicted in **Scheme 24** and shows metabolism *via* pyruvate **4** and acetyl-CoA **3**, and the other is *via* formation of a carbon dioxide feedstock from the decarboxylation of oxaloacetic acid **63** and pyruvate **4**, which can be metabolised by the Wood-Ljungdahl pathway.

Chapter II – Synthesis of acetate kinase inhibitors

Acetate kinase (ACK) was identified as a target for transient inhibition using a small molecule inhibitor in *C. autoethanogenum* because the *ack* gene, which encodes the ACK enzyme, cannot be knocked out whilst retaining growth, as shown by Cooksley *et al.* in a sister organism, *C. acetobutylicum*.[39]



*Scheme 25 – The hydrolytic cleavage of acetyl phosphate **2** to acetate **1** and free phosphate, and its blockage by the two potential classes of inhibitor.*

As mentioned previously, no crystal structure for acetate kinase from *C. autoethanogenum* was available, therefore an *in silico* docking approach was not taken to inhibitor design. Instead, ligand-based structure design was adopted for synthesis of the inhibitors. The inhibitors chosen were modelled upon the substrate, acetyl phosphate **2,** and the two families of compounds that were identified as potential inhibitors and their potential mode of action can be seen in **Scheme 25**.

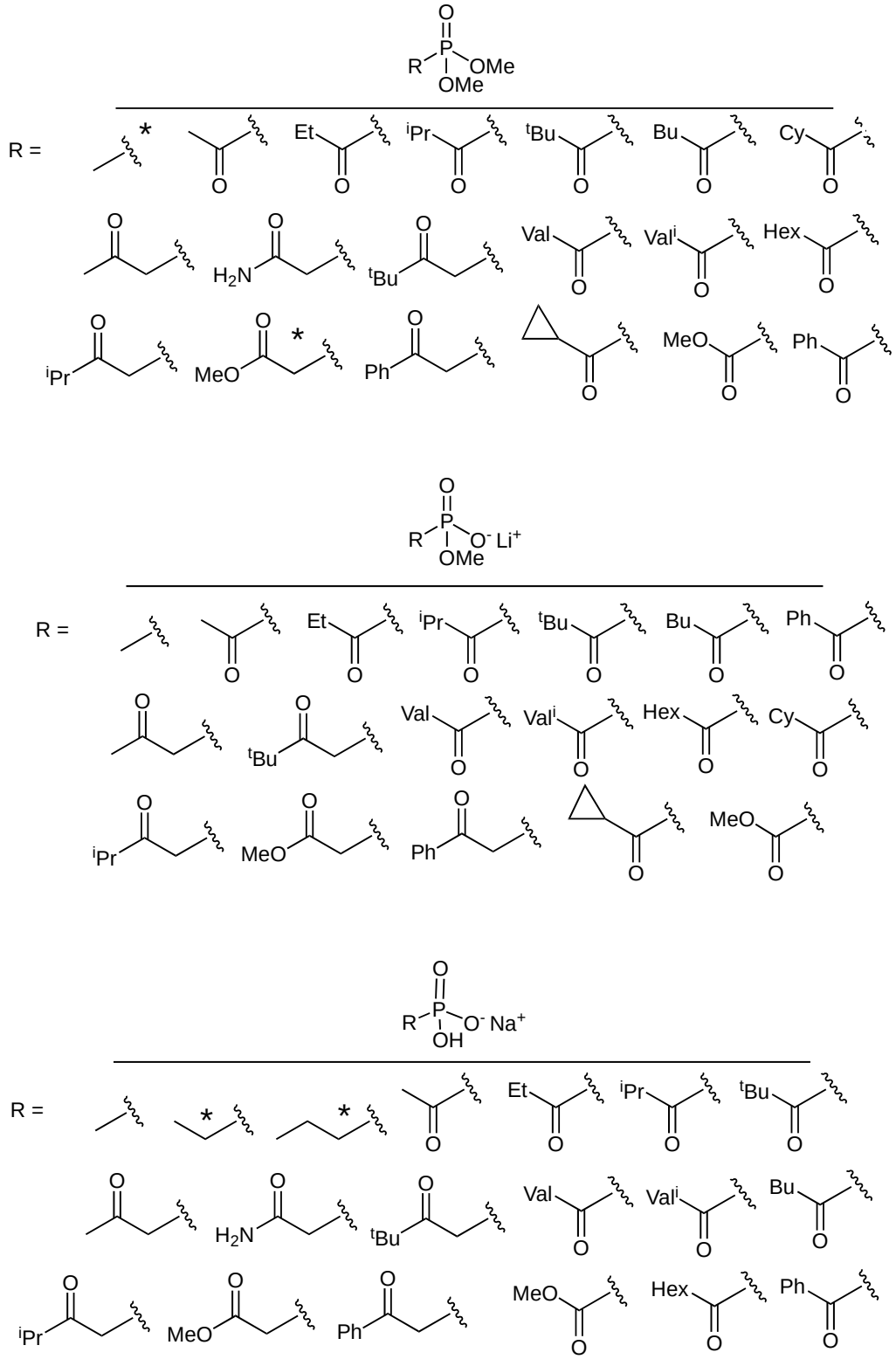


*Figure 25 – The compounds to be tested as part of the inhibition screen against acetate kinase from C. autoethanogenum. Those marked with a (*) were purchased from Sigma-Aldrich. Diethyl pyrocarbonate **139** was also tested as a known inhibitor of acetate kinases.*[67-70]

The inhibitor compounds which were synthesised to be tested against ACK during this work can be seen in **Figure 25**. Initially, these compounds were synthesised as the dimethyl phosphonate esters, which were all oils, as well as lithium monoester and phosphonic acid sodium salt derivatives to yield white or coloured powders. These de-esterified compounds were thought to more closely resemble the substrate in terms of polarity and potential interactions. Additionally, solid group I metal salts allowed easier handling when making solutions for inhibition studies.

Initially, both synthetic routes followed literature procedures, which were later modified to give better yields or easier purifications as required.[71-77] A number of the methods and compounds were developed with assistance from MSci or summer project students Crystal Quist, Freya F. H. Taday and Diana Kołos. The range of compounds chosen was to probe the interactions in the active site. Therefore, a number of functional groups were used and are listed in **Figure 25**, to test for a range of interactions. Compounds incorporating aromatic R-groups were included to investigate π-π stacking or π-cation interactions, others incorporating both hydrogen bond

donors and/or acceptors to look for hydrogen bonding and some with large alkyl chains to obtain information on any hydrophobic interactions. This structural scanning was performed because the nature of the active site was unknown. It could potentially lead to further development of the inhibitor if there was a successful 'hit' with any of the compounds.

Synthesis of putative β-ketophosphonate inhibitors

Initial synthetic targets for inhibitors of acetate kinase utilised a methylene isostere of the oxygen atom linking the phosphonate moiety to the carbonyl, examples of which are shown in **Figure 26**. This removes a hydrolytically labile bond, and means that if the active site recognises the inhibitor, it may undergo similar binding, but would not cleave into acetate **1** and free phosphate for ATP regeneration.

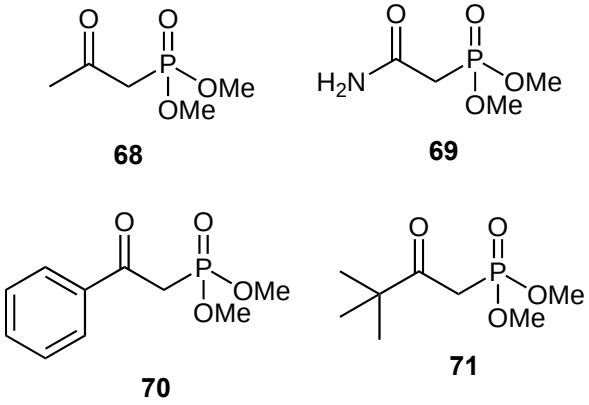
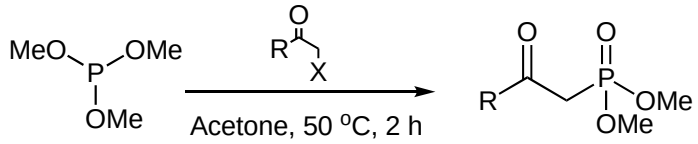


Figure 26 – The four initial synthetic targets for inhibition to mimic acetyl phosphate, containing a methylene isostere.

Initially four synthetic targets were identified, as seen in **Figure 26**, which would be the starting point for synthesis and provide a range of R-groups which could probe the interactions within the active site. These diester compounds, along with their subsequent monoester and phosphonic acid derivatives, were modelled on the substrate with an isostere, and were expected to be the most effective competitive inhibitors. Dimethyl (2-oxopropyl)phosphonate **68** is an analogue containing the methylene isostere and same functional group α- to the carbonyl as the substrate; compound **69** was designed to test for hydrogen bond acceptors close to the binding pocket which can be exploited; compound **70** was designed to probe for any π-π stacking or π-cation interactions; and compound **71** was designed to determine if larger alkyl groups with greater hydrophobicity would be tolerated within the active site.

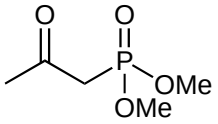
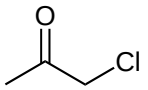
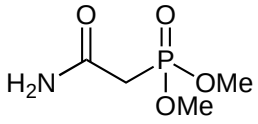
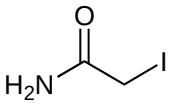
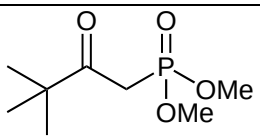
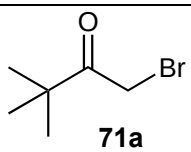
These four compounds were synthesised based on work by Arbuzov, utilising methods adapted from Kosobokov *et al.* and Sirivolu *et al.*[72, 76, 78] The general reaction scheme for synthesis of β-ketophosphonate compounds from trimethyl phosphite and the corresponding α-haloketone can be seen in **Scheme 26**. Initially, compounds requiring conversion to the

corresponding iodide underwent a Finkelstein reaction by addition of potassium iodide to the reaction shown in **Scheme 26**.[79] Reactions with a bromide starting material were used without halogen exchange as they were sufficiently reactive to form good yields of products.



Scheme 26 – The general reaction to produce β-ketophosphonates using trimethyl phosphite and the corresponding α-haloketone, where X denotes the halide used.

The four compounds (**68-71**) were successfully synthesised in acceptable yields for further synthesis and testing. The compounds produced from this general method and the yields of the reactions can be seen in **Table 1**.

Compound	Corresponding α-haloketone	Yield (%)
 **68**	 **68a**	58 %
 **69**	 **69a**	62 %
 **71**	 **71a**	48 %*

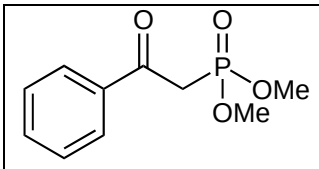
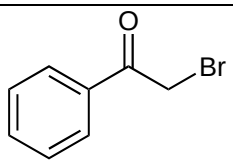
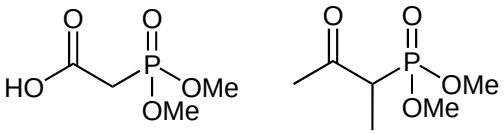
 **70**	 **70a**	55 %

*Table 1 – The library of β-ketophosphonate compounds synthesised via the reaction in Scheme 26. Starred compound (*) **71** was synthesised by an MSci student, Crystal Quist.*

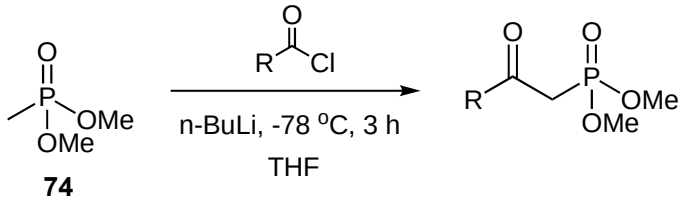
Synthesis of further compounds incorporating other functionality was also attempted *via* this method, which can be seen in **Figure 27**. Synthesis of a phosphonate ester **72** based on iodoacetic acid, designed to test hydrogen bonding similar to the amide derivative **69,** was unsuccessful. If the amide compound was to show promising results in inhibiting acetate kinase, further development could go into the method to synthesise the acid derivative. The unsuccessful synthesis could be caused by the presence of the free acid, which may hinder the reaction by acidifying the reaction. Attempts were also made to react 3-chloro-2-butanone with trimethyl phosphite to yield an α-methyl substituent **73**. The reaction was successful and the product was seen *via* mass spectrometry but unfortunately, the product was inseparable from the starting material even after two rounds of flash column chromatography.



72 **73**

Figure 27 – Two compounds that could not be prepared by the initial synthesis method.

The quantities of phosphonate esters obtained were sufficient for synthesis of phosphonic acid and/or monomethyl ester derivative synthesis to be performed. However, purification of the compounds synthesised *via* this methodology proved difficult due to co-elution of the phosphonate esters with the starting material, trimethyl phosphite. This led to multiple rounds of flash column chromatography being required for single products and numerous purifications proving unsuccessful.



74

Scheme 27 – The revised synthesis method for the β-ketophosphonate esters, used for the synthesis of compounds with bulkier functional groups.

A new synthetic method, depicted in **Scheme 27**, to synthesise compounds with bulkier R-groups was investigated based upon work by Lundy *et al.* This consisted of utilising n-butyllithium in tetrahydrofuran at -78 °C to deprotonate methyl dimethyl phosphonate followed by reaction with acid

chlorides to give the corresponding phosphonate esters.[74] A new putative inhibitor **75** was synthesised based on the compound taken from the work of Lundy *et al.,* as well as the preparation of two previously reported compounds, the *tert*-butyl **71** and phenyl **70** derivatives. The compounds produced from this general method and the yields of the reactions can be seen in **Table 2**.

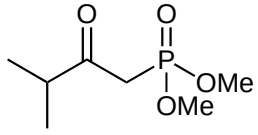
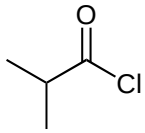
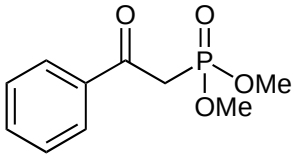
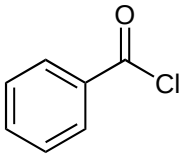
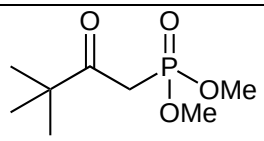
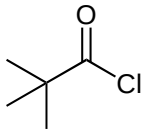
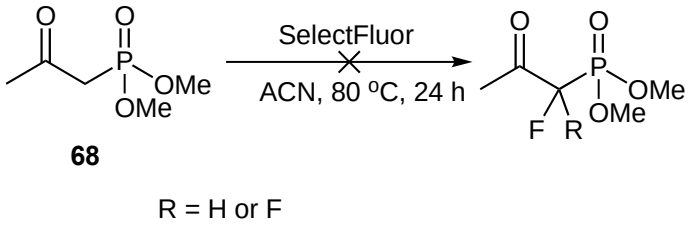
Compound	Corresponding acid chloride	Yield (%)
 **75**		43 %
 **70**		58 %
 **71**		56 %

Table 2 – Synthesis via a modified method from Lundy et al.[74]

Synthesis of the compounds listed in **Table 2** was achieved in good yields, with a simpler purification step, consisting of gradient column chromatography allowing the product to elute in pure ethyl acetate after removal of impurities. Synthesis was also attempted using smaller acid chlorides to obtain

inhibitor compounds bearing less bulky functional groups, such as methyl or ethyl derivatives. However, these smaller acid chlorides lead to some addition of two R-groups and an inseparable mixture of products. It would appear that this method is only suitable for bulkier acid chlorides, as they may block a second deprotonation, or slow the deprotonation sufficiently that only one addition of the acid chloride occurs.



68

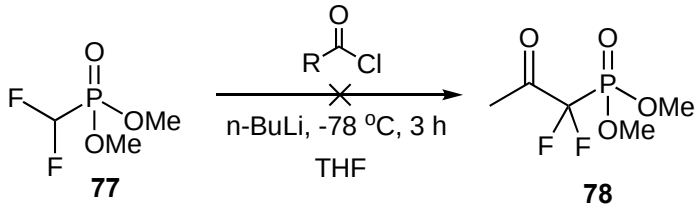
R = H or F

*Scheme 28 – The unsuccessful fluorination of dimethyl (2-oxopropyl)phosphonate **68**.*

Further analogues containing fluorine atoms on the carbon α-to the phosphorous atom were attempted, as depicted in **Scheme 28**. Synthesis of both the mono-fluoro **76** and di-fluoro **78** compound were tried, according to the procedure by Radwan-Olszewka *et al.*; however, the products were never purified successfully *via* flash column chromatography to a purity suitable for inhibitor testing, generating a mixture of mono and di-fluoro products.[80]

It is unclear why this reaction was not successful as it was in the literature, although the only difference between the literature and the procedure documented here is the use of

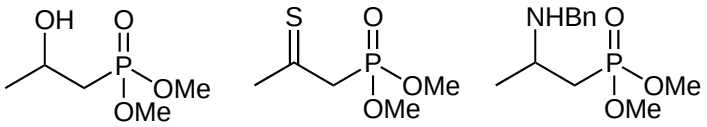
trimethyl phosphite as opposed to triethyl phosphite in the literature.



77 **78**

*Scheme 29 – Attempt to synthesise the difluoro compound **78** using the method adapted from Lundy et al.[74]*

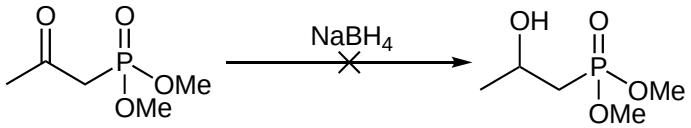
Synthesis of the difluoro compound **78** was also attempted *via* a method adapted from that of Lundy *et al.*[74] However, this was not successful as it had been previously with the compounds shown in **Table 2**. The mixture obtained did not contain any desirable products, and starting material could not be seen, as determined by mass spectrometry.



79 **80** **81**

Figure 28 – Other modifications which were attempted to test the interactions of the active site of the enzyme.

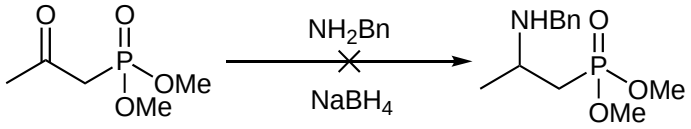
Further modifications to dimethyl (2-oxopropyl)phosphonate **68** were to attempt transformations of the carbonyl to investigate the effect of modification of a hydrogen bond acceptor in the compounds, as seen in **Figure 28**. Therefore, as seen in **Scheme 30**, a reduction of the carbonyl was attempted to yield compound **79**.



68 **79**

*Scheme 30 – The attempted reduction of the carbonyl of dimethyl (2-oxopropyl)phosphonate **68**.*

The desired reduction was partially successful using sodium borohydride, shown in **Scheme 30**, where limited conversion to the product was seen by mass spectrometry and NMR spectroscopy. However, the product was never fully converted and contained starting material, which following purification attempts using flash column chromatography, still contained a large amount of impurities and would be unsuitable for testing *in vitro* and *in vivo.* To rule out issues with the reagent affecting the reduction, different batches of sodium borohydride were used in an attempt to produce pure compound **79**; unfortunately, these were also unsuccessful.

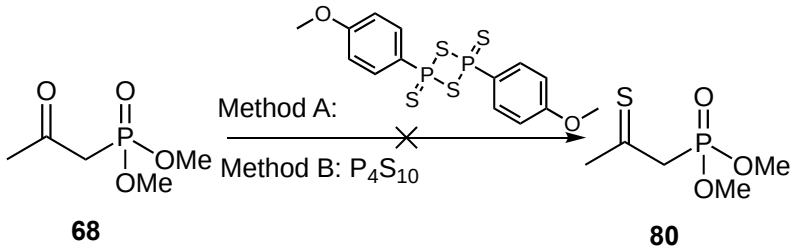


68 **81**

*Scheme 31 – Attempted reductive amination of the carbonyl of dimethyl (2-oxopropyl)phosphonate **68**.*

Reductive amination of dimethyl (2-oxopropyl)phosphonate **68** using benzylamine was attempted, which would allow for easy

hydrogenolysis of the benzyl group to yield an amine in the β-position. However, when attempted as in **Scheme 31,** a mixture of products was obtained which included starting material, a β-hydroxyl and the β-amine, which were not separated through flash column chromatography. An alternative method could be to attempt to achieve the amine synthesis *via* an oxime, similar to the work of Radwan-Olszewska *et al*.[80]



68 **80**

Scheme 32 – Unsuccessful attempts to produce a β-thioketone compound.

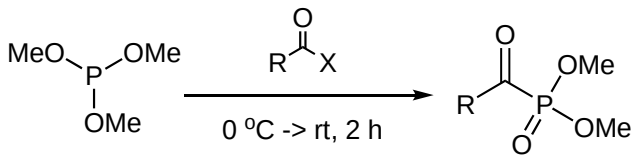
As in **Scheme 32**, two methods were attempted to synthesise a thioketone derivative **80**, a weaker H-bond acceptor, as a potential inhibitor. Both methods, utilising either Lawesson reagent or phosphorous pentasulfide as the thionating agent, were unsuccessful, generating no product, and resulting in only starting material being present at the end of the reaction. It is possible that the conditions used for this thionation were too mild, as some literature suggests that use of a microwave reactor gives the desired reaction, whereas these attempts at

thionation were performed using room temperature or reflux conditions similar to the protocol by Curphey *et al.*[81]

Five β-ketophosphonate compounds were successfully synthesised following two different routes, which were determined to be pure by NMR. These compounds can be used in enzyme assays as well as undergoing subsequent reactions to give the monoester and phosphonic acid derivatives, which are described later.

Synthesis of putative α-carbonylphosphonate inhibitors

The other class of compounds of interest in this research were the α-carbonylphosphonate compounds, which differ from the acetyl phosphate substrate by removal of the oxygen linkage between the carbonyl and phosphorous moieties. This species has very different properties to the substrate, such as the orientation of the α-carbonyl, making them of particular interest should they be shown to be successful inhibitors. As stated previously, this change to the molecule removes a hydrolytically labile bond, and should the active site recognise the inhibitor as a potential substrate, it may undergo competitive binding. However, when bound into the active site these substrates would not cleave, as seen in **Scheme 25**.



Scheme 33 – The reaction used to synthesise α-carbonylphosphonate compounds as part of this research.

Synthesis of a library of these compounds was based upon work by Whitteck *et al.* who had previously synthesised dimethyl acetylphosphonate **82**.[77] The methodology used in the original paper was adapted to a longer reaction time of 2 h after addition of trimethyl phosphite, to ensure a complete reaction and afford higher product yields. A general reaction scheme for the procedure can be seen in **Scheme 33**. Acid chlorides were stirred at 0 °C, with the slow, dropwise addition of trimethyl phosphite key to the success of the reaction, resulting in the evolution of chloromethane gas as a by-product of the reaction. If the trimethyl phosphite was added too quickly, the exothermic reaction resulting from the addition to acid chlorides appears to cause the reaction to give lower product yield. Generation of chloromethane indicates the reaction is progressing successfully.

Of interest is that the slow addition of neat materials required means that smaller scale reactions were more difficult, therefore all of the α-carbonyl phosphonates were synthesised

on multigram scale. Removal of any remaining starting materials *in vacuo* was sufficient to afford suitably pure material to be used in further reactions as well as taking the compounds forward for testing *in vitro* in enzyme assays.

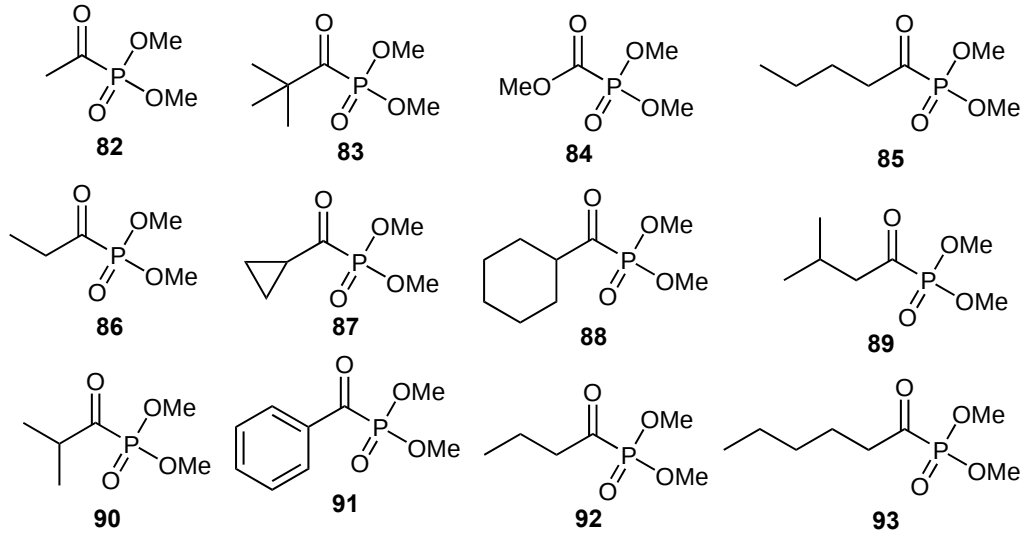


Figure 29 – The range of α-carbonylphosphonates successfully synthesised as part of this work.

The compounds generated had a range of functional groups, as indicated in **Figure 29**. As with the previous series of inhibitors, the β-ketophosphonate compounds, the R-group variation was to investigate the binding in the active site of acetate kinase. The full library of compounds developed for this class of compounds can be seen in **Figure 29**. These compounds were all successfully synthesised in good yields, which are illustrated in **Table 3**.

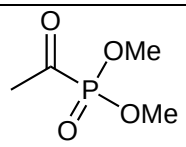
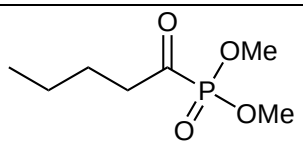
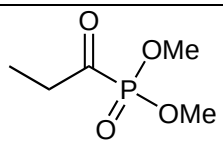
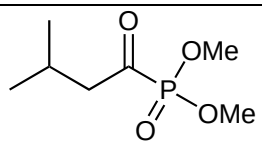
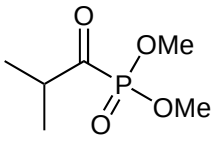
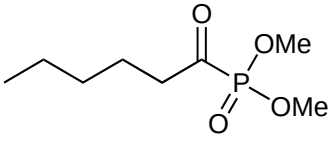
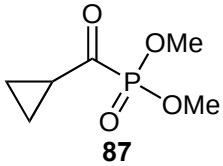
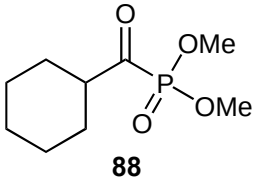
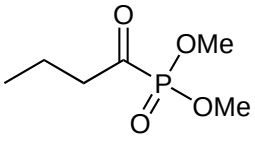
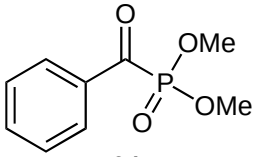
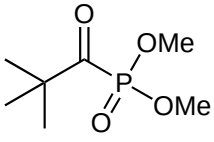
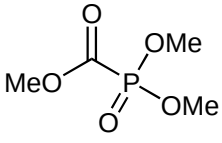
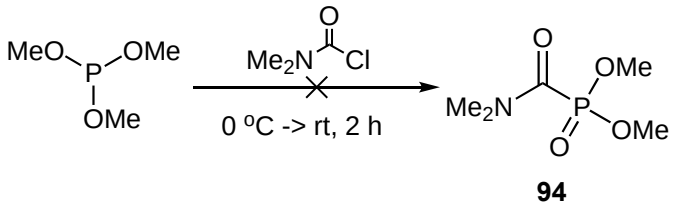
Compound	Yield	Compound	Yield
 **82**	90 %	 **85**	93 %
 **86**	90 %	 **89**	90 %
 **90**	94 %	 **93**	92 %
 **87**	91 %	 **88**	89 %
 **92**	96 %	 **91**	93 %
 **83**	66 %	 **84**	74 %

Table 3 – Yields of synthesis of α-carbonylphosphonate compounds. Product yield was lost due to lack of conversion of starting materials potentially arising from addition being too quick, leading to increased reaction temperatures.

One method to further enhance the yield of these reactions could potentially be the use of a syringe pump to reduce the addition speed of the trimethyl phosphite to the acid chloride,

as opposed to the pressure equalising dropping funnel used as part of this work.



94

Scheme 34 – The unsuccessful reaction of dimethyl carbamoylchloride with trimethyl phosphite.

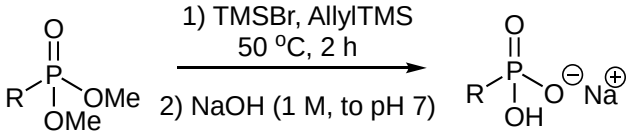
The synthesis of one further analogue was attempted using dimethyl carbamoyl chloride as the acid chloride, but was unsuccessful as seen in **Scheme 34**. This result could be due to the decreased electrophilicity of the acid chloride due to the donating effect of the substituted nitrogen atom, leading to isocyanate formation. No reaction was observed, which was confirmed by mass spectrometry, where no product was observed.

Twelve compounds in this series were successfully synthesised which were taken forwards for further reaction, and enzyme assays *in vitro*. Their purity was assessed by NMR.

Synthesis of putative phosphonic acid inhibitors

To more closely mimic the di-acid moiety of the acetyl phosphate substrate, it was necessary for the phosphonate

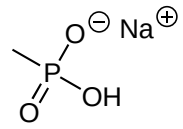
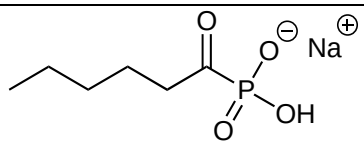
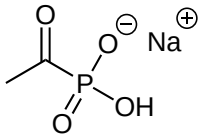
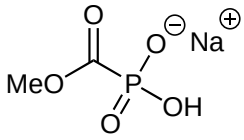
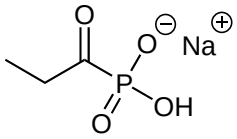
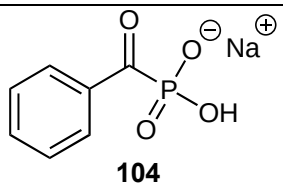
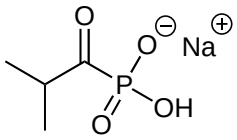
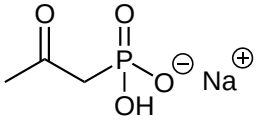
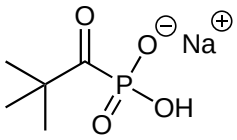
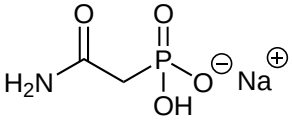
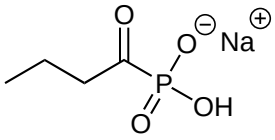
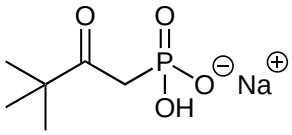
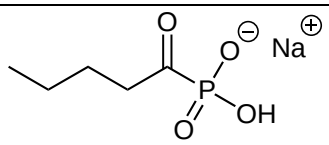
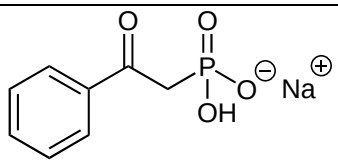
esters to undergo double dealkylation to yield phosphonic acids. Because the phosphonic acids have the same acid group as the acetyl phosphate **2** substrate, it was predicted that they may possess greater inhibition in the enzyme assay.



Scheme 35 – The method used for dealkylation of phoshonate esters to yield phosphonic acids as their sodium salts.

The dealkylation of phosphonate esters is well documented in the literature, and is based upon the original work of McKenna *et al*.[82] Initially, a procedure modified by Karaman *et al.* was attempted to afford the phosphonic acid compounds using trimethylsilyl bromide, but this method was unsuccessful in generating pure compounds that could be used in enzyme assays.[71] To overcome this, a modified procedure described by McGrath *et al.* lead to synthesis using trimethylsilyl bromide and allyltrimethylsilane to yield the phosphonic acids as sodium salts, as seen in **Scheme 35**.[75]

Of the eighteen phosphonate compounds to hand, sixteen were successfully converted to phosphonic acids using this procedure. These were obtained in good yields, which can be seen in **Table 4**.

Compound	Yield	Compound	Yield
 **95**	86 %	 **102**	88 %
 **96**	85 %	 **103**	91 %
 **97**	97 %	 **104**	61 %
 **98**	82 %	 **105**	94 %
 **99**	87 %	 **106**	93 %
 **100**	87 %	 **107**	91 %
 **101**	78 %	 **108**	84 %*

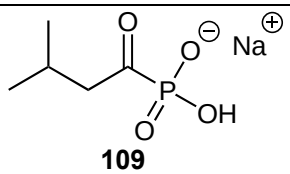
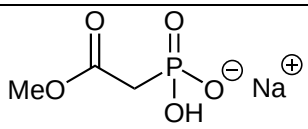
 **109**	83 %
 **110**	88 %

Table 4 – The yields of the successfully synthesised phosphonic acid inhibitor compounds.

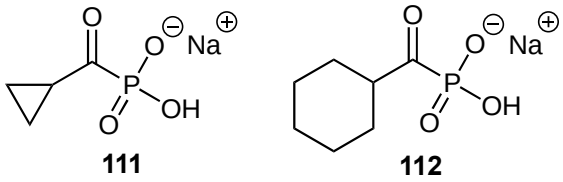


Figure 30 – The two compounds which were not successfully synthesised via this method.

The two compounds that were not successfully reacted by this procedure were the cyclopropyl **87** and cyclohexyl **88** derivatives of the α-carbonylphosphonates, seen in **Figure 30**. It is unclear as to why they were unsuccessful, as the reaction proceeded to completion, but upon lyophilisation, yielded impure material which could not be purified by recrystallisation, as it was not readily soluble in many organic solvents. Both ethanol and acetone mixes with ether were used for the attempted recrystallisation, but this was unsuccessful, and these two compounds were not taken forwards for testing.

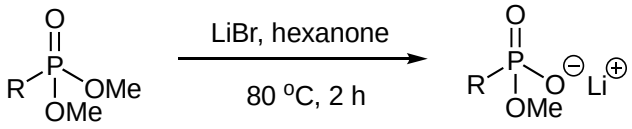
However, sixteen compounds were successfully synthesised and went forward into enzyme assays. Their purity was assessed by NMR. It was predicted that these compounds

would most closely mimic the acid groups of acetyl phosphate **2**, however they may encounter issues in whole cell growth experiments due to a potential inability to cross cell membranes due to their polarity.

Synthesis of putative monomethyl phosphonate ester inhibitors

Following on from the synthesis of the phosphonic acid sodium salts, the monomethyl esters were also targeted as potential inhibitors. The reason for this is their single negative charge and decreased polarity when compared to the phosphonic acids, which may have difficulty crossing cell membranes.

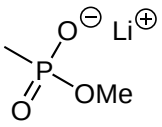
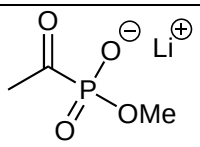
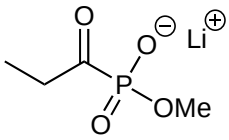
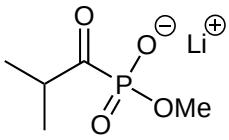
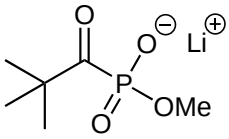
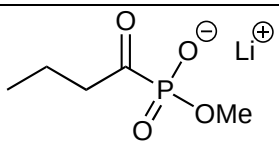
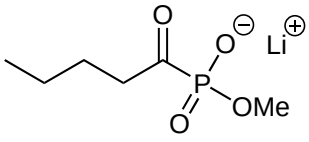
Following procedures similar to the work of Karaman *et al.* and Krawczyk, lithium bromide was used to mono-dealkylate the phosphonate ester compounds synthesised previously.[71, 73] This yielded the monomethyl esters as lithium salts as white solid powders, which were easily handled and purified, therefore were not reacted further to yield the free mono-acid, as in Karaman *et al.*[71]

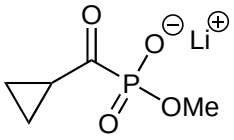
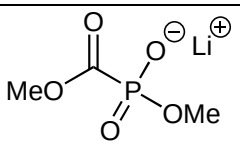
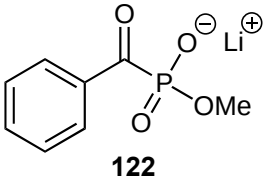
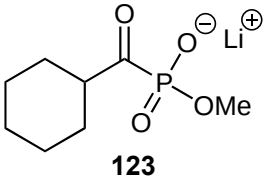
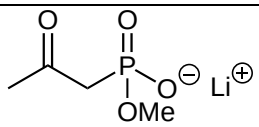
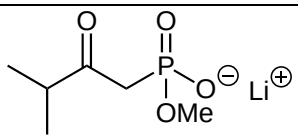
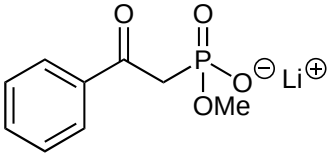


Scheme 36 – The reaction of phosphonate esters with lithium bromide to yield monomethyl ester inhibitor compounds.

Of the 18 phosphonate esters to hand, 17 underwent reaction with lithium bromide to afford the corresponding monomethyl ester. The general reaction scheme for the cleavage can be seen in **Scheme 36**. The reaction afforded a range of monomethyl ester compounds of good purity. The reaction with the α-benzoylphosphonate derivative contained a small impurity which could not be removed during recrystallisation using ethanol/ether or acetone/ether as suggested in the work by Krawczyk.[73] Therefore, the α-benzoylphosphonate monomethyl ester inhibitor was used with this minor impurity in *in vitro* assays.

The successfully synthesised monomethyl ester cleavage products are listed in **Table 5**.

Compound	Yield
 **113**	90 %
 **114**	87 %
 **115**	97 %
 **116**	77 %
 **117**	59 %
 **118**	72 %
 **119**	80 %

Compound	Yield
 **120**	70 %
 **121**	91 %
 **122**	74 %
 **123**	51 %
 **124**	50 %
 **125**	68 %
 **126**	52 %

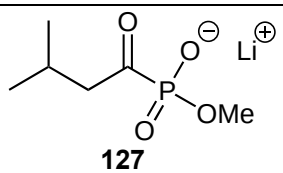
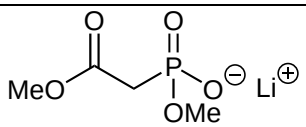
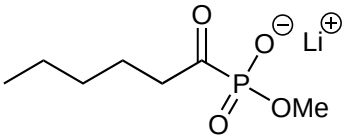
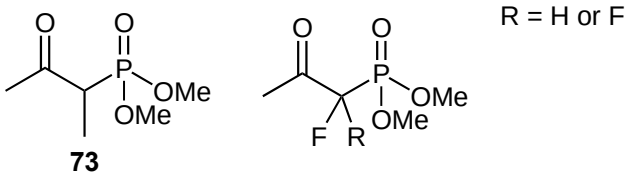
 **127**	67 %	 **129**	76 %
 **128**	90 %		

Table 5 – The successfully synthesised α-carbonyl phosphonate monomethyl ester compounds.

Seventeen compounds were successfully synthesised *via* the procedure shown in **Scheme 36**, affording compounds which were evaluated in enzyme assays. Their purity was assessed by NMR. The monomethyl esters may be key to any whole cell growth work, as they are less polar than the corresponding phosphonic acids, and may be taken up by cells more readily due to this. *C. autoethanogenum* grows at acidic pH, between 4.5 – 5 in a normal growth experiment on fructose, meaning that the monomethyl ester may be protonated in the media. This would be favourable as it would mean that there was not a charged species attempting to pass through the cell membrane.

Summary

Acetate kinase from *C. autoethanogenum* was identified as the target for inhibition in this work, and fifty-three compounds have been generated as putative inhibitors. These compounds have been synthesised using a non-cleavable bioisostere or omission of the oxygen atom linking the two moieties of acetyl phosphate.



73

Figure 31 – The three compounds which would be the most promising future targets for synthesis.

There were also a number of unsuccessful syntheses which could be worked towards further to generate a larger library for testing. Of particular promise would be the fluoro-substituted compounds, which were synthesised but not purified sufficiently, as well as the introduction of a methyl group adjacent to the phosphonate moiety,

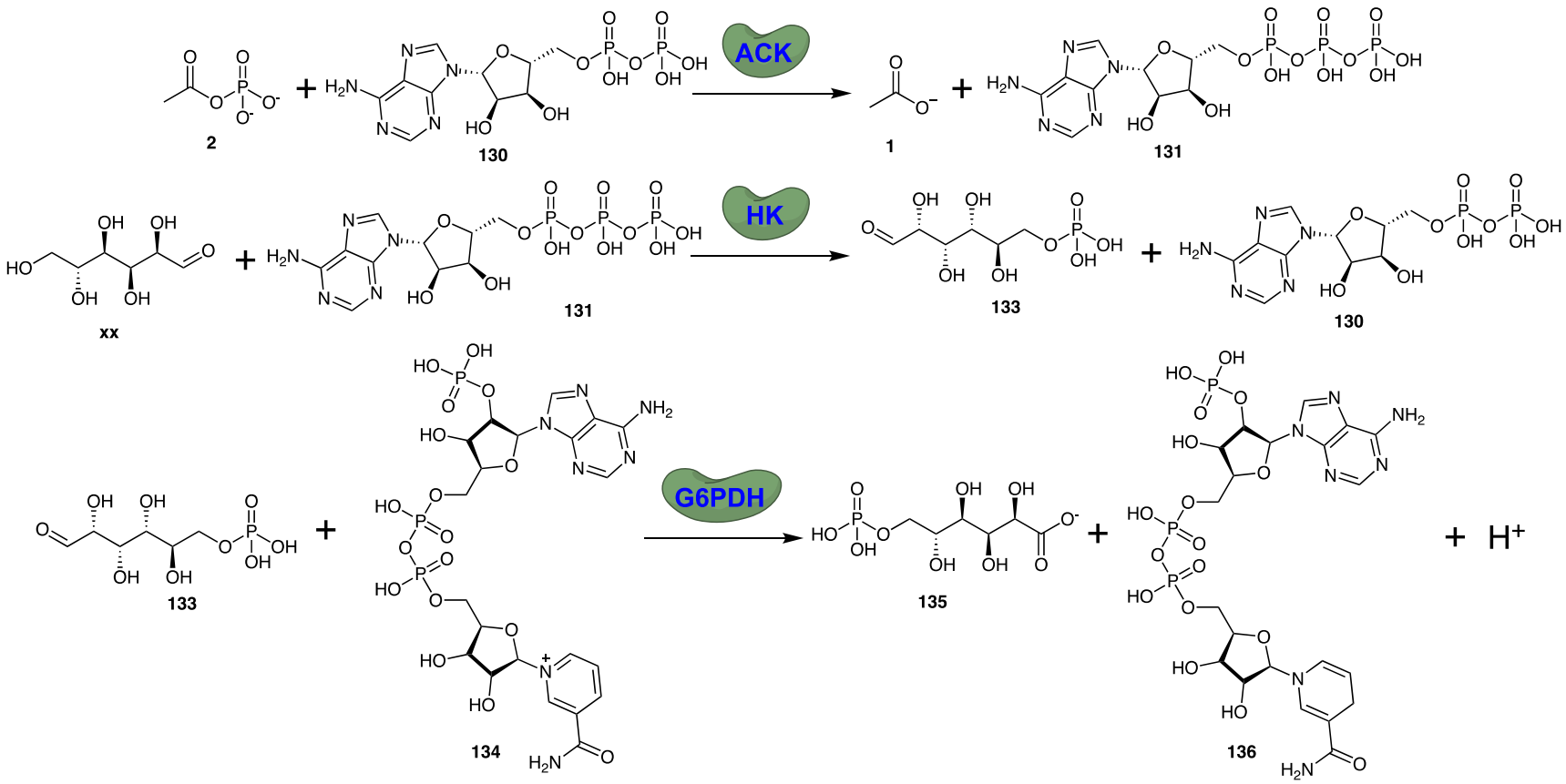
The successfully synthesised compounds were taken forward into enzyme assays, to identify functional groups which interact with acetate kinase for potential further inhibitor

development or whole cell growth. Growth experiments were not completed during this PhD due to time constraints.

Chapter III – Acetate kinase

With the inhibitors successfully synthesised, it was also necessary to express and purify the enzyme acetate kinase from *C. autoethanogenum* to allow isolated enzyme testing *in vitro* before potential testing *in vivo* in the organism. To obtain the protein, the gene sequence was obtained from the known genome sequence, suitable primers designed, which were then amplified by PCR, cloned and sequenced by the Sanger method. pT primers were used (included in experimental) to extract the genomic DNA from *C. autoethanogenum*, in order to amplify the gene encoding the acetate kinase which was then transformed into *E. coli* as a host expression system.

Therefore, primers were designed to amplify the gene encoding acetate kinase, which was then cloned into the pT plasmid which was transformed into *E. coli* BL21 DE3 as the host, chosen for its characteristic ability for overexpression of proteins. A pT vector system was used, which is a derivative of the plasmid pThydA taken from previously successful work by Girbal *et al*.[83] This allowed successful expression and purification *via* C-terminal Strep-Tag® to yield purified acetate kinase for testing.



*Scheme 37 – The enzyme cascade used in the assay by Nakajima et al.[84] It details the conversion of acetyl phosphate **2** to acetate **1**. This can then be visualised using further reactions in a coupled assay to generate NADH **136**, which can be visualised by UV at 340 nm.*

Following on from the expression and purification of the acetate kinase, it was necessary to establish the kinetic parameters of the acetate kinase for comparison with acetate kinase from other organisms. This was performed using an established assay for acetate kinase, developed by Nakajima *et al.*, shown in **Scheme 37**, which measures the product formed at various substrate concentrations.[84]

Following on from kinetic experiments to determine K_M and V_{max}, an inhibition assay was designed modified from the assay by Nakajima *et al.* as shown in **Scheme 37**.[84] This allowed the previously synthesised compounds to be tested against the recombinantly purified enzyme from *C. autoethanogenum*, as well as values for I_{50}, the concentration at which 50 % inhibition is achieved, and K_i, the concentration of inhibitor where reaction rate is half that of V_{max} under saturating substrate conditions. The IC_{50} and K_i values were determined for any compounds which showed promise in initial assays.

Purification of recombinant acetate kinase from *C. autoethanogenum*

The plasmid chosen for expression of acteate kinase was the pT vector system that had been previously used to

overexpress acetate kinase and other enzymes from another acetogen, *Moorella thermoacetica* in *E. coli*.[85] Initially, primers were designed for use with the pT vector system, derived from the pThydA vector plasmid, which incorporates a native Strep-Tag® sequence. The pT vector also contains a thiolase promoter from *C. acetobutylicum*, which upregulates expression.

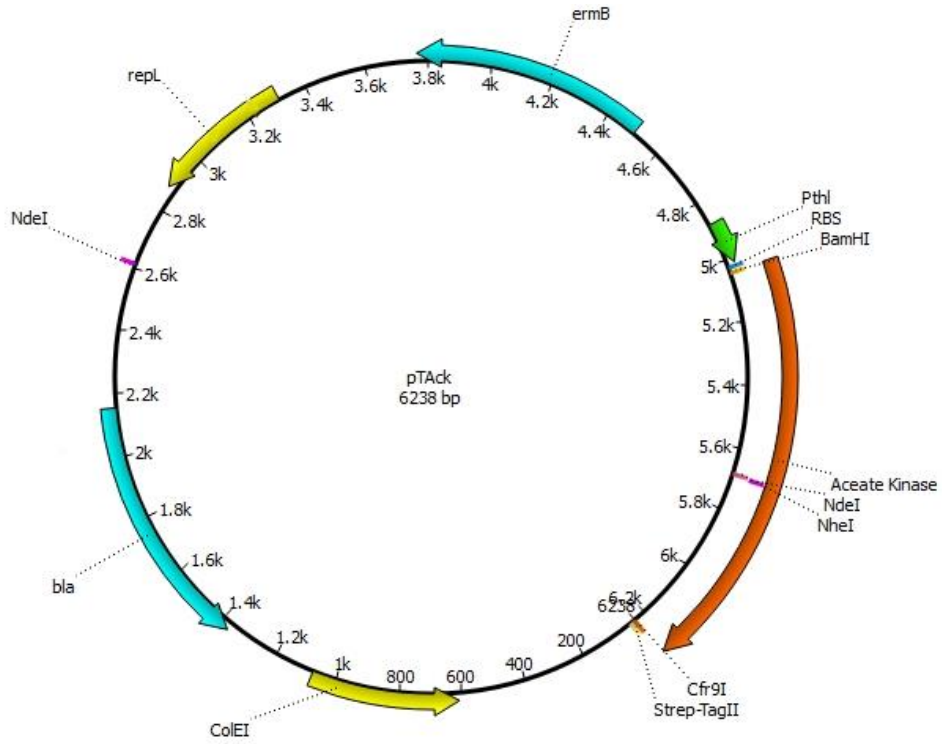


Figure 32 – Plasmid map of pTAck used for the expression of acetate kinase in E. coli BL21 DE3 cells. The map was obtained using Benchling.

Following primer design, genomic DNA was extracted from *C. autoethanogenum* strain DSM 10061 and amplified by PCR. The fragment was purified and underwent restriction digest using the BamHI and Cfr9I restriction enzymes. Subsequently,

the insert was ligated into the pT vector by T4 DNA ligase and the resulting plasmid was transformed into chemically competent *E. coli* BL21 DE3 cells and plated on LB agar plates containing ampicillin for plasmid screening. From this successful growth, colonies were picked and inoculated in liquid LB medium. To ensure that the desired acetate kinase insert had been correctly ligated, a plasmid miniprep was performed, leading to a control digest using fast digest enzymes BamHI and SmaI, which was performed to show that the insert present, as seen in **Figure 33**.

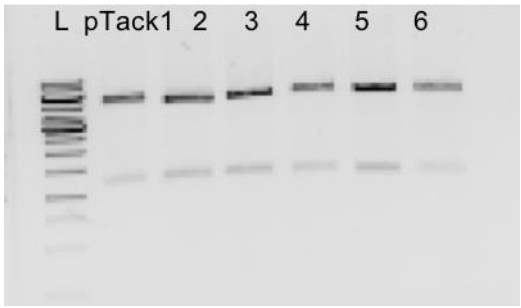


Figure 33 – The gel from the control digest showing bands for the complete plasmid and inserted acetate kinase observed before sequencing was undertaken. L = ladder, pTAck1-6 show the 6 plasmids generated during this work.

Following this control digest an agarose gel was run, showing a band for the backbone plasmid (~5 kb) and the inserted acetate kinase (~1.2 kb), which can be seen in **Figure 33**. To ensure the correct sequence of the insert, the plasmid was sent for sequencing to Source Bioscience, which confirmed the correct sequence had been ligated. Thus, the previously screened colonies could be used to inoculate pre-cultures for

overnight growth, which were then used to inoculate a larger culture, grown to an OD of approximately 1.0.

Protein purification was successfully performed by affinity column using Strep-Tactin sepharose, which binds the Strep-Tag, resulting in purified acetate kinase. The SDS-PAGE and Western Blot in **Figure 34** show successful purification of the protein. To determine an approximate concentration of the acetate kinase, a Bradford Assay was conducted which resulted in a concentration of approximately 0.5 mg/mL as determined by spectrophotometer at 595 nm.

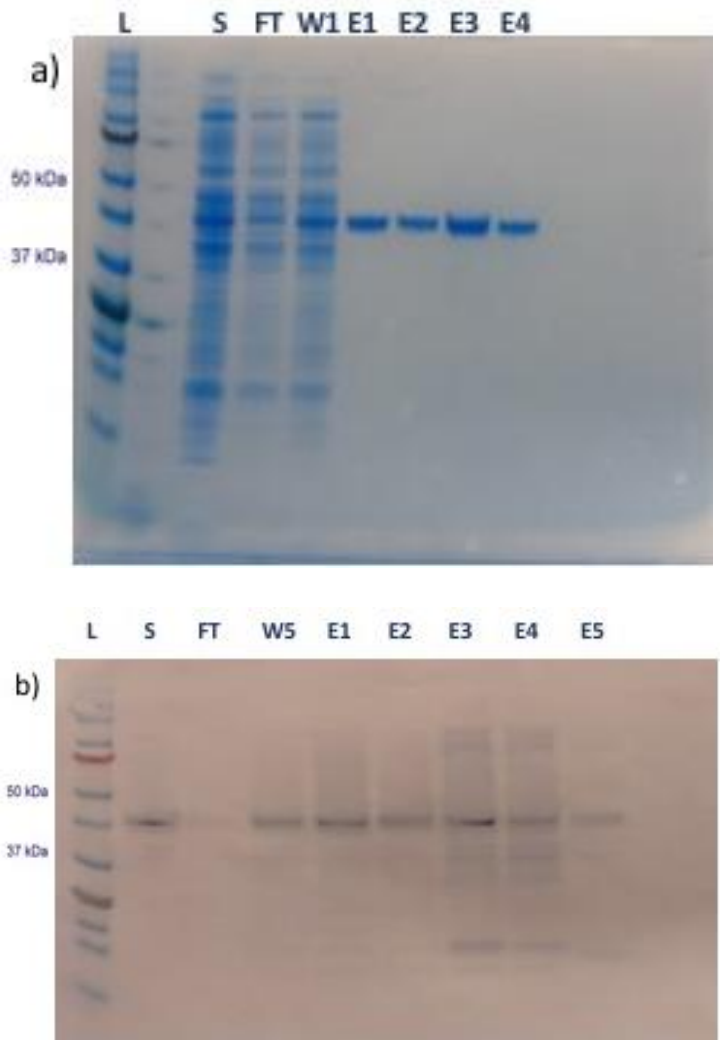


Figure 34 – a) SDS-PAGE and b) Western Blot analysis showing a band corresponding to the expected size of acetate kinase (43 kDa) from C. autoethanogenum. L = ladder; S = supernatant; FT = flowthrough; W1 = wash 1; W5 = wash 5; E1-5 = elutions 1-5.

Figure 34 shows the successful purification of acetate kinase from *E. coli* BL21 DE3 cells. The SDS-PAGE shows the correct band for acetate kinase, which has a molecular weight of 43 kDa. Western Blot analysis confirms the presence of the Strep-Tag which is bound to the acetate kinase protein. The Western Blot also contains other bands as contaminants, which are potentially explained by the Strep-Tag purification system,

which can bind other proteins from the host system that are biotinylated, such as the biotin carboxyl carrier protein (~17 kDa).[86, 87] In addition, the use of milk powder with the Strep-Tag system used can lead to a high background reading, which may also account for some of the impurities seen. However, the presence of a single band on the SDS-PAGE with no contaminants suggests pure protein.

The successful purification of acetate kinase provided a platform for enzyme kinetics to be determined and allowed testing of the inhibitors.

Determination of kinetic parameters of acetate kinase

To determine the kinetic parameters of ACK from *C. autoethanogenum*, the purified enzyme was subjected to the assay used by Nakajima *et al.*[84] This assay operates in the physiological direction in *C. autoethanogenum*, showing the conversion of acetyl phosphate **2** to acetate **1**. This assay couples the turnover of ACK to two other enzymes, hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6PDH), as shown in **Scheme 37.** Following this cascade of enzymatic reactions, NADPH formation can be measured at 340 nm by a

UV spectrophotometer, which can be used to determine the turnover of acetyl phosphate to acetate.

The assay shown in **Scheme 37** allows the concentration of either acetyl phosphate or ADP to be altered, dependent upon which kinetic parameter is being determined. In this case, the concentration of acetyl phosphate would be varied to determine the enzyme affinity for this substrate.

Before acquisition of kinetic parameters at various concentrations, a standard curve was obtained, which allowed the absorbance values at OD_{340} in the enzyme assay to be converted to a corresponding concentration of NADPH. It was necessary to obtain a standard curve of NADPH using various concentrations of NADPH ranging from 0.05 mM to 1.0 mM. This gave a straight line of absorbance at 340 nm versus NADPH concentration, allowing for values obtained from the spectrophotometer as part of the enzyme assays to be converted into concentrations of NADPH formed as part of the enzyme assay. The standard curve can be seen in **Figure 35**.

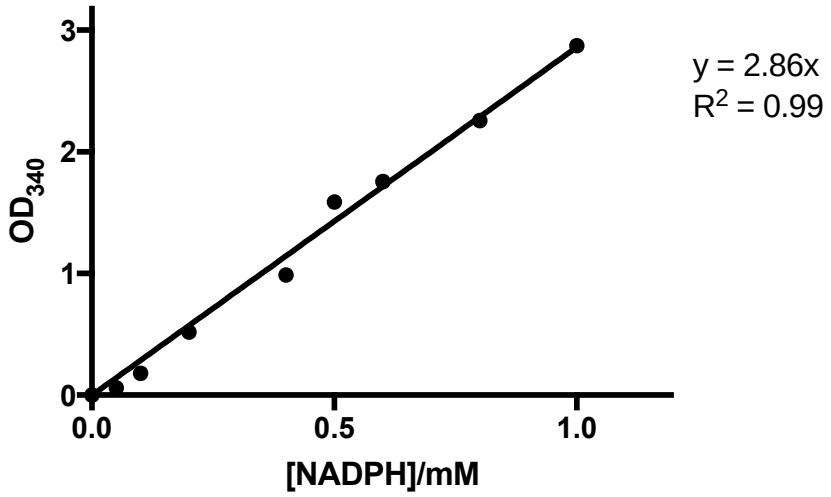


Figure 35 – The standard curve of NADPH concentration versus OD_{340} reading.

With the data for the standard curve obtained, the kinetic assay to determine V_{max} and K_M could be performed, which would also allow inhibition assays to be performed to test the putative inhibitors against acetate kinase to assess their activity.

To establish an effective method for determination of the kinetic parameters of purified acetate kinase, the assay by Nakajima *et al.* was modified by variation of acetyl phosphate **2** concentration, while the concentration of the other components of the assay were kept constant.[84] Initially, the enzyme assays were performed in cuvettes in a UV spectrophotometer with detection set to 340 nm to establish a method which could be adapted to 96-well plates for inhibition studies. Variation of the concentration (S) of acetyl phosphate

2 measured against rate of reaction (v) could then be plotted into a Michaelis-Menten curve, a well-known model of enzyme kinetics developed by Michaelis and Menten which compares the two parameters.[88] This plot allows V_{max} and K_M to be determined, as well as transformed into Lineweaver-Burk, Eadie-Hofstee and Hanes-Woolf plots, which can be seen in **Figure 36**. The different plots allow the kinetic parameters to be determined with different degrees of accuracy and error.

Lineweaver-Burk plots were originally developed to allow kinetic parameter determination to be easily calculated before computer software was widely used to calculate non-linear regression.[89] One criticism of the Lineweaver-Burk plot is the distortion of errors in data, caused by the larger concentrations being grouped together on the graph, determinding the position of the line. More recently, the Eadie-Hofstee and Hanes-Woolf plots have been developed as more accurate methods of kinetic parameter determination because they give equal weighting to each point of the line.[90-92] One drawback to the Eadie-Hofstee plot is that both axes are determined based upon rate of reaction (v), so any error is multiplied. In addition, experimental error can cause poor goodness of fit values for the curve. As with the Eadie-Hofstee

plot, the Hanes-Woolf plot has both axes reliant upon substrate concentration (S), and therefore goodness of fit may not be applicable. It is therefore important for there to be a degree of correlation between all of the plots used to determine kinetic parameters.

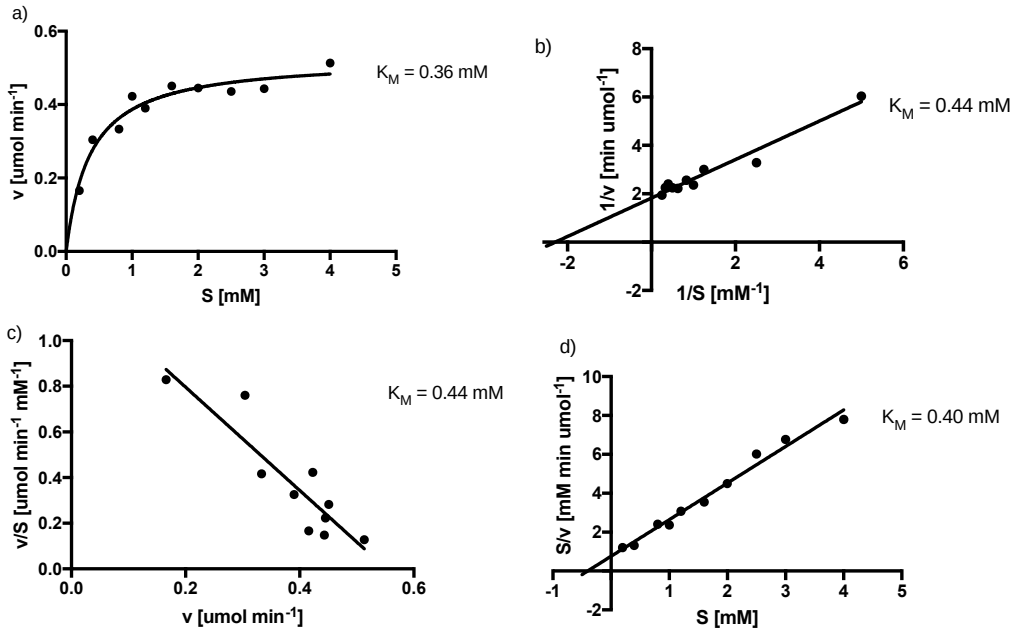


Figure 36 – Graphs showing kinetic parameters of acetate kinase from C. autoethanogenum using the assay by Nakajima et al.: a) Michaelis-Menten curve of rate of reaction (v) plotted against substrate concentration (S); b) Lineweaver-Burk plot showing 1/v plotted against 1/S, where 1/x-intercept gives K_M; c) Eadie-Hofstee plot showing v plotted against v/S, where –(1/slope) gives K_M; d) Hanes-Woolf plot showing S/v plotted against S, where –(x-intercept) gives K_M.[84]

The plots in **Figure 36** suggest that the K_M value of ACK from *C. autoethanogenum* is 0.40 ± 0.04 mM. K_M, the Michaelis constant, is the substrate concentration at half maximum rate of reaction, V_{max}.[88, 93] It is also a good indicator of a start point for testing of compounds for inhibition, to ensure that the concentration of the inhibitors is suitable to yield results. This

is because it is known at which points the enzyme has maximum activity and 50 % activity, so inhibitor concentration can be added accordingly. The K_M value for acetyl phosphate for *M. thermophila* been determined by Gorrell *et al.* to be 0.47 mM.[56] This is comparable to the value determined for *C. autoethanogenum*, which allows confidence in the result.

The Eadie-Hofstee plot shown in Figure 2**Figure 36, Graph (c)** has a significantly lower R^2 value than the other three plots, so the values given may be less accurate. However, in this case the K_M value correlates with the other graphs for the kinetic data.

Kinetic parameters were also determined for ACK using a TECAN plate reader, to ensure that the kinetic data obtained was accurate for inhibition assays, which would also be conducted using 96-well plates. As before, the assay conditions involved varying the concentration of acetyl phosphate, which was injected by the plate reader and readings taken to give kinetic information. As with the cuvette method, this allowed determination of a Michaelis-Menten curve, as well as Lineweaver-Burk, Eadie-Hofstee and Hanes-

Woolf plots. The results of the 96 well-plate assay can be seen in **Figure 37**.

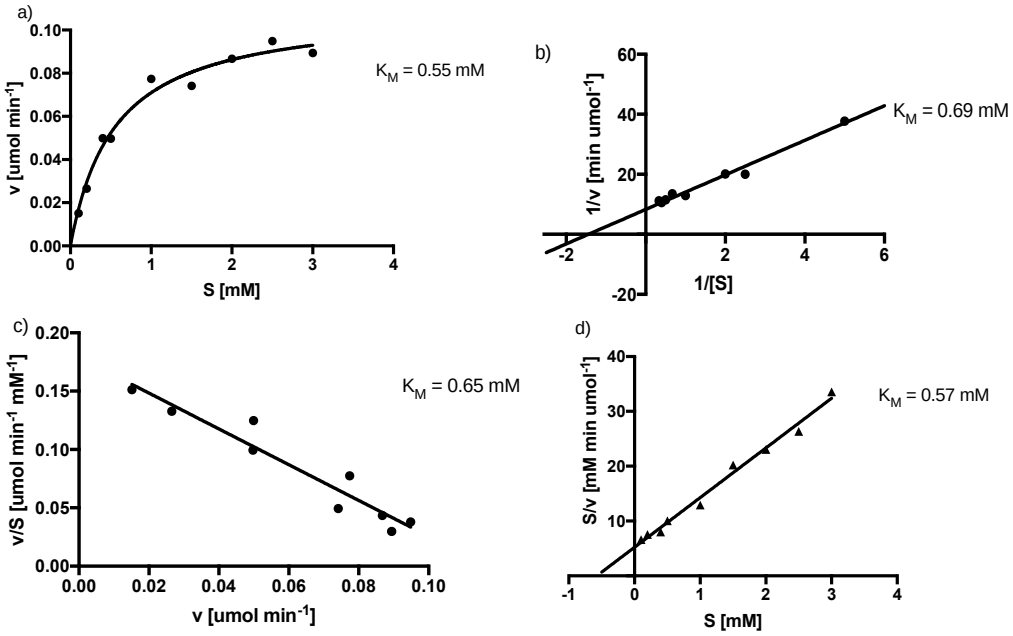


Figure 37 – 96 well-plate measurements of kinetic parameters of acetate kinase from C. autoethanogenum using the assay by Nakajima et al.: a) Michaelis-Menten curve of rate of reaction (v) plotted against substrate concentration (S); b) Lineweaver-Burk plot showing 1/v plotted against 1/S, where 1/x-intercept gives K_M; c) Eadie-Hofstee plot showing v plotted against v/S, where –(1/slope) gives K_M; d) Hanes-Woolf plot showing S/v plotted against S, where –(x-intercept) gives K_M.[84]

The kinetic plots in **Figure 37** give K_M values of 0.62 ± 0.07 mM. These results are slightly higher than the K_M values determined in the cuvettes, but not significantly different of the cuvette assay and published literature values from Gorrell *et al.* for *M. thermophila*.[56] One potential reason for the difference would be the different apparatus used for the kinetic determination, which may give different results.

Determining the kinetic parameters allows for the putative inhibitors to be tested for activity. The kinetic parameters give a measure of the affinity the enzyme has for the substrate, and in the presence of an inhibitor, the change in that value can be measured. The change in the kinetic plots and K_M value in the presence of an inhibitor can indicate the type of inhibition as well as giving the magnitude of the inhibition effect.

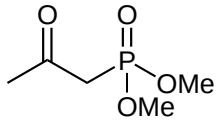
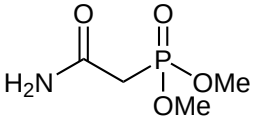
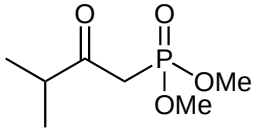
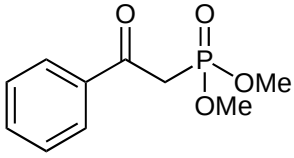
Inhibition of acetate kinase from *C. autoethanogenum*

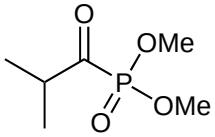
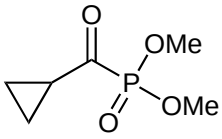
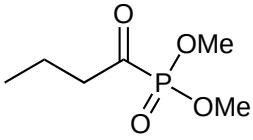
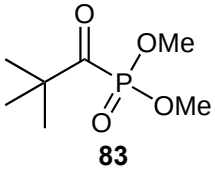
To determine the effectiveness of inhibition of the compounds synthesised, they were screened against the recombinantly purified acetate kinase from *C. autoethanogenum*. This was performed with the same assay used to determine the kinetic parameters, using the modified protocol from Nakajima *et al.*[84] In the assay, which was scaled down to 200 μL for use in 96-well plates, 2 μL water was substituted for 2 μL inhibitor to an effective concentration of 1 mM of inhibitor compound. This was to be used as a starting point for inhibition in the assay to find compounds active at this concentration.

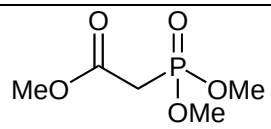
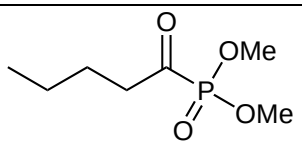
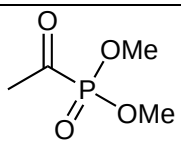
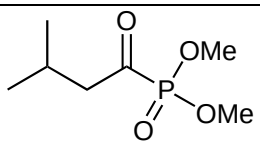
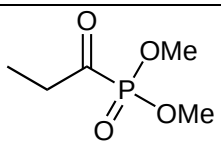
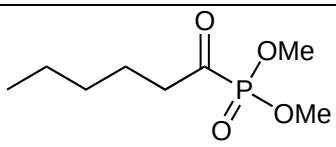
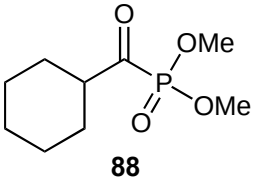
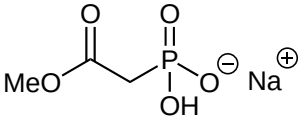
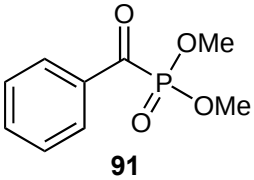
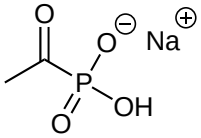
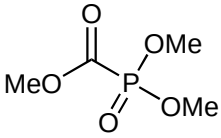
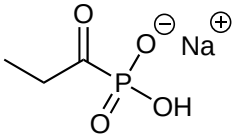
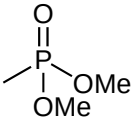
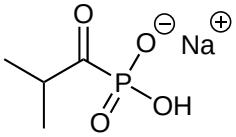
The concentration of acetyl phosphate in the assay was 2.5 mM as in the original assay by Nakajima *et al.* Although this is

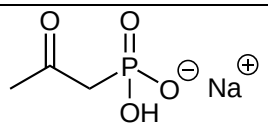
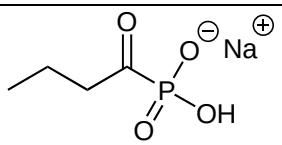
higher than the determined K_M values, it is a saturating concentration of substrate, therefore the enzymatic activity should be close to V_{max}.

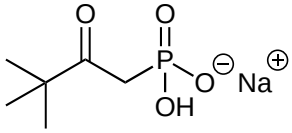
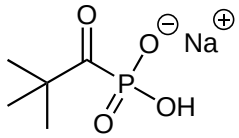
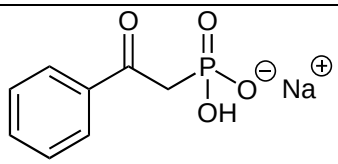
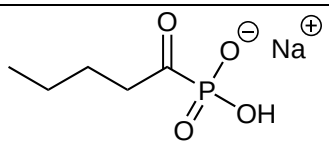
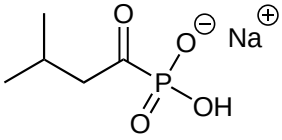
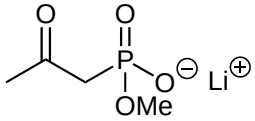
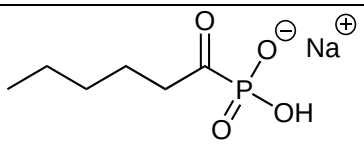
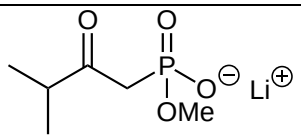
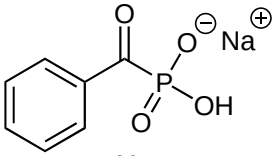
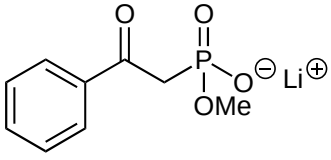
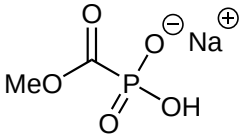
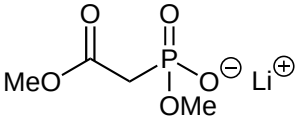
A total of 53 compounds were screened against the purified enzyme, including diethyl pyrocarbonate **139**, a known inhibitor of acetate kinase, shown in *M. thermophila*.[68-70] This allowed for comparison between the synthesised inhibitors and a known inhibitor compound.

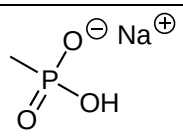
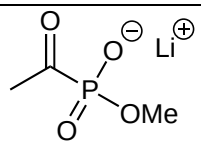
Compound	SAR
 **68**	89 %
 **69**	70 %
 **75**	90 %
 **70**	92 %

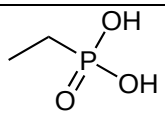
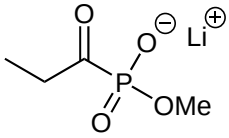
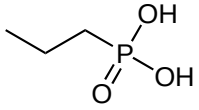
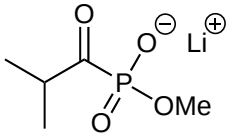
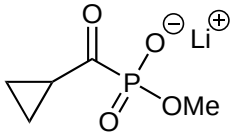
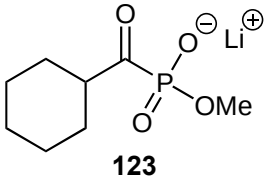
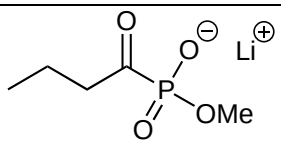
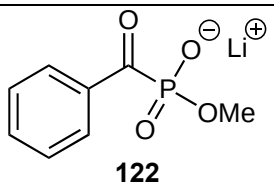
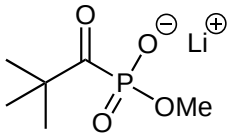
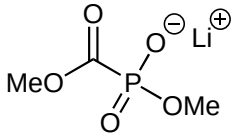
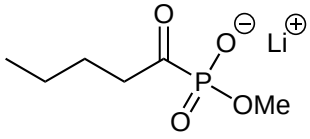
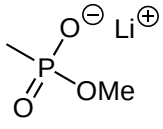
Compound	SAR
 **90**	71 %
 **87**	62 %
 **92**	75 %
 **83**	66 %

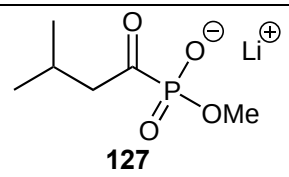
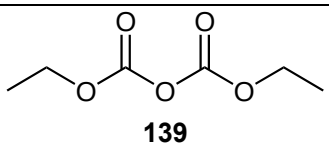
 **144**	78 %	 **85**	42 %
Compound	SAR	Compound	SAR
 **82**	59 %	 **89**	77 %
 **86**	85 %	 **93**	69 %
 **88**	68 %	 **110**	75 %
 **91**	54 %	 **96**	57 %
 **84**	60 %	 **97**	77 %
 **74**	59 %	 **98**	65 %

 **105**	77 %	 **100**	66 %

Compound	SAR	Compound	SAR
 **107**	74 %	 **99**	68 %
 **108**	71 %	 **101**	35 %
 **109**	67 %	 **124**	73 %
 **102**	49 %	 **125**	68 %
 **104**	36 %	 **126**	77 %
 **103**	63 %	 **129**	67 %

 **95**	80 %	 **114**	65 %

Compound	SAR	Compound	SAR
 **137**	47 %	 **115**	81 %
 **138**	46 %	 **116**	77 %
 **120**	95 %	 **123**	67 %
 **118**	89 %	 **122**	63 %
 **117**	65 %	 **121**	68 %
 **119**	63 %	 **113**	74 %

 **127**	68 %
 **139**	0 %

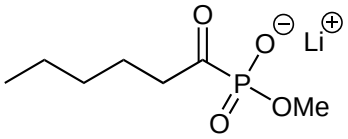
Compound	SAR
 **128**	67 %

Table 6 – The results of the enzyme inhibition assay with inhibitor concentrations at 1 mM, modified from the original protocol by Nakajima et al.; SAR = Specific Activity Remaining, defined as the change in OD_{340} per minute versus the uninhibited enzyme in the presence of 1 mM inhibitor compound.[84]

The specific activity remaining (SAR) of the enzyme in the assay shown in **Table 6** is an averaged result from triplicates of technical triplicates. Values of 50 % activity remaining or less, denoting candidates which had shown good inhibition in multiple assays were taken forward as compounds of interest and tested further. Those taken forward were compounds **101**, **102**, **104**, **137**, **138** and diethyl pyrocarbonate **139**.

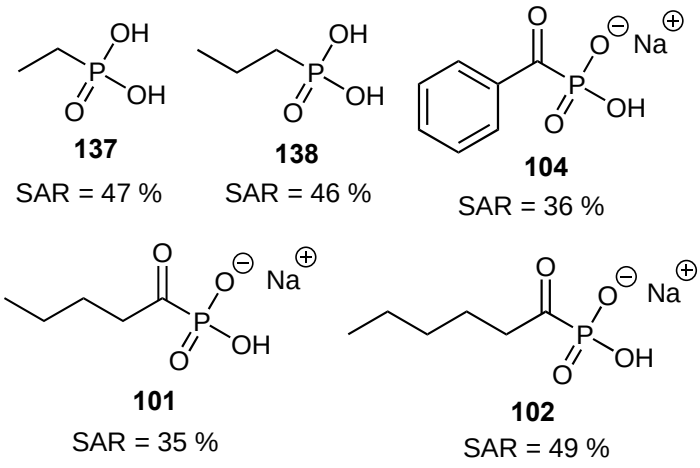


Figure 38 – The compounds of interest for further testing as potential inhibitors of acetate kinase. SAR = Specific activity remaining at 1 mM of inhibitor.

Compounds **137** and **138** are of interest because they lack the carbonyl group of the substrate and other inhibitors that have been screened against the enzyme. In addition, it is notable that they have alkyl chain lengths of 2 and 3 carbons, analogous to the carbon chain length of the library of putative inhibitor compounds tested against the enzyme.

The R-group variation between compounds **101**, **102** and **104** is of interest because they would undergo different interactions within the active site if the inhibition were competitive. While compounds **101** and **102** differ only in carbon chain length, it is interesting to note that the isovaleryl **109** and n-butyryl derivative **100** did not show the same degree of inhibition as these compounds. Compound **104** is of interest because of its phenyl ring R-group, which may

suggest π-π stacking or π-cation interactions if the mode of inhibition was competitive.

Notably, these inhibitor compounds have a different molecular structure and geometry to the substrate, due to the phosphorous atom of the carbonylphosphonic acids bonding directly α- to the carbonyl group.

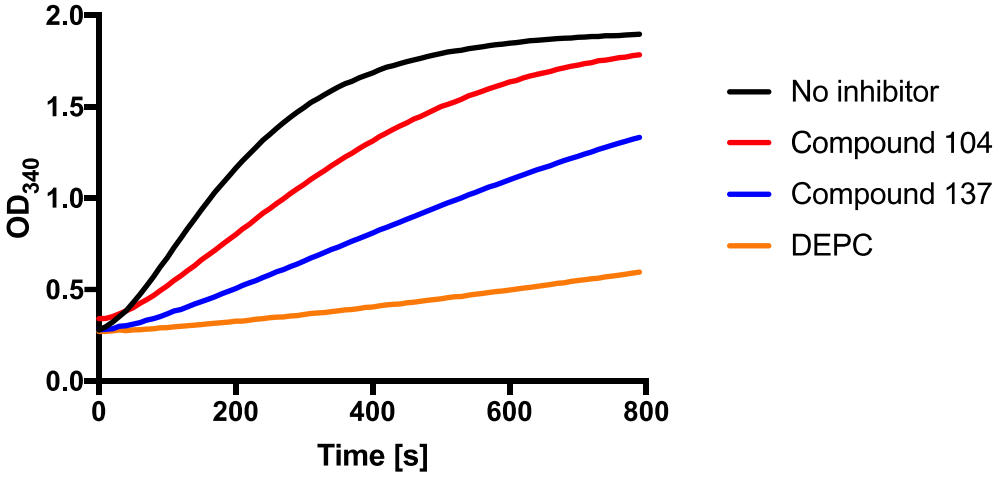


Figure 39 – A graph to show the effect of the inhibitor compounds on the rate of reaction of the purified acetate kinase versus the non-inhibited enzyme.

The graph in **Figure 39** shows the inhibition of acetate kinase by compounds **104**, **137** and diethyl pyrocarbonate **139**. It can be seen that the inhibitors depicted show inhibition of the acetate kinase at 1 mM of inhibitor. The results show a slower rate of turnover of acetyl phosphate **2** to acetate **1** in the presence of compound **104** and compound **137**, evidenced by

the lower gradient of the slope of the curve. For example, compound **104** shows a ~50% increase in time to achieve the same change in OD_{340}.

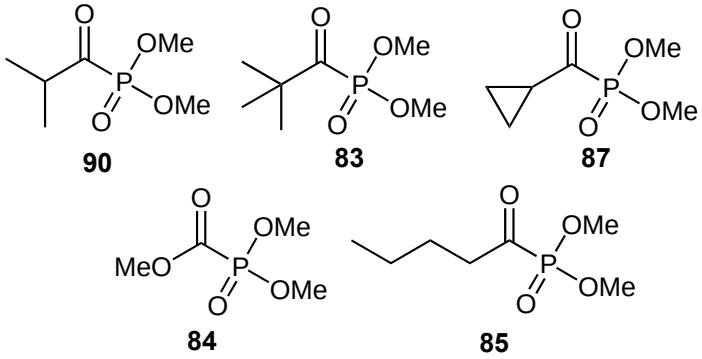


Figure 40 – The compounds which showed activity in a later replicate of the assay, which were seen to have degraded by NMR.

There were a number of compounds which showed activity in a later replicate of the assay; however, upon NMR analysis they were revealed to have undergone degradation in solution over time. These compounds can be seen in **Figure 40**. Interestingly, compound **85** was seen to be undergoing degradation to form compound **101**, which may explain the positive result in the later assay, as this compound **101** was one of the successful candidates shown in **Figure 38**. This also suggests that it could be of interest to test the phosphonate ester and monomethyl derivatives of the successful inhibitors if the compounds were to be tested in whole cell growth experiments. It also highlighted that the

compounds were not stable in solution at -20 °C for longer than 1 month.

The successful compounds from the initial screen, shown in **Figure 38,** were taken forwards into assays to determine IC_{50} and K_i values. IC_{50} is the concentration at which the enzyme can be seen to be inhibited by 50 %, therefore will turn over the substrate to product at 50 % of the rate of the uninhibited enzyme. This is determined using a dose-response curve using a logarithmic value for the concentration of inhibitor, which is demonstrated in **Figure 41**. K_i, the concentration at which the rate of reaction of the enzyme is reduced by 50 %, is measured at various concentrations of substrate in the presence of inhibitor, allowing K_M to be measured and compared between the inhibited and uninhibited enzyme. The change in the slope or intersecting points of the line on the kinetic plots shown previously can also give information towards the type of inhibitor binding that is occurring, whether competitive, non-competitive or uncompetitive.[55]

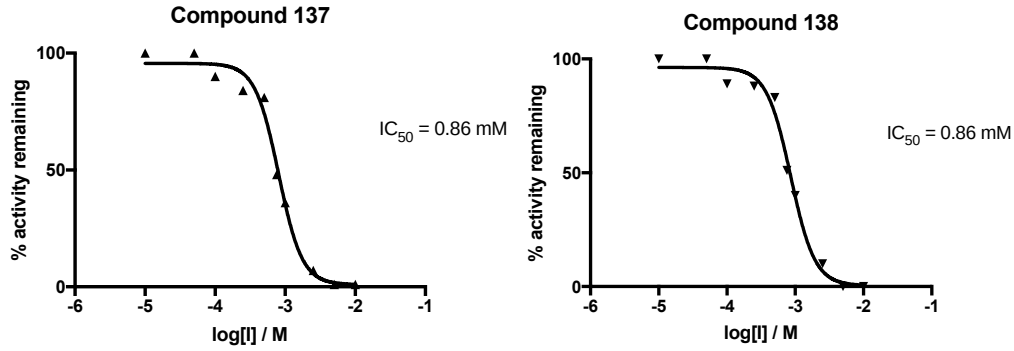


Figure 41 – Example of the graphs to determine the IC_{50} values of the successful compounds.

With the five compounds and diethyl pyrocarbonate **139** found to be active in the screen at a concentration of 1 mM, further assays were undertaken to determine their IC_{50} values. These were performed by plotting dose-response curves, by varying the concentration of the inhibitor in the assay between 0.01 mM and 10 mM, with the other assay substituents remaining constant. An example of the graphs generated to determine IC_{50} can be seen in **Figure 41**. The graphs shown in **Figure 41** are good examples of dose-response curves, showing both maximum and minimum concentrations for inhibition, as well as allowing IC_{50} to be determined from the slope of the curve.

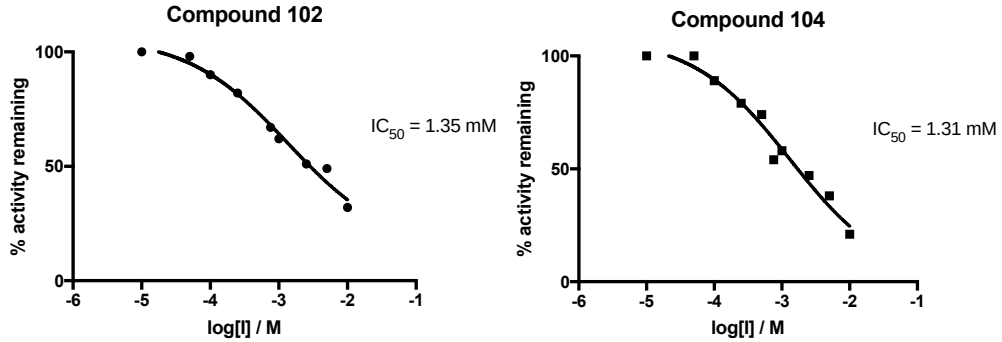
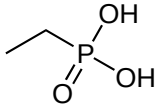
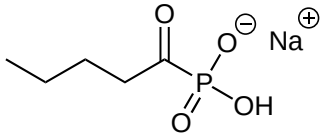


*Figure 42 – The dose-response curves for compounds **102** and **104**, which show minimum inhibition and an inhibition curve, but do not show 100 % inhibition.*

The dose-response curves in **Figure 42** show compounds **102** and **104**, which appear to give the slope of the curve from the minimum inhibition concentration, but did not reach 100 % inhibition. However, the curves generated allowed for an IC_{50} value to be determined computationally using Prism, which can be seen in **Table 7**. Additionally, the dose-response curve for compound **101** did not give a complete inhibition curve, but inhibition appeared to be attenuating at ~60 % inhibition, which suggests that the IC_{50} value determined computationally may not be accurate for this compound.

Compound	IC_{50} / mM
 **137**	0.81

Compound	IC_{50} / mM
 **101**	0.70*

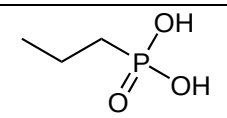
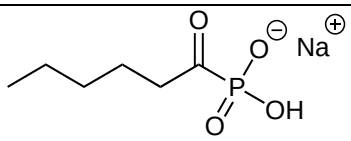
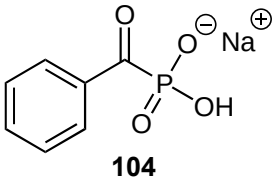
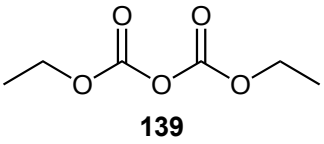
 **138**	0.86	 **102**	1.35
 **104**	1.31	 **139**	0.10

*Table 7 – The IC_{50} values determined for the six compounds showing 50 % or greater inhibition in initial assays. Compound **101** (*) had a dose-response curve attenuating at 60 % inhibition, so may not be accurate.*

The IC_{50} values for compounds **102** and **104** are slightly higher than the 1 mM used in the initial assay. This could be due to a small amount of experimental error in one of the replicates of the assay, causing an increase to the value determined initially. However, the data suggests that all of the compounds with the exception of diethyl pyrocarbonate **139** have an IC_{50} value of ~1 mM.

Following assays to determine IC_{50} values for the compounds, enzyme kinetics assays in the presence of each inhibitor compound were performed to attempt to determine K_i values for the compounds. These assays also suggest the mode of binding for the inhibitor compounds, whether competitive, non-competitive or uncompetitive. This can be performed using the reciprocal Lineweaver-Burk plot, whereby the

characteristics of the line will change according to the type of inhibition.[55]

Assays were performed by varying the concentration of the acetyl phosphate added to the wells, to determine the kinetic parameters over a range of concentrations in the presence of the inhibitor. The final concentration of inhibitor in the wells for compounds **101**, **102** and **104** was 2 mM, and for compounds **137** and **138** was 1 mM. These concentrations were used to ensure sufficient change in rate of reaction for K_i to be determined. Diethyl pyrocarbonate **139** was not successfully assayed to determine a K_i value, because it inhibits irreversibly by acylating residues of the active site.[69]

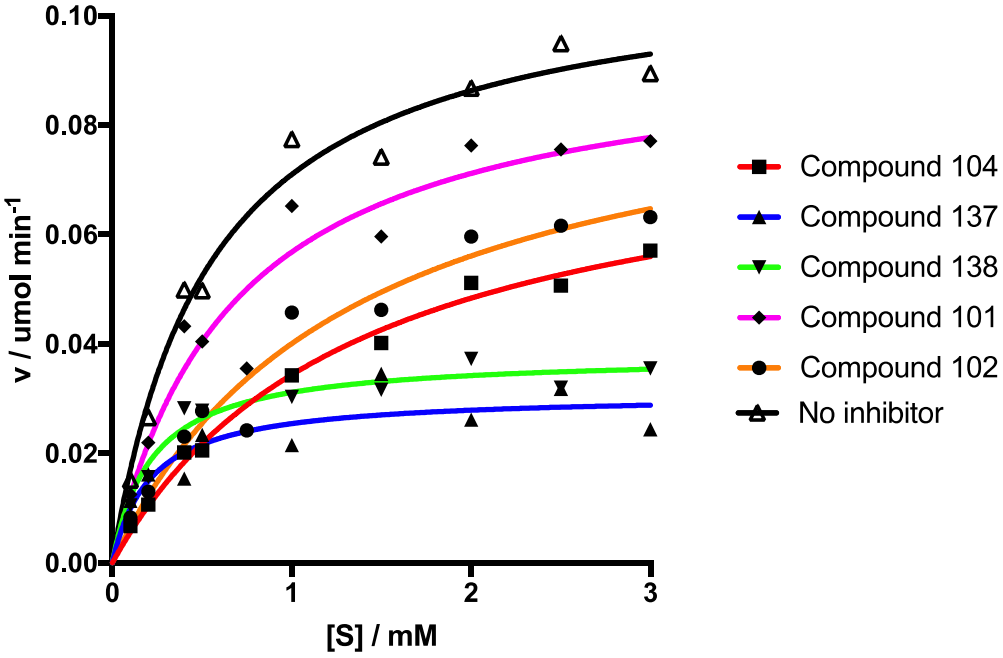


Figure 43 – A graph showing the rates of reaction of the acetate kinase enzyme in the presence of the inhibitor compounds at varying substrate concentrations.

Figure 43 shows the rate of reaction of the enzyme in the presence of inhibitors. This illustrates that the inhibitors are effective and reduce the rate of reaction in the assay. Compounds **101** and **102** showed lower activity than expected from previous assays at the concentration added, however, they did exhibit a degree of inhibition of acetate kinase so the results are included.

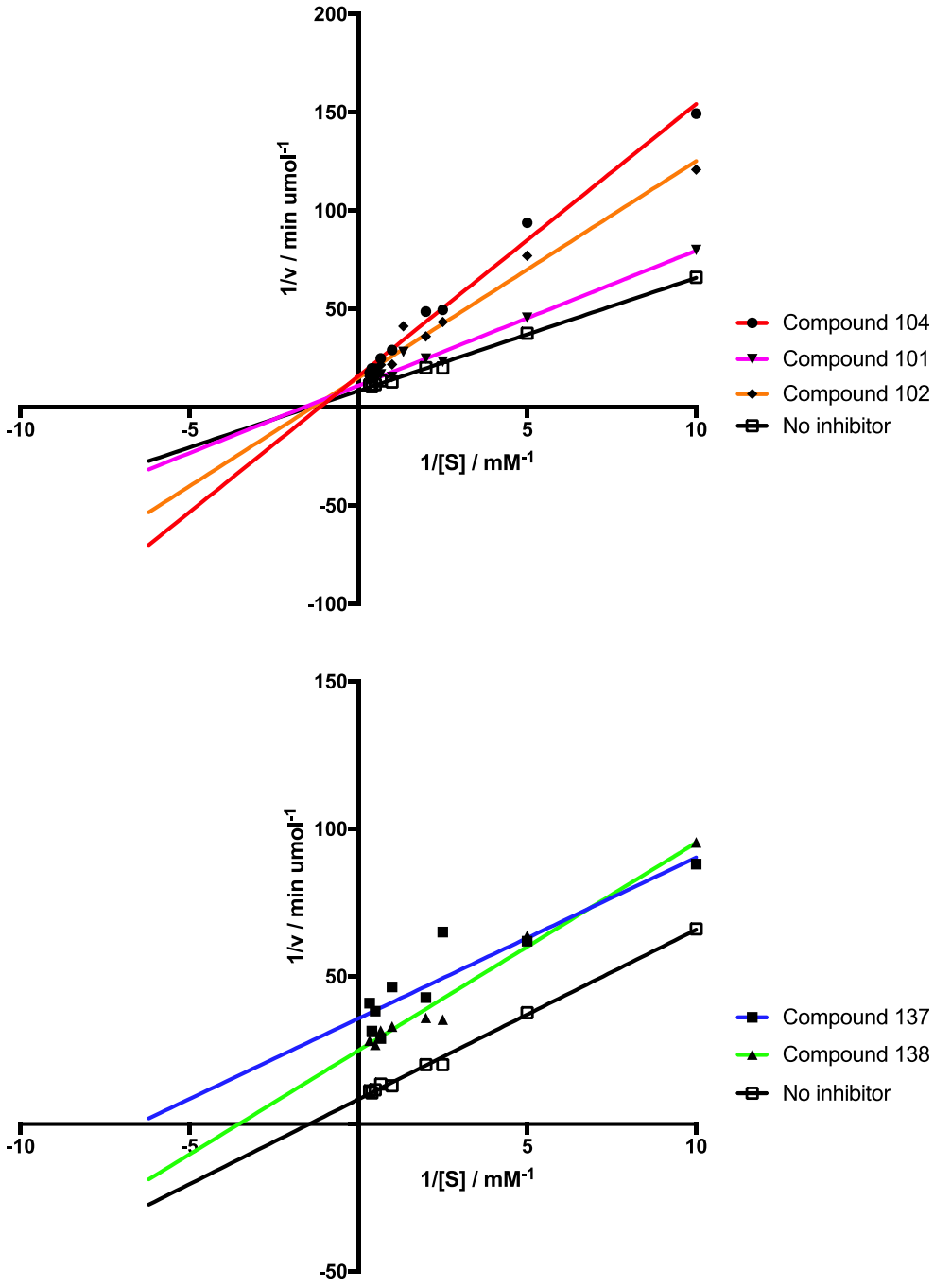


Figure 44 – The double reciprocal Lineweaver-Burk plots showing the kinetic activity of acetate kinase in the presence of inhibitor compounds.

Figure 44 shows the double reciprocal Lineweaver-Burk plot of the data obtained for enzyme kinetics in the presence of the

inhibitor compounds of the relevant concentration of 1 mM and 2 mM.

The top graph shows that compounds **101**, **102** and **104** appear to be primarily non-competitive inhibitors of acetate kinase, which was given by the slopes of the line intersecting the y-axis at different points, giving a different V_{max} value whilst maintaining their K_M. There was a small change in K_M, suggesting there could be a small amount of mixed inhibition between competitive and non-competitive inhibition. Competitive inhibition means that the inhibitor is blocking the substrate's access to the active site, which may be overcome at increased substrate concentrations. Non-competitive inhibition means that the compounds may be binding to the enzyme at an allosteric site, affecting the conformation of the enzyme and reducing its rate of reaction. This is not overcome by increased substrate concentration, as V_{max} is lower for the enzyme in the presence of these compounds.[55] Due to the small change in K_M, it was assumed that there was a degree of mixed inhibition with these compounds, so two K_i values were calculated.

The bottom graph shows that compounds **137** and **138** appeared to be inhibiting uncompetitively, evidenced by a decreased K_M and V_{max} value. This means that these compounds are binding to the enzyme-substrate complex, which is not unusual in reactions containing more than one substrate, as in this one, where ADP is also required by the enzyme for the turnover of acetyl phosphate **2** to acetate **1**.[55]

With the mode of inhibition for each inhibitor compound established, it was possible to determine K_M and K'_M mathematically using equations based upon the graph in **Figure 44**. From this graph, it is possible to calculate the K_i values for the compounds **101**, **102** and **104**, using the equations below.

For non-competitive inhibition:

$$V'_{max} = \frac{V_{max}}{\left(1 + \frac{[I]}{K_i}\right)}$$

V'_{max} – V_{max} in the presence of inhibitor

For competitive inhibition:

$$K'_M = K_M\left(1 + \frac{[I]}{K_i}\right)$$

K'_M – K_M in the presence of inhibitor

For compounds **137** and **138** which show uncompetitive inhibition, the following equation is used:

$$V'_{max} = \frac{V_{max}}{\left(1 + \frac{[I]}{K_i}\right)} \text{ or } K'_M = \frac{K_M}{\left(1 + \frac{[I]}{K_i}\right)}$$

The K_i values for all of the inhibitor compounds can be seen in **Table 8**.

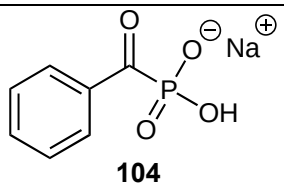
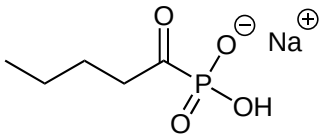
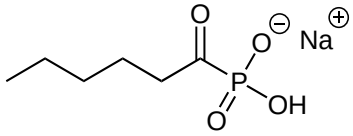
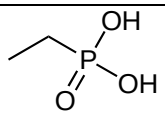
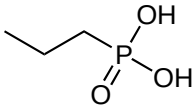
Compound	Non-competitive K_i / mM	Competitive K_i / mM
 **104**	2.22	7.67
 **101**	6.25	20.0
 **102**	2.52	33.3
	Uncompetitive K_i / mM	
 **137**	0.282	
 **138**	0.692	

Table 8 – The K_i values calculated from the assays used to determine the kinetics of the enzyme in the presence of inhibitor compounds.

Compound **101** exhibited very little inhibition during the assays to determine K_i. This can be seen in **Figure 43** which shows only a small change in the curve for the rate of reaction, at a concentration greater than was determined in the IC_{50} assays. However as previously noted, the data for the IC_{50} assay for compound **101** was poor, showing attenuation at ~60 % inhibition. This suggests that the value calculated computationally for IC_{50} for this compound was inaccurate.

The values for K_i are higher than those given for IC_{50} for compounds **101**, **102**, and **104**, converse to the rules of the Cheng-Prusoff equation, which calculates the correlation between the two.[94]

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}}$$

Equation 4 – The Cheng-Prusoff equation, which can be used to calculate the correlation between IC_{50} and K_i.[94]

The Cheng-Prusoff equation, shown in **Equation 4**, states that IC_{50} must always be equal to or larger than K_i. However, it does not apply for mixed inhibition, where the two cannot be correlated which may explain the disparity in the two results.[95] Compounds **101**, **102** and **104** appear to inhibit acetate kinase via mixed inhibition which explains the reason that the values for IC_{50} and K_i cannot be correlated.

Compounds **137** and **138** appear to inhibit acetate kinase by uncompetitive inhibition so can be correlated by the Cheng-Prusoff equation, but are dependent upon the concentration of the substrates due to the assay being a two-substrate assay, where the inhibitor is competing with one of those substrates. The values obtained for K_i are lower than those for IC_{50}, which satisfy the Cheng-Prusoff equation.

The data obtained in the assays suggests that the most potent inhibitors tested, with the exception of diethyl pyrocarbonate **139**, were ethylphosphonic acid **137** and propylphosphonic acid **138**. These compounds had K_i values of 0.28 mM and 0.69 mM respectively, as well as having IC_{50} values of 0.81 mM and 0.86 mM and have been shown to bind uncompetitively. As uncompetitive inhibition is more common in assays which involve more than one substrate, it is possible that these inhibitors are interacting with ADP, or its binding site.[55]

Of the compounds which were designed as part of the two classes of ligand based inhibitors, benzoylphosphonic acid **104** was the most potent, binding with mixed inhibition, with K_i

values of 2.22 mM (non-competitive) and 7.67 mM (competitive) and an IC_{50} value of 1.31 mM.

It is interesting to note that the compounds were designed as competitive inhibitors, but have not shown a significant amount of competitive inhibition. This was unexpected due to the ligand-based design of the inhibitors being based upon the substrate, acetyl phosphate **2**.

The synthesised compounds are less potent than diethyl pyrocarbonate **139**, which was previously known to inhibit acetate kinase and has a calculated IC_{50} value of 0.1 mM.[68-70] However, diethyl pyrocarbonate **139** cannot be used in liquid media over an extended period of time, particularly with *C. autoethanogenum* which requires a mildly acidic environment for growth, as it is susceptible to degradation in water, which was observed during the course of the assays. In addition, it is non-selective for acetate kinase, as it is an inhibitor for a number of enzymes.

Summary

The gene for acetate kinase from *C. autoethanogenum* was successfully cloned into a pT vector system containing a native C-terminal Strep-Tag, transformed into *E. coli* BL21 DE3,

expressed and purified *via* a Strep-Tactin purification system. SDS-PAGE and Western Blot analysis confirmed successful purification and enzyme concentration was determined to be >0.5 mg/mL.

To determine kinetic parameters of acetate kinase, an assay developed by Nakajima *et al.* was modified to incorporate varying concentrations of acetyl phosphate **2**.[84] This modified assay allowed the kinetic parameters K_M and V_{max} to be determined, in both cuvette and 96 well-plate assays.

The putative inhibitor compounds were tested against acetate kinase in the modified assay from Nakajima *et al.*[84] Initial assays were performed at 1 mM inhibitor against 2.5 mM substrate. Five compounds were shown to inhibit the enzyme at this concentration by 50% or more, and were taken forward for further assays to determine their IC_{50} and K_i values. Two compounds were identified to inhibit acetate kinase by uncompetitive inhibition and three further compounds were shown to inhibit it by mixed inhibition. The most potent, ethylphosphonic acid **137**, had a K_i value of 0.28 mM and an IC_{50} value of 0.81 mM.

3.0 Conclusions

The two-fold aims of the project were the design and synthesis of isotopically labelled metabolites of the native pathways in *C. autoethanogenum*, which was performed by synthesis of putative mimic compounds. The second aim was inhibitor design and synthesis of small molecule inhibitors to test against the enzymes of this pathway, of which acetate kinase was chosen.

Initial isotopic labelling studies were performed using SNAC thioesters, putative mimics of acetyl-CoA **3** with increased hydrophobicity for cellular uptake. These have previously been tested as part of polyketide biosynthesis studies and successfully recognised and incorporated into the enzymes of this pathway.[47] 2-$^{2}H_3$-acetyl-SNAC **52** and 1-^{13}C-acetyl-SNAC **53** were synthesised and tested in whole cell growth experiments, where timepoints were taken and the supernatant samples analysed. The results of the NMR analysis showed no uptake of the SNAC thioesters into *C. autoethanogenum*, suggesting that they may not be suitable for whole cell growth in clostridia.

Following on from the unsuccessful growth experiments with SNAC thioesters, and work on pantothenate uptake experiments by Annan *et al.*,[51] a further mimic based upon pantothenate was devised and synthesised with a ^{13}C isotopic label. 1-^{13}C-acetyl pantetheine **54** was synthesised, but was not tested in whole cell growth experiments due to time constraints.

The logical next step in this work is to conduct growth experiments with 1-^{13}C-acetyl pantetheine **54** to determine its effectiveness as an acetyl-CoA **3** building block as part of its biosynthesis. This could be performed by using a modified vitamin solution omitting pantothenate from the media, replacing it with the labelled compound.

C. autoethanogenum growth experiments with isotopically labelled fumaric acid **65** have been conducted and two pathways have been discovered to be responsible for the conversion of fumaric acid **65** to acetate **1**. The first of these pathways is the incomplete TCA cycle, converting fumaric acid **65** to acetate **1**, *via* maleic acid **64**, oxaloacetic acid **63**, pyruvate **4**, acetyl-CoA **3** and acetyl phosphate **2**. Additionally, a micromolar concentration of alanine containing a ^{13}C label

was detected in the growth experiments, suggesting conversion of ^{13}C-fumaric acid to pyruvate **4**, where alanine is a downstream product generated *via* transaminase reaction. The second pathway active in fumarate metabolism involves re-assimilation of carbon dioxide produced from the decarboxylation of oxaloacetic acid **63** during conversion to pyruvate **4**, making use of the reducing equivalent produced by the first route. This was shown by increased ^{12}C-acetate signals during growth experiments with 2,3-$^{13}C_2$-fumaric acid **66**, and later proved by the differences in growth experiments with $^{13}C_4$-fumaric acid **67**.

Acetate kinase was identified as the target for enzyme inhibition because it was shown by Cooksley *et al.* that the genetic knockout of the *ack* gene led to reduced survivability in *C. acetobutylicum*. Therefore, transient inhibition may be more a more useful regulation of acetate kinase activity because it would have a knockdown effect which may drive metabolism towards higher value chemical products, while maintaining the survival of *C. autoethanogenum*.

To this end, fifty-three compounds were synthesised for testing against acetate kinase. These were designed based

upon the ligand, acetyl phosphate **2**. Two families were chosen to build the library of compounds, α-carbonylphosphonates and β-ketophosphonates. These were designed to remove a labile bond present in the substrate and to have different functional R-groups to explore the interactions of the unknown active site.

Acetate kinase was recombinantly expressed in *E. coli* BL21 (DE3) using a pT plasmid taken from Girbal *et al.*[83] Purification was achieved using a Strep-Tag system native in the pT plasmid, and yields of protein were determined to be >0.5 mg/mL by Bradford assay. Following purification, enzyme kinetics were determined using a modified assay from Nakajima *et al.*[84] K_M for acetate kinase from *C. autoethanogenum* was determined to be 0.4 ± 0.04 mM in cuvette assays and later determined to be 0.62 ± 0.07 mM in well plate assays. These values are comparable to published values for acetate kinase from *M. thermophila*, which has a similarity of 63 % and a published crystal structure.[57, 68, 69]

With the enzyme purified and inhibitors synthesised, the assay was modified to incorporate inhibitors and the enzyme was tested against fifty-three compounds. Five compounds showed

>50 % inhibition at 1 mM in an initial screen, and compounds **101**, **102**, **104**, **137**, **138** and diethyl pyrocarbonate **139** were taken forward for IC_{50} determination. IC_{50} values were determined and compounds **137** and **138** were found to be those with the lowest IC_{50} values of 0.81 and 0.86 mM respectively. K_i experiments were conducted on the five compounds to determine the most potent inhibitor and to determine the mode of inhibition.

This K_i assays revealed that compounds **101**, **102** and **104** were mixed inhibitors and **137** and **138** were uncompetitive inhibitors. Ethylphosphonic acid **137** was shown to be the most potent overall inhibitor with the lowest IC_{50} value of 0.81 and K_i value of 0.28 mM. Of the mixed inhibitors, the benzoylphosphonic acid sodium salt **104** was the most potent inhibitor.

Further work on the inhibition of acetate kinase may be to investigate the uncompetitive inhibitors and their mechanism of action. Uncompetitive inhibition is more common in two substrate enzyme assays such as the one used here, so the inhibitors may be interacting with the other substrate, ADP **130** or its binding site. Additionally, because of the low

relative potency of the inhibitor compounds, which showed millimolar inhibition, further synthesis of derivatives which were not completed during this project may be of use. These methods included reduction, thionation, reductive amination and fluorination.

4.0 Experimental

General procedures

General chemical procedures

Commercial reagent grade solvents from Fischer Scientific® were used for all reactions. Anhydrous solvents were purchased from Fischer Scientific® or Sigma Aldrich®. All reagents used were from Sigma Aldrich®, VWR International or Fisher Scientific®.

High-resolution mass spectra were recorded on a Bruker™ microTOF, an orthogonal Time of Flight instrument with an electro spray ionisation (ESI, positive and negative ion) source. Values of mass-to-charge ratio, denoted m/z, given to four decimal places. The mass of counter ions; H^+, $^+NH_4$ and Na^+, Im^+ ($C_3H_5N_2$); 1.0078, 18.0338, 22.9898 and 69.0453 respectively.

IR spectroscopy was obtained on a Nicolet FT-IR spectrometer using sodium chloride discs, with samples made up as nujol mulls, or liquid films using dichloromethane if necessary.

Metlting points were obtained from a Stuart Scientific Melting point apparatus and are given to the nearest temperature as a range of temperatures.

General NMR procedures

^{1}H NMR (400/500 MHz), ^{13}C NMR (100/125 MHz) spectra were recorded on Bruker™ AV400, AV(III)400, AV(III)500 or DPX400 spectrometers using Wilmad 507-PP-7 or 535-PP-7 NMR tubes. Chemical shifts, δ, given in parts per million (ppm) and *J* values in Hertz (Hz). Multiplets were designated by the following notations: s, singlet; d, doublet; t, triplet; q, quartet; quin, quintuplet; sex, sextet; sept, septet; m, multiplet; dd, double doublet; dt, double triplet; tt, triple triplet; d.sep, double septet; appt is used to denote peaks which appear to be a certain multiplicity but first order theory suggests they should be a different multiplicity.

^{1}H, ^{13}C, ^{31}P NMR procedures

Spectra were obtained using Bruker DPX400, AV(III)400, AV(III)400HD, or AV(III)500 MHz spectrometers. Samples of compound (~20 mg) were dissolved in an appropriate deuterated solvent (600 μL) and analysed according to the automated parameters. Solvents used were $CDCl_3$, D_2O, MeOD, DMSO-d_6 or $(CD_3)CO$.

Quantitative ^{1}H NMR procedure for isotopic labelling studies

^{1}H spectra with and without ^{13}C decoupling were obtained using a Bruker AV(III)800 MHz spectrometer. To biological supernatant samples (540 µL) were added deuterium oxide (55 µL) and DSS (5 µL). For whole cell growth experiments in the NMR spectrometer, deuterium oxide (55 µL) and DSS (5 µL) were added to PETC-MES culture media (540 µL) containing *C. autoethanogenum* from pre-cultures (50 µL).

Quantitative NMR experiments were achieved by manual solvent suppression, using a NOESY 1D experiment with a D1 delay of 15 s to ensure complete relaxation of metabolites for determination of concentration. This was calculated by a T1 relaxation experiment performed on the metabolites in culture, and D1 was used as 5*T1 for the metabolite with the largest T1 value.

Quantitative $^{1}H,^{13}C$ 2D NMR procedure for isotopic labelling studies

$^{1}H,^{13}C$-HSQC experiments were obtained using a Bruker AV(III)800 MHz spectrometer. To biological supernatant samples (540 µL) were added deuterium oxide (55 µL) and

DSS (5 µL). For whole cell growth experiments in the NMR spectrometer, deuterium oxide (55 µL) and DSS (5 µL) were added to PETC-MES culture media (540 µL) containing *C. autoethanogenum* from pre-cultures (50 µL). The experiments were performed with parameters matching those of the quantitative 1D experiments.

^{13}C 1D NMR procedure for isotopic labelling studies

^{13}C NMR experiments were obtained using a Bruker AV(III)800 MHz spectrometer. To biological supernatant samples (540 µL) were added deuterium oxide (55 µL) and DSS (5 µL). For whole cell growth experiments in the NMR spectrometer, deuterium oxide (55 µL) and DSS (5 µL) were added to PETC-MES culture media (540 µL) containing *C. autoethanogenum* from pre-cultures (50 µL).

Additionally, for SNAC thioester experiments, spectral width was increased to 210 ppm to ensure the ^{13}C-labelled acetyl SNAC peak was incorporated into the spectrum.

General biological procedures

The bacterial strain used was *Escherichia coli* BL21 DE3 as an intermediate host for transformation of genomic DNA from *C. autoethanogenum* DSM 10061. Cultures were grown in liquid or solid (1.5 % agar) media. *E. coli* strains were cultivated aerobically at 37 °C in LB medium. Chemically competent *E. coli* cells were grown previously and stored at -80 °C. *C. autoethanogenum* was grown in an anaerobic cabinet (MG1000 Anaerobic Work Station, Don Whitley Scientific) containing an atmosphere consisting of nitrogen (80 %), hydrogen (10 %), carbon dioxide (10 %), or externally by use of anaerobically gassed and sealed Hungate tubes or anaerobic culture bottles sealed with aluminium caps (Dunnlab). For cultivation of *C. autoethanogenum,* optimised *C. ljungdahlii* medium (ATCC medium number 1754) was used. In 1 litre: sodium formate (1 g), MES (5 g), NaOH (5 M, 2 mL), 50x macronutrient solution (25 mL), 1000x acidic trace element solution (1 mL), 1000x base trace element solution (1 mL) and 1000x vitamin solution (1 mL).

For heterotrophic growth, yeast extract (2 g L^{-1}) and fructose (40 mM) were added. The addition of fructose was conducted from an anaerobic and sterile filtered stock solution after

autoclaving. The medium, as well as the supplements, were prepared under anaerobic conditions and purged with nitrogen before autoclaving. For reduction of the medium, L-cysteine (4 %) was added as reducing agent. The growth temperature for *C. autoethanogenum* was 37 °C and the starting pH was 5.8 to 6.0. Ampicillin was used for verification of positive transformation in *E. coli*. Lysozyme was provided by Affymetrix.

For well plate assays, a TECAN Infinite M1000 Pro was used with Corning 96 well-plates (Black plate, clear bottom, polystyrene) for both kinetic parameter determination and inhibition studies.

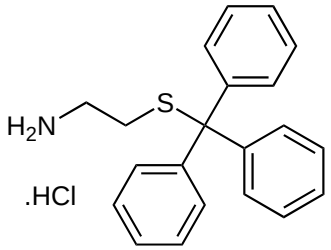
Initial expression of acetate kinase was done in collaboration with Ronja Breitkopf (Synthetic Biology Research Centre, Nottingham).

Whole cell lysis procedure

Cell culture (1 mL) was centrifuged (10,000 *g* for 5 min) at 4 ^{o}C. The supernatant was removed and frozen in liquid N_2. Methanol:water (100 μL, 50:50 % v/v) was added to the pellet sample. The re-suspended cells were lysed by freezing in liquid nitrogen, thawing on ice and vortexing for 30 s. The freeze-thaw method was repeated three times. Chloroform (100 μL) was added followed by centrifugation to remove cell debris. Supernatant was and stored 4 $^{\circ}$C prior to analysis.

4.1 Pathway metabolites and mimic synthesis

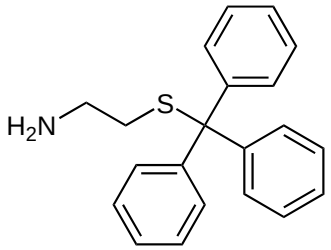
S-tritylcysteamine hydrochloride (140)[59]



140

Cysteamine hydrochloride **40** (822 mg, 7.26 mmol) was dissolved in a mixture of dimethylformamide and dichloromethane (1:1, 40 mL). Trityl chloride (2.93 g, 10.5 mmol) was added and stirred for two hours, followed by concentration *in vacuo*. The residue was then partitioned between dichloromethane and water (1:1, 500 mL) so that it formed a white precipitate between the two layers. The precipitate was isolated between the organic and aqueous phases and filtered off. This was dried in vacuo to yield the product **140** as a white solid (2.33 g, 6.53 mmol, 90 %) to be taken forward into next step. HRMS m/z (ESI+) found 342.1290 $[MNa]^+$ (319.1392) $C_{21}H_{21}NS$ requires 319.1395.

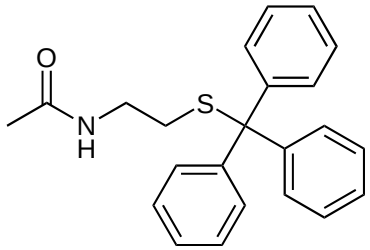
S-tritylcysteamine (41)[59]



41

S-tritylcysteamine hydrochloride **140** (1.57 g, 4.40 mmol) was solubilised in diethyl ether (270 mL) and potassium hydroxide (17.5 mM, 100 mL). To this was added sodium hydroxide (9.6 M) dropwise until the precipitate disappeared. This was repeated, dried over sodium sulfate and concentrated *in vacuo* to yield the product **41** as a white solid (1.13 g, 3.52 mmol, 80 %). M.p. 96 – 97 °C. ν_{max}(nujol)/cm^{-1} 3056 ($CH_{aromatic}$), 2923 (CH), 1658 (NH); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.34 (2H, t, *J* 6.5 Hz, C*H*$_2$), 2.62 (2H, t, *J* 6.5 Hz, C*H*$_2$), 7.23 (3H, tt, *J* 7.2, 2.2 Hz, ArC*H*), 7.31 (6H, tt, *J* 7.8, 1.6 Hz, ArC*H*), 7.45 (6H, dt, *J* 7.8, 1.6 Hz, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 36.3 (*C*H$_2$), 41.1 (*C*H$_2$), 66.6 (*C*Ph$_3$), 126.7 (Ar*C*H), 127.9 (Ar*C*H), 129.6 (Ar*C*H), 144.9 (Ar*C*); HRMS m/z (ESI+) found 342.1288 [MNa]$^+$ (319.1390) $C_{21}H_{21}NS$ requires 319.1395.

***N*-(2-(tritylthio)ethyl)acetamide (55)**

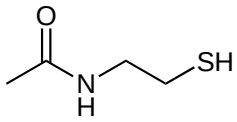


55

To a stirred solution of acetic acid (0.04 mL, 0.625 mmol) in dichloromethane (5 mL) at 0 °C, 4-dimethylaminopyridine (15.0 mg, 0.125 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (96.0 mg, 0.500 mmol) and *S*-tritylcysteamine **41** (159 mg, 0.500 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride (5 mL) and extracted with dichloromethane (3 x 10 mL). The extracts were combined and washed with 1 M HCl (2 x 20 mL), water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. This was filtered and concentrated *in vacuo* and purified by flash column chromatography using petroleum ether:ethyl acetate (1:1) to yield the product **55** as a white solid (136 mg, 0.377 mmol, 75 %). M.p. 145 – 147 °C. ν_{max}(nujol)/cm^{-1} 3058 ($CH_{aromatic}$), 1634 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.88 (3H, s, C*H*$_3$), 2.42 (2H, t, *J* 6.2 Hz, C*H*$_2$), 3.08, appt. q, *J* 6.2 Hz, C*H*$_2$), 5.56 (1H, br s, N*H*), 7.22 (3H, t,

J 7.8 Hz, ArC*H*), 7.29 (6H, t, *J* 7.8 Hz, ArC*H*), 7.42 (2H, d, *J* 7.8 Hz, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 23.3 (*C*H_3), 32.0 (*C*H_2), 38.3 (*C*H_2), 66.9 (*C*Ph_3), 126.9 (Ar*C*H), 128.0 (Ar*C*H), 129.5 (Ar*C*H), 144.7 (Ar*C*), 169.9 (*C*O); HRMS m/z (ESI+) found 384.1395 $[MNa]^+$ (361.1497) $C_{23}H_{23}NOS$ requires 361.1500.

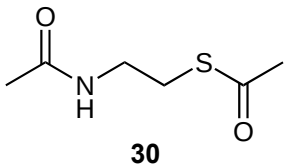
_N_-acetylcysteamine (56)



56

N-(2-(tritylthio)ethyl)acetamide **55** (200 mg, 0.55 mmol) was dissolved in dichloromethane (10 mL) and triethylsilane (0.19 mL, 1.10 mmol) was added. The mixture was cooled to 0 °C, trifluoroacetic acid (8.22 mL, 15.1 mmol) was added dropwise, stirred for 5 min at 0 °C, then 1 h at r.t., before concentration *in vacuo*. The crude product was then purified by flash column chromatography using dichloromethane, eluting with MeOH (10 %) to yield the product **56** as a colourless oil (51 mg, 0.43 mmol, 78 %). ν_{max}(nujol)/cm^{-1} 2930 (CH), 1653 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.19 (3H, s, C*H*$_3$), 2.30 (2H, q, *J* 7.5 Hz, C*H*$_2$), 3.61 (2H, m, C*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 9.38 (*C*H_3), 29.1 (*C*H_2), 49.6 (*C*H_2); HRMS m/z (ESI+) found 156.0450 $[MNa]^+$ (133.0552) $C_5H_{11}NOS$ requires 133.0561.

Acetyl-SNAC (30)[48]

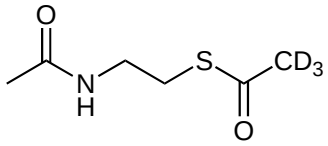


30

To a stirred solution of acetic acid (0.10 mL, 1.70 mmol) in dichloromethane (10 mL) at 0 °C, 4-dimethylaminopyridine (61.0 mg, 0.50 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (382 mg, 2.00 mmol) and *N*-acetylcysteamine **56** (0.21 mL, 2.00 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride (10 mL) and extracted three times with dichloromethane (3 x 20 mL). The extracts were combined and washed with water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using ethyl acetate to yield the product **30** as a colourless oil (171 mg, 1.03 mmol, 61 %). ν_{max}(nujol)/cm^{-1} 1692 (C=O), 1655 (C=O), 1553 (N-H); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.96 (3H, s, C*H*$_3$), 2.34 (3H, s, C*H*$_3$), 3.01 (2H, t, *J* 6.4 Hz, C*H*$_2$), 3.41 (2H, appt q, *J* 5.3 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 23.2 (*C*H_3), 28.8 (*C*H_2), 30.6 (*C*H_3), 39.6 (*C*H_2), 170.5 (N*C*O), 196.3

(S*C*O); HRMS m/z (ESI+) found 184.0414 $[MNa]^+$ (161.0516) $C_6H_{11}NO_2S$ requires 161.0510.

S-(2-acetamidoethyl) ethanethioate-2-$^2H_3$ (52)

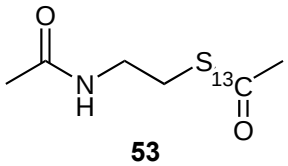


52

To a stirred solution of d_4-acetic acid (0.50 mL, 8.73 mmol) in dichloromethane (20 mL) at 0 °C, 4-dimethylaminopyridine (305 mg, 2.50 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (1.91 g, 10.0 mmol) and *N*-acetylcysteamine **56** (0.92 mL, 8.64 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride and extracted three times with dichloromethane. The extracts were combined and washed with water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using ethyl acetate to yield the product **52** as a colourless oil (482 mg, 2.94 mmol, 34 %). ν_{max}(nujol)/cm^{-1} 2927 (CH), 1691 (CO), 1643 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.97 (3H, s, C*H*$_3$), 3.00 (2H, t, *J* 6.3 Hz, C*H*$_2$), 3.39 (2H, q, *J* 6.3 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 22.9 (*C*H_3), 28.6 (*C*H_2), 29.9 (m, *C*D_3), 39.6 (*C*H_2), 170.9 (NCO), 196.2 (SCO);

HRMS m/z (ESI+) found 187.0606 $[MNa]^+$ (164.0708) $C_6H_8D_3NO_2S$ requires 164.0699.

1-^{13}C-*S*-(2-acetamidoethyl) ethanethioate (53)

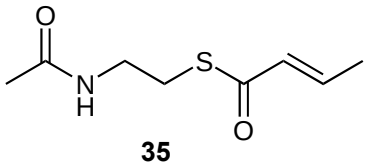


53

To a stirred solution of 1-^{13}C-acetic acid (0.17 mL, 3.00 mmol) in dichloromethane (20 mL) at 0 °C, 4-dimethylaminopyridine (122 mg, 1.00 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (767 mg, 4.00 mmol) and *N*-acetylcysteamine **56** (0.31 mL, 3.00 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride and extracted three times with dichloromethane. The extracts were combined and washed with water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using ethyl acetate to yield the product **53** as a colourless oil (266 mg, 1.64 mmol, 55 %). ν_{max}(nujol)/cm^{-1} 2930 (CH), 1692 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.95 (3H, s, C*H*$_3$), 2.34 (3H, d, $J^{13C\text{-}1H}$ 6.2 Hz, C*H*$_3$), 3.01 (2H, dt, *J* 4.8, 6.9 Hz, C*H*$_2$), 3.41 (2H, appt q, *J* 6.3 Hz, C*H*$_2$) 5.97 (1H, br s, N*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 23.2 (*C*H$_3$), 28.9 (*C*H$_2$),

30.6 (d, *J* 47.3 Hz, *C*H_3), 39.6 (*C*H_2), 170.3 (*C*O), 196.3 (*C*O); HRMS m/z (ESI+) found 185.0438 $[MNa]^+$ (162.0540) $C_5{}^{13}CH_{11}NO_2S$ requires 162.0544.

_S_-(2-acetamidoethyl) (_E_)-but-2-enethioate (35)

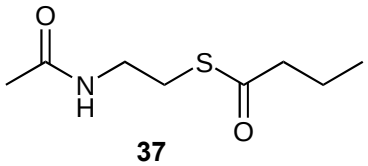


35

To a stirred solution of *trans*-crotonic acid (155 mg, 1.80 mmol) in dichloromethane (10 mL) at 0 °C, 4-dimethylaminopyridine (61.0 mg, 0.50 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (382 mg, 2.00 mmol) and *N*-acetylcysteamine **56** (0.21 mL, 2.00 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride and extracted three times with dichloromethane. The extracts were combined and washed with water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using ethyl acetate to yield the product **35** as a colourless oil (226 mg, 1.21 mmol, 67 %). ν_{max}(film)/cm^{-1} 3017 (C-H_{alkene}), 1665 (C=O), 1637 (C=O), 1529 (N-H); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.89 (3H, dd, *J* 7.0, 1.2, Hz, C*H*$_3$), 1.96 (3H, s, C*H*$_3$), 3.08

(3H, t, *J* 6.0 Hz, C*H*$_2$), 3.44 (3H, appt q, *J* 6.7 Hz, C*H*$_2$), 6.15 (1H, dd, *J* 15.4 and 1.2 Hz, C*H*), 6.19 (1H, br s, N*H*), 6.93 (1H, dq, *J* 15.4, 7.0 Hz, CH); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 18.0 (*C*H$_3$), 23.2 (*C*H$_3$), 28.2 (*C*H$_2$), 39.8 (*C*H$_2$), 129.9 (*C*H), 141.8 (*C*H), 170.5 (*C*O), 190.2 (*C*O); HRMS m/z (ESI+) found 210.0552 $[MNa]^+$ (187.0654) $C_8H_{13}NO_2S$ requires 187.0667.

_S_-(2-acetamidoethyl) butanethioate (37)

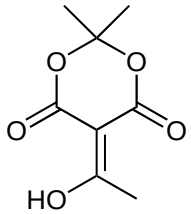


37

To a stirred solution of butyric acid **23** (0.16 mL, 1.76 mmol) in dichloromethane (10 mL) at 0 °C, 4-dimethylaminopyridine (61.0 mg, 0.50 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (382 mg, 2.00 mmol) and *N*-acetylcysteamine **56** (0.21 mL, 2.00 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride and extracted three times with dichloromethane. The extracts were combined and washed with water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using ethyl acetate to yield the product **37** as a white oily solid (220 mg, 1.16 mmol, 66 %). ν_{max}(nujol)/cm$^-$

[1] 1691 (C=O), 1655 (C=O), 1553 (N-H); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.92 (3H, t, *J* 6.0 Hz, C*H*$_3$), 1.66 (2H, appt. sex, *J* 7.2 Hz, C*H*$_2$), 1.93 (3H, s, C*H*$_3$), 2.52 (2H, t, *J* 6.0 Hz, C*H*$_2$), 2.99 (2H, t, *J* 6.0 Hz, C*H*$_2$), 3.39 (2H, q, *J* 6.7 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 13.7 (*C*H$_3$), 19.4 (*C*H$_2$), 23.3 (*C*H$_3$), 28.6 (*C*H$_2$), 39.9 (*C*H$_2$), 46.1 (*C*H$_2$), 170.7 (N*C*O), 200.1 (S*C*O); HRMS m/z (ESI+) found 212.0728 $[MNa]^+$ (189.0830) $C_8H_{15}NO_2S$ requires 189.0823.

5-(1-hydroxyethylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (58)[58]

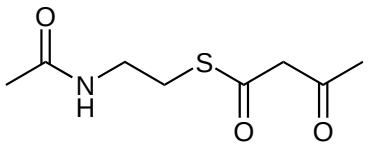


58

Meldrum's acid **57** (1.82 g, 12.7 mmol) was dissolved in dichloromethane (20 mL) before addition of 4-dimethylaminopyridine (3.09 g, 25.3 mmol) and stirred at 0 °C under argon. After 5 minutes stirring, acetyl chloride (0.90 mL, 12.7 mmol) was added and stirred for 30 min at 0 °C before stirring to room temperature for 30 min. Hydrochloric acid (1 M, 10 mL) was added and extracted with dichloromethane (3 x 10 mL). The organic layer was separated and washed with hydrochloric acid (1 M, 2 x 10 mL) and brine

(2 x 20 mL), dried over magnesium sulphate and concentrated *in vacuo* to yield a yellow solid **58** (1.98 g, 10.6 mmol, 84 %). M.p. 84 – 96 °C. ν_{max}(nujol)/cm^{-1} 2991 (CH_{alkane}), 2852 (CH_{alkene}), 1740 (CO), 1663 (CO), 1578 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.73 (6H, s, C*H*$_3$), 2.68 (3H, s, C*H*$_3$); HRMS m/z (ESI+) found 209.0416 $[MNa]^+$ (186.0518) $C_8H_{10}O_5$ requires 186.0528.

S-(2-acetamidoethyl) 3-oxobutanethioate (31)[96]

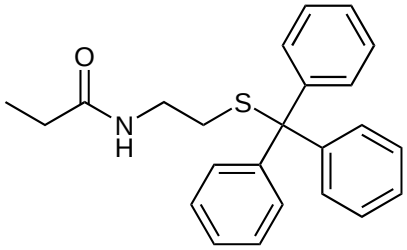


31

5-(1-hydroxyethylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione **58** (1.06 g, 5.70 mmol) was dissolved in toluene (16 mL). N-acetylcysteamine **56** (0.4 mL, 3.80 mmol) in toluene (4 mL) was added and stirred at 80 °C for 16 h before concentrating *in vacuo*. The residue was purified by flash column chromatography using petroleum ether:ethyl acetate 1:1 eluting with ethyl acetate to yield the product **31** as a colourless oil (660 mg, 3.25 mmol, 86 %). ν_{max}(nujol)/cm^{-1} 2923 (CH), 1720 (CO), 1673 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.00 (3H, s, C*H*$_3$), 2.30 (3H, s, C*H*$_3$), 3.12 (2H, t, *J* 6.1 Hz, C*H*$_2$), 3.49 (2H, q, *J* 6.1 Hz, C*H*$_2$), 3.75 (2H, s, C*H*$_2$) 5.90 (1H,

br s, N*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 23.2 (*C*H_3), 29.3 (*C*H_3), 30.4 (*C*H_2), 39.2 (*C*H_2), 58.0 (*C*H_2), 174.1 (*C*O), 192.3 (*C*O), 199.8 (*C*O); HRMS m/z (ESI+) found 226.0521 $[MNa]^+$ (203.0623) $C_8H_{13}NO_3S$ requires 203.0616.

***N*-(2-(tritylthio)ethyl)propionamide (59)**

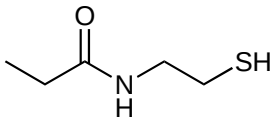


59

To a stirred solution of propionic acid (0.19 mL, 2.50 mmol) in dichloromethane (10 mL) at 0 °C, 4-dimethylaminopyridine (61.0 mg, 0.50 mmol), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (382 mg, 2.00 mmol) and *S*-tritylcysteamine **41** (637 mg, 2.00 mmol) were added, and the mixture stirred at r.t. for 16 h. The reaction mixture was quenched with saturated aqueous ammonium chloride and extracted three times with dichloromethane. The extracts were combined and washed with 1 M HCl (2 x 20 mL), $NaHCO_3$ (sat., 2 x 20 mL), water (2 x 20 mL) and brine (2 x 20 mL) and dried over magnesium sulfate. These were then filtered and concentrated *in vacuo* and purified by flash column chromatography using petroleum ether:ethyl acetate (1:1) to

yield the product **59** as a white solid (647 mg, 1.72 mmol, 86 %). M.p. 176 – 177 °C. ν_{max}(film)/cm^{-1} 3061 ($CH_{aromatic}$), 2923 (CH_{alkane}), 1642 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.12 (3H, t, *J* 7.6 Hz, C*H*$_3$), 2.13 (2H, q, *J* 7.6 Hz, C*H*$_2$), 2.43 (2H, t, *J* 6.0 Hz, C*H*$_2$), 3.11 (2H, appt q, *J* 5.3 Hz, C*H*$_2$), 7.22 (3H, tt, *J* 7.1, 2.0 Hz, ArC*H*), 7.29 (6H, tt, *J* 7.1, 2.0 Hz , ArC*H*), 7.41 (6H, dt, *J* 7.7, 1.5 Hz, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 9.8 (*C*H$_3$), 29.7 (*C*H$_2$), 32.1 (*C*H$_2$), 38.0 (*C*H$_2$), 66.8 (*C*Ar$_3$), 126.8 (Ar*C*H), 128.0 (Ar*C*H), 129.5 (Ar*C*H), 144.6 (Ar*C*), 173.5 (*C*O); HRMS m/z (ESI+) found 398.1548 [MNa]$^+$ (375.1650) $C_{24}H_{25}NOS$ requires 375.1657.

N-(2-mercaptoethyl)propionamide (60)

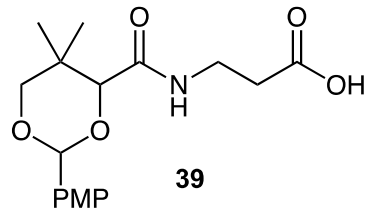


60

N-(2-(tritylthio)ethyl)propionamide **59** (207 mg, 0.55 mmol) was dissolved in dichloromethane (20 mL) and triethylsilane (0.19 mL, 1.10 mmol) was added. After cooling to 0 °C, trifluoroacetic acid (8.22 mL, 15.1 mmol) was added dropwise and stirred at 0 °C for 5 min, then stirred at r.t. for 30 min. The solvent was removed *in vacuo* and the product purified by flash column chromatography using dichloromethane, and eluting with dichloromethane:methanol (95:5) to yield the

product **60** as a waxy colourless solid (26 mg, 0.20 mmol, 36 %). ν_{max}(film)/cm^{-1} 2928 (CH), 1652 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.20 (3H, t, *J* 7.6 Hz, C*H*$_3$), 1.37 (1H, t, *J* 8.3 Hz, S*H*), 2.28 (2H, appt q, *J* 7.7 Hz, C*H*$_2$), 2.71 (2H, dt, *J* 6.2, 8.3 Hz, C*H*$_2$), 3.47 (2H, q, *J* 6.2 Hz, C*H*$_2$), 5.96 (1H, br s, N*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 9.8 (*C*H$_3$), 24.7 (*C*H$_2$), 29.7 (*C*H$_2$), 42.3 (*C*H$_2$), 174.3 (*C*O); HRMS m/z (ESI+) found 156.0505 [MNa]$^+$ (133.0607) $C_5H_{11}NOS$ requires 133.0561.

3-(2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamido)propanoic acid (39)[52, 60]

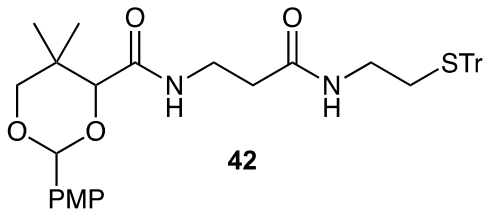


39

Calcium D-pantothenate **38** (5 g, 21.0 mmol) was dissolved in anhydrous dimethylformamide (50 mL) and concentrated sulfuric acid (0.65 mL, 21.0 mmol) was added dropwise and stirred for 30 minutes. If a gel formed, an additional portion of anhydrous dimethylformamide (50 mL) was added and sonicated. *p*-anisaldehyde dimethyl acetal (3.6 mL, 21.0 mmol) and 10-camphorsulfonic acid (244 mg, 1.05 mmol) was added and stirred for 16 h, then the solvent removed *in vacuo*. The residue was dissolved in ethyl acetate (500 mL) and water

(50 mL) followed by the organic layer being washed with water (2 x 50 mL), dried over sodium sulfate and concentrated *in vacuo.* The white solid is then washed with cold dichloromethane to remove any residual *p*-anisaldehyde dimethyl acetal to yield a white, crystalline solid as the product **39** (1.9 g, 5.66 mmol, 27 %). M.p. 158 – 159 °C. ν_{max}(nujol)/cm^{-1} 3072 ($CH_{aromatic}$), 1729 (CO_{acid}), 1623 (CO_{amide}); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.12 (3H, s, C*H*$_3$), 1.13 (3H, s, C*H*$_3$), 2.64 (2H, t, *J* 6.2 Hz, C*H*$_2$), 3.56 (2H, m, C*H*$_2$), 3.68 (1H, d, *J* 11.7 Hz, C*H*$_2$), 3.74 (1H, d, *J* 11.7 Hz, C*H*$_2$), 3.88 (3H, s, OC*H*$_3$), 4.12 (1H, s, C*H*), 5.49 (1H, s, C*H*), 6.94 (2H, d, *J* 8.8 Hz, ArC*H*), 7.06 (1H, t, *J* 5.5 Hz, N*H*), 7.44 (1H, d, *J* 8.8 Hz, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 19.1 (*C*H$_3$), 21.8 (*C*H$_3$), 33.1 (*C*CH$_3$), 33.7 (*C*H$_2$), 34.1 (*C*H$_2$), 55.3 (O*C*H$_3$), 78.5 (*C*H$_2$), 83.7 (*C*H), 101.3 (*C*H), 113.8 (Ar*C*H), 127.4 (Ar*C*H), 130.2 (Ar*C*), 160.2 (Ar*C*), 169.5 (*C*O), 175.9 (*C*O); HRMS m/z (ESI+) found 360.1418 [MNa]$^+$ (337.1520) $C_{17}H_{23}NO_6$ requires 337.1525.

2-(4-methoxyphenyl)-5,5-dimethyl-N-(3-oxo-3-((2-(tritylthio)ethyl)amino)propyl)-1,3-dioxane-4-carboxamide (42)

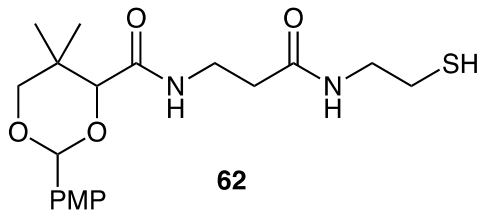


42

3-(2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamido)propanoic acid **39** (375 mg, 1.05 mmol) was dissolved in dichloromethane (10 mL) and *S*-tritylcysteamine **41** (336 mg, 1.05 mmol) was added, followed by addition of 4-dimethylaminopyridine (cat.), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (202 mg, 1.3 mmol) and triethylamine (0.18 mL, 1.3 mmol). This was stirred for 16 h, before addition of ammonium chloride (10 mL, sat.). Dichloromethane (40 mL) was added and the organic layer was extracted sequentially with hydrochloric acid (1 M, 2 x 50 mL), sodium hydrogen carbonate (2 x 50 mL, sat.), water (1 x 50 mL) and brine (1 x 100 mL). The organic layer was dried over sodium sulfate before being concentrated *in vacuo* to yield the product **42** as a white solid (551 mg, 0.86 mmol, 82 %). M.p. 88 – 90 °C. ν_{max}(nujol)/cm^{-1} 3313 ($CH_{aromatic}$), 2927 (CH), 1659 (CO), 1250 ($CO_{aromatic}$); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.08 (3H, s, C*H*$_3$), 1.09 (3H, s, C*H*$_3$), 1.60 (3H, s, C*H*$_3$),

2.35 (2H, dt, *J* 1.6, 6.0 Hz, C*H*$_2$), 2.43 (2H, m, C*H*$_2$), 2.64 (2H, t, *J* 6.2 Hz, C*H*$_2$), 3.04 (2H, m, C*H*$_2$), 3.51 (2H, appt q, *J* 6.3 Hz, C*H*$_2$), 3.64 (1H, d, *J* 11.3 Hz, C*H*$_2$), 3.71 (1H, d, *J* 11.3 Hz, C*H*$_2$), 3.84 (3H, s, OC*H*$_3$), 4.07 (1H, s, C*H*), 5.44 (1H, s, C*H*), 5.70 (1H, t, *J* 5.6 Hz, N*H*), 6.94 (2H, d, *J* 8.8 Hz, ArC*H*), 7.00 (1H, t, *J* 6.1 Hz, N*H*), 7.24 (3H, m, ArC*H*), 7.31 (6H, m, ArC*H*), 7.43 (8H, m, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 19.1 (*C*H$_3$), 21.9 (*C*H$_3$), 31.8 (*C*H$_2$), 33.1 (*C*CH$_3$), 34.8 (*C*H$_2$), 35.9 (*C*H$_2$), 38.2 (*C*H$_2$), 55.4 (O*C*H$_3$), 67.0 (*C*Ph$_3$), 78.5 (*C*H$_2$), 83.8 (*C*H), 101.3 (*C*H), 113.8 (Ar*C*H), 126.9 (Ar*C*H), 127.5 (Ar*C*H), 128.0 (Ar*C*H), 129.5 (Ar*C*H), 130.2 (Ar*C*), 144.6 (Ar*C*), 160.2 (Ar*C*), 169.5 (*C*O), 170.6 (*C*O); HRMS m/z (ESI+) found 707.3823 $[M+Im]^+$ (638.3370) $C_{38}H_{42}N_2O_5S$ requires 638.2814.

PMP-pantetheine (62)[61]

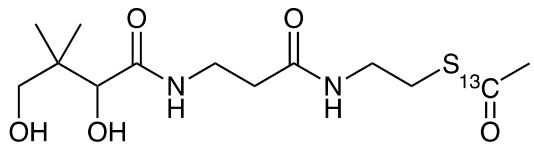


62

3-(2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamido)propanoic acid **39** (774 mg, 2.00 mmol) was dissolved in dichloromethane (10 mL) and cysteamine hydrochloride **40** (452 mg, 4.00 mmol) was added, followed

by addition of PyBOP (1.50 g, 3.00 mmol) and diisopropylethylamine (1.58 mL, 10.0 mmol). This was stirred for 16 h, before concentration *in vacuo.* The compound was purified by gradient column chromatography using hexane:ethyl acetate (1:1) to ethyl acetate (100 %) to yield the product **62** as a colourless oil (689 mg, 1.74 mmol, 87 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.10 (3H, s, C*H*$_3$), 1.12. (3H, s, C*H*$_3$), 1.37 (1H, t, *J* 8.4 Hz, S*H*), 2.49 (2H, t, *J* 6.0 Hz, C*H*$_2$), 2.62 (2H, dt, *J* 6.7, 8.0 Hz, C*H*$_2$), 3.42 (2H, m, C*H*$_2$), 3.59 (2H, t, *J* 6.6 Hz, C*H*$_2$), 3.70 (2H, d, *J* 10.7 Hz, C*H*$_2$), 3.85 (3H, s, OC*H*$_3$), 4.12 (1H, s, C*H*), 5.48 (1H, s, C*H*), 6.44 (1H, br s, N*H*), 6.94 (2H, d, *J* 8.5 Hz, ArC*H*), 7.09 (1H, t, *J* 6.0 Hz, N*H*), 7.45 (2H, d, *J* 8.5 Hz, ArC*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 19.1 (*C*H$_3$), 21.9 (*C*H$_3$), 24.3 (*C*H$_2$), 33.1 (*C*CH$_3$), 34.9 (*C*H$_2$), 36.1 (*C*H$_2$), 42.5 (*C*H$_2$), 55.4 (O*C*H$_3$), 78.4 (*C*H$_2$), 83.8 (*C*H), 101.4 (*C*H), 113.8 (Ar*C*H), 127.5 (Ar*C*H), 130.0 (Ar*C*), 160.3 (Ar*C*), 170.0 (*C*O), 171.2 (*C*O); HRMS m/z (ESI+) found 419.1594 $[MNa]^+$ (396.1696) $C_{19}H_{28}N_2O_5S$ requires 396.1719.

2-^{13}C-acetyl pantetheine (54)

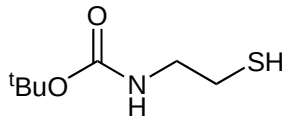


54

PMP-pantetheine **62** (680 mg, 1.71 mmol) was dissolved in dichloromethane (10 mL) and 2-^{13}C-acetic acid (0.11 mL, 1.90 mmol) was added, followed by addition of 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (493 mg, 2.57 mmol) and diisopropylethylamine (0.40 mL, 3.40 mmol). This was stirred for 16 h, before addition of ammonium chloride (10 mL, sat.). Dichloromethane (40 mL) was added and the organic layer was extracted sequentially with hydrochloric acid (1 M, 2 x 50 mL), sodium hydrogen carbonate (2 x 50 mL, sat.), water (1 x 50 mL) and brine (1 x 100 mL). The organic layer was dried over sodium sulfate before being concentrated *in vacuo* to yield the coupled product as a colourless oil (611 mg, 1.39 mmol, 81 %). HRMS m/z (ESI+) found 462.1739 $[MNa]^+$ (439.1841) $C_{20}{}^{13}CH_{30}N_2O_6S$ requires 439.1858.

The colourless oil was subsequently dissolved in acetic acid:water (2:1, 30 mL) and stirred for 16 h. Solvent was removed *in vacuo* and purified by flash column chromatography using dichloromethane:methanol (20:1) to yield 2-^{13}C-acetyl pantetheine **54** as a light yellow oil (374 mg, 1.17 mmol, 84 %). ν_{max}(film)/cm^{-1} 2931 (CH), 2875 (CH), 1652 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.94 (3H, s, C*H*$_3$), 1.04 (3H, s, C*H*$_3$), 1.23 (2H, t, *J* 6.8 Hz, O*H*), 2.38 (3H, d, *J* 6.2 Hz, C*H*$_3$), 2.45 (2H, t, *J* 5.9 Hz, C*H*$_2$), 3.05 (2H, m, C*H*$_2$), 3.41 (1H, sep, *J* 6.9 Hz, C*H*$_2$), 3.49 (5H, m, C*H*$_2$), 3.59 (2H, q, *J* 6.1 Hz, C*H*$_2$), 4.01 (1H, s, C*H*), 6.32 (1H, br s, N*H*), 7.40 (1H, br s, N*H*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 20.4 (*C*H$_3$), 21.6 (*C*H$_3$), 28.6 (*C*H$_2$), 30.6 (d, *J* 46.6 Hz, *C*H$_3$), 35.2 (*C*H$_2$), 35.6 (*C*H$_2$), 39.4 (d, *J* 20.8 Hz, *C*H$_2$), 65.8 (*C*CH$_3$), 71.0 (*C*H$_2$), 77.7 (*C*H), 170.5 (*C*O), 171.7 (*C*O), 196.6 (*C*O); HRMS m/z (ESI+) found 344.1330 [MNa]$^+$ (321.1432) $C_{12}{}^{13}CH_{24}N_2O_5S$ requires 321.1439.

_N_-Boc-cysteamine (141)

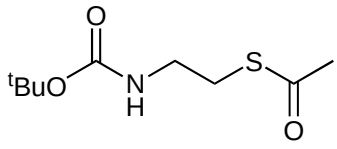


141

Cysteamine hydrochloride **40** (2.27 g, 20.0 mmol) was dissolved in dichloromethane (100 mL). Di-*tert*-butyl

dicarbonate (4.80 g, 22.0 mmol) was added before dropwise addition of triethylamine (6.10 mL, 44.0 mmol) in dichloromethane (100 mL) and stirred at r.t. for 16 h, before concentrating *in vacuo*. The residue was redissolved in diethyl ether and washed with hydrochloric acid (2 x 50 mL, 1 M), sodium hydrogen carbonate (2 x 50 mL, sat.) and brine (2 x 50 mL, sat.) before drying over magnesium sulfate and concentration *in vacuo* to yield product **141** as a colourless oil (2.93 g, 16.6 mmol, 83 %). ν_{max}(film)/cm^{-1} 2977 (CH), 2931 (CH), 1695 (CO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.51 (9H, s, ^{t}Bu), 2.93 (2H, t, *J* 6.4 Hz, C*H*$_2$), 3.38 (2H, m, C*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 25.0 (*C*H$_2$), 28.4 (^{t}Bu), 40.6 (*C*H$_2$), 85.1 (*C*^{t}Bu), 155.7 (*C*O); HRMS m/z (ESI+) found 200.0725 [MNa]$^{+}$ (177.0827) $C_7H_{15}NO_2S$ requires 177.0823.

S-(2-((tert-butoxycarbonyl)amino)ethyl) ethanethioate (142)



142

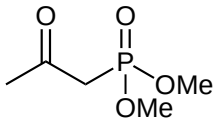
N-Boc-cysteamine **141** (1.0 g, 5.4 mmol) was dissolved in dichloromethane (10 mL) and 4-dimethylaminopyridine (cat.), 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (1.3 g, 7.0 mmol) and triethylamine (0.84 mL, 6 mmol) were

added. The mixture was cooled to 0 °C and glacial acetic acid (0.34 mL) added dropwise and the mixture stirred to rt over 16 h. Dichloromethane (40 mL) and saturated ammonium chloride (50 mL) were added and the two layers separated before washing the organic layer with hydrochloric acid (2 x 50 mL, 1 M), sodium hydrogen carbonate (2 x 50 mL, sat.), water (50 mL) and brine (50 mL). This was then dried over sodium sulfate and concentrated *in vacuo* to yield the product **142** as a colourless oil (1.0 g, 4.56 mmol, 84 %). ^{1}H NMR (400 MHz, D_2O) δ_H 1.46 (9H, s, C*H*$_3$), 2.37 (3H, s, C*H*$_3$), 3.02 (2H, t, *J* 6.5 Hz, C*H*$_2$), 3.32 (2H, appt q, *J* 5.8 Hz, C*H*$_2$), 4.82 (1H, br s, N*H*); ^{13}C NMR (100 MHz, D_2O) δ_C 28.4 (*C*H$_3$), 29.4 (*C*H$_3$), 30.6 (*C*H$_2$), 40.3 (*C*H$_2$), 79.5 (*C*^{t}Bu), 155.8 (*C*O), 195.8 (*C*O); HRMS m/z (ESI+) found 242.0820 $[MNa]^+$ (219.0942) $C_9H_{17}NO_3S$ requires 219.0929.

4.2 Acetate kinase inhibitors

β-Ketophosphonate inhibitor synthesis

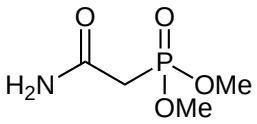
Dimethyl (2-oxopropyl)phosphonate (68)[72]



68

Chloroacetone (0.43 mL, 5.35 mmol) and potassium iodide (820 mg, 4.90 mmol) in anhydrous acetone (4 mL) and anhydrous acetonitrile (5 mL) was stirred at r.t. for 2 h under argon. Trimethyl phosphite (0.58 mL, 4.90 mmol) was added, stirred for 16 h at 56 °C to give the crude product. The mixture was cooled to r.t., filtered through Celite® and concentrated *in vacuo* before flash column chromatography using dichloromethane:methanol 99:1 to yield the product **68** as a light yellow oil (473 mg, 2.85 mmol, 58 %). ν_{max}(film)/cm^{-1} 2959 (CH), 1716 (CO), 1462 (CP), 1263 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.26 (3H, s, C*H*$_3$), 3.04 (2H, d, *J* 22.7 Hz, C*H*$_2$), 3.73 (6H, d, *J* 11.3 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 31.4 (*C*H$_3$), 42.3 (d, *J* 127.9 Hz, *C*H$_2$), 53.0 (d, *J* 6.3 Hz, O*C*H$_3$), 199.7 (*C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P 22.4; HRMS m/z (ESI+) found 189.0297 [MNa]$^+$ (166.0399) $C_5H_{11}O_4P$ requires 166.0395.

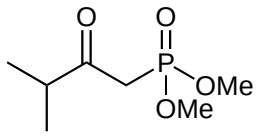
Dimethyl (2-amino-2-oxoethyl)phosphonate (69)



69

To a solution of iodoacetamide (314 mg, 1.70 mmol) in anhydrous acetonitrile (2 mL) was added trimethyl phosphite (0.1 mL, 0.86 mmol) and stirred for 16 h at 56 °C. The mixture was purified *via* flash column chromatography using dichloromethane:methanol 96:4 to yield the product **69** as a white waxy solid (89 mg, 0.53 mmol, 62 %). ν_{max}(nujol)/cm^{-1} 2923 (CH), 1666 (CO), 1227 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.94 (2H, d, *J* 21.0 Hz, C*H*$_2$), 3.83 (6H, d, *J* 11.3 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 34.4 (d, *J* 131.7 Hz, *C*H_2), 53.3 (d, *J* 6.2 Hz, O*C*H_3), 166.6 (*C*O); HRMS m/z (ESI+) found 190.0242 $[MNa]^+$ (167.0344) $C_4H_{10}NO_4P$ requires 167.0347.

Dimethyl (3-methyl-2-oxobutyl)phosphonate (75)[74]

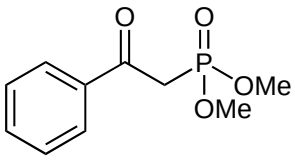


75

To a solution of n-butyllithium (4.7 mL, 1.6 M in hexanes, 7.5 mmol) in anhydrous tetrahydrofuran (30 mL) at -78 °C was

added dimethyl methyl phosphonate **74** (0.54 mL, 5.0 mmol) in anyhydrous tetrahydrofuran (20 mL) dropwise over a period of and stirred for 1 h. Isobutyryl chloride (0.52 mL, 5.0 mmol) was added dropwise, stirred at -78 °C for 2 h before stirring at r.t. for 16 h. Water (20 mL) and chloroform (100 mL) were added before filtering the resulting suspension through cotton wool, rinsing with chloroform. The organic layer was separated and dried over anhydrous sodium sulfate before concentration *in vacuo*. The resulting oil was purified by gradient flash column chromatography hexane:ethyl acetate 95:5, eluting with ethyl acetate to give product **75** as a colourless oil (417 mg, 2.15 mmol, 43 %). ν_{max}(film)/cm^{-1} 2967 (CH), 1710 (CO), 1466 (CP), 1255 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.12 (6H, d, *J* 6.9 Hz, C*H*$_3$), 2.83 (1H, quin, *J* 6.9 Hz, C*H*), 3.13 (2H, d, *J* 22.6 Hz, C*H*$_2$), 3.78 (6H, d, *J* 11.2 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 17.9 (*C*H_3), 39.0 (d, *J* 130.5 Hz, *C*H_2), 41.9 (d, *J* 2.0 Hz, *C*H), 53.0 (d, *J* 6.2 Hz, O*C*H_3), 205.9 (d, *J* 6.5 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P 23.1; HRMS m/z (ESI+) found 217.0606 [MNa]$^+$ (194.0708) $C_7H_{15}O_4P$ requires 194.0708.

Dimethyl (2-oxo-2-phenylethyl)phosphonate (70)[74]



70

To a solution of n-butyllithium (3.0 mL, 2.5 M in hexanes, 7.5 mmol) in anhydrous tetrahydrofuran (30 mL) at -78 °C was added dimethyl methyl phosphonate **74** (0.54 mL, 5.0 mmol) in anyhydrous tetrahydrofuran (20 mL) dropwise over a period of and stirred for 1 h. Benzoyl chloride (0.58 mL, 5.0 mmol) was added dropwise, stirred at -78 °C for 2 h before stirring at r.t. for 16 h. Water (20 mL) and chloroform (100 mL) were added before filtering the resulting suspension through cotton wool, rinsing with chloroform. The organic layer was separated and dried over anhydrous sodium sulfate before concentration *in vacuo*. The resulting oil was purified by gradient flash column chromatography hexane:ethyl acetate 95:5, eluting with ethyl acetate to give product **70** as a colourless oil (666 mg, 2.9 mmol, 58 %). ν_{max}(film)/cm^{-1} 3061 ($CH_{aromatic}$), 2955 (CH_{alkane}), 1679 (CO), 1449 (CP), 1254 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 3.66 (2H, d, *J* 22.6 Hz, C*H*$_2$), 3.80 (6H, d, *J* 11.1 Hz, OC*H*$_3$), 7.50 (2H, t, *J* 7.6 Hz, ArC*H*), 7.61 (1H, t, *J* 7.6 Hz, ArC*H*), 8.01 (2H, d, *J* 7.6 Hz, ArC*H*); ^{13}C NMR (100

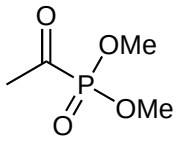
MHz, $CDCl_3$) δ_C 37.0 (d, *J* 131.3 Hz, *C*H_2), 53.2 (d, *J* 6.5 Hz, O*C*H_3), 128.7 (Ar*C*H), 129.0 (Ar*C*H), 133.8 (Ar*C*), 136.4 (d, *J* 2.4 Hz, Ar*C*), 191.8 (d, *J* 6.6 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P (162 MHz, $CDCl_3$) 22.8; HRMS m/z (ESI+) found 251.0442 $[MNa]^+$ (228.0544) $C_{10}H_{13}O_4P$ requires 228.0551.

α-Carbonylphosphonate inhibitor synthesis

General procedure

To a solution of the corresponding acid chloride (52 mmol) at 0 °C, trimethyl phosphite (5.9 mL, 50 mmol) was added dropwise over a period of 1 h and stirred under positive nitrogen pressure. Following addition, the solution was stirred at r.t. for 2 h. Unreacted starting materials were then removed *in vacuo* to yield the product.

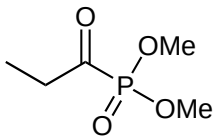
Dimethyl acetylphosphonate (82)



82

Reaction yielded the product **82** as a colourless oil (6.87 g, 45.2 mmol, 90 %). ν_{max}(film)/cm^{-1} 2961 (CH), 1697 (CO), 1450 (CP), 1262 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.40 (3H, d, *J* 5.8 Hz, C*H*$_3$), 3.78 (6H, d, *J* 10.7 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 30.8 (d, *J* 60.3 Hz, *C*H$_3$), 53.9 (d, *J* 7.6 Hz, O*C*H$_3$), 208.2 (d, *J* 170.9 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -1.12; HRMS m/z (ESI+) found 175.0131 [MNa]$^+$ (152.0233) $C_4H_9O_4P$ requires 152.0238.

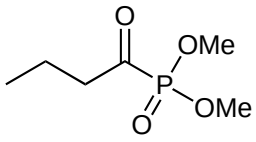
Dimethyl propionylphosphonate (86)



86

Reaction yielded the product **86** as a colourless oil (7.47 g, 45.0 mmol, 90 %). ν_{max}(film)/cm^{-1} 2960 (CH), 1697 (CO), 1458 (CP), 1260 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.08 (3H, dt, *J* 0.7, 7.2 Hz, C*H*$_3$), 2.83 (2H, dq, *J* 1.0, 7.2 Hz, C*H*$_2$), 3.84 (6H, d, *J* 10.8 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 6.24 (d, *J* 4.1 Hz, *C*H$_3$), 37.1 (d, *J* 57.0 Hz, *C*H$_2$), 53.9 (d, *J* 8.1 Hz, O*C*H$_3$), 210.8 (d, *J* 166 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.8; HRMS m/z (ESI+) found 189.0258 [MNa]$^+$ (166.0360) $C_5H_{11}O_4P$ requires 166.0395.

Dimethyl butyrylphosphonate (92)

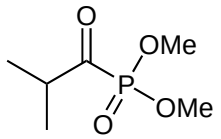


92

Reaction yielded the product **92** as a colourless oil (8.65 g, 47.9 mmol, 96 %). ν_{max}(film)/cm^{-1} 2960 (CH), 1695 (CO), 1463 (CP), 1257 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.95 (3H, dt, *J* 7.5, 0.6 Hz, C*H*$_3$), 1.68 (2H, t, *J* 7.1 Hz, C*H*$_2$), 2.77 (2H, t, *J* 7.1 Hz, C*H*$_2$), 3.87 (6H, d, *J* 10.7 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 10.9 (*C*H$_3$), 13.3 (d, *J* 3.8 Hz, *C*H$_2$), 42.9 (d, *J*

54.0 Hz, *C*CH_2), 51.4 (d, *J* 7.4 Hz, O*C*H_3), 208.0 (d, *J* 164.5 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -3.51; HRMS m/z (ESI+) found 181.0622 $[MNa]^+$ (180.0544) $C_6H_{13}O_4P$ requires 180.0551.

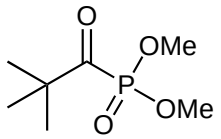
Dimethyl isobutyrylphosphonate (90)



90

Reaction yielded the product **90** as a colourless oil (8.67 g, 47.9 mmol, 96 %). ν_{max}(film)/cm^{-1} 2961 (CH), 1691 (CO), 1464 (CP), 1259 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.17 (6H, d, *J* 6.9 Hz, C*H*$_3$), 3.12 (1H, d.sept, *J* 0.4, 6.9 Hz, C*H*), 3.84 (6H, d, *J* 10.7 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 16.9 (*C*H_3), 41.5 (d, *J* 54.4 Hz, *C*H), 53.8 (d, *J* 7.1 Hz, O*C*H_3), 214.0 (d, *J* 154.7 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.84; HRMS m/z (ESI+) found 203.0440 $[MNa]^+$ (180.0542) $C_6H_{13}O_4P$ requires 180.0551.

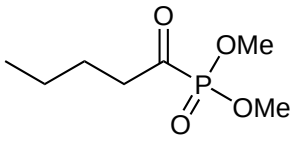
Dimethyl trimethylacetylphosphonate (83)



83

Reaction yielded the product **83** as a colourless oil (6.37 g, 32.8 mmol, 66 %). ν_{max}(film)/cm^{-1} 2960 (CH), 1680 (CO), 1462 (CP), 1257 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.26 (9H, s, C*H*$_3$), 3.83 (6H, d, *J* 10.3 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 26.7 (*C*H$_3$), 54.4 (d, *J* 6.0 Hz, O*C*H$_3$), 80.4 (d, *J* 8.1 Hz, C^{t}Bu); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.63; HRMS m/z (ESI+) found 217.0605 $[MNa]^+$ (194.0707) $C_7H_{15}O_4P$ requires 194.0708.

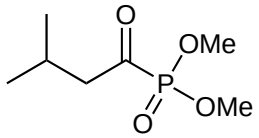
Dimethyl valerylphosphonate (85)



85

Reaction yielded the product **85** as a colourless oil (9.03 g, 46.5 mmol, 93 %). ν_{max}(film)/cm^{-1} 2959 (CH), 1696 (CO), 1463 (CP), 1259 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.89 (3H, t, *J* 7.5 Hz, C*H*$_3$), 1.32 (2H, sex, *J* 7.5 Hz, C*H*$_2$), 1.59 (2H, quin, *J* 7.5 Hz, C*H*$_2$), 2.80 (2H, t, *J* 7.2 Hz, C*H*$_2$), 3.84 (6H, d, *J* 10.7 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 13.7 (*C*H$_3$), 22.0 (*C*H$_2$), 24.3 (d, *J* 3.7 Hz, *C*H$_2$), 43.3 (d, *J* 54.3 Hz, *C*H$_2$), 53.9 (d, *J* 7.2 Hz, O*C*H$_3$), 210.6 (d, *J* 163.8 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.93; HRMS m/z (ESI+) found 217.0593 $[MNa]^+$ (194.0695) $C_7H_{15}O_4P$ requires 194.0708.

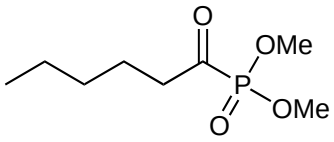
Dimethyl isovalerylphosphonate (89)



89

Reaction yielded the product **89** as a colourless oil (8.81 g, 45.4 mmol, 90 %). ν_{max}(film)/cm^{-1} 2959 (CH), 1696 (CO), 1467 (CP), 1260 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.95 (6H, dd, *J* 6.7, 0.5 Hz, C*H*$_3$) 2.26 (1H, d.sept, *J* 6.7, 0.3 Hz, C*H*), 2.71 (2H, dd, J 6.8, 0.6 Hz, C*H*$_2$), 3.86 (6H, d, *J* 10.7 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 22.4 (*C*H$_3$), 23.6 (d, *J* 3.8 Hz, *C*H), 52.3 (d, *J* 52.2 Hz, *C*H$_2$), 53.9 (d, *J* 7.3 Hz, O*C*H$_3$), 210.4 (d, *J* 163.8 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -1.12; HRMS m/z (ESI+) found 217.0599 [MNa]$^+$ (194.0701) $C_7H_{15}O_4P$ requires 194.0708.

Dimethyl hexanoylphosphonate (93)

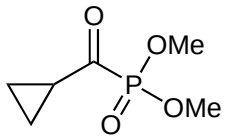


93

Reaction yielded the product **93** as a colourless oil (9.54 g, 45.8 mmol, 92 %). ν_{max}(film)/cm^{-1} 2957 (CH), 1693 (CO), 1458 (CP), 1257 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.89 (3H, t, *J* 6.9 Hz, C*H*$_3$), 1.31 (4H, m, C*H*$_2$), 1.63 (2H, quin, *J* 1.9 Hz, C*H*$_2$), 2.82 (2H, t, *J* 6.8 Hz, C*H*$_2$), 3.86 (6H, d, *J* 10.7 Hz,

OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 13.8 (*C*H$_3$), 21.9 (d, *J* 4.0 Hz, *C*H$_2$), 22.3 (*C*H$_2$), 31.0 (s, *C*H$_2$) 43.6 (d, *J* 57.2 Hz, *C*H$_2$), 53.9 (d, *J* 7.9 Hz, O*C*H$_3$), 210.7 (d, *J* 165.0 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.88; HRMS m/z (ESI+) found 209.0939 $[MH]^+$ (208.0861) $C_8H_{17}O_4P$ requires 208.0864.

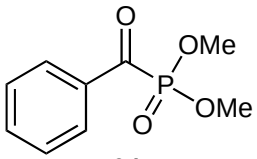
Dimethyl (cyclopropanecarbonyl)phosphonate (87)



87

Reaction yielded the product **87** as a light pink oil (8.13 g, 45.6 mmol, 91 %). ν_{max}(film)/cm^{-1} 2959 (CH), 1673 (CO), 1449 (CP), 1263 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.13 (2H, m, C*H*$_2$), 1.26 (2H, m, C*H*$_2$), 2.60 (1H, tt, *J* 4.6, 3.3 Hz, C*H*), 3.84 (6H, d, *J* 10.8 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 14.1 (*C*H$_2$), 22.5 (d, *J* 72.8 Hz, *C*H), 53.9 (d, *J* 7.2 Hz, O*C*H$_3$), 210.0 (d, *J* 174.9 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.77; HRMS m/z (ESI+) found 201.0292 $[MNa]^+$ (178.0394) $C_6H_{11}O_4P$ requires 178.0395.

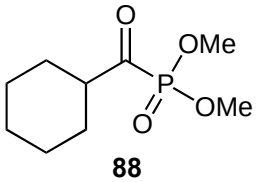
Dimethyl benzoylphosphonate (91)



91

Reaction yielded the product **91** as a yellow oil (9.94 g, 4.40 mmol, 88 %). ν_{max}(film)/cm^{-1} 2958 (ArCH), 1711 (CO), 1655 ($CC_{aromatic}$), 1450 (CP), 1260 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 3.89 (6H, d, *J* 10.7 Hz, O*CH*$_3$), 7.49 (2H, m, Ar*CH*), 7.62 (1H, tt, *J* 7.4, 1.2 Hz, Ar*CH*), 8.22 (2H, dd, *J* 8.5, 1.2 Hz, Ar*CH*); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 54.2 (d, *J* 7.2 Hz, O*C*H$_3$), 128.9 (Ar*C*H), 129.8 (d, *J* 1.8 Hz, Ar*C*H), 135.0 (Ar*C*H), 198.2 (d, *J* 174.7 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P 0.68; HRMS m/z (ESI+) found 237.0278 [MNa]$^+$ (214.0380) $C_9H_{11}O_4P$ requires 214.0395.

Dimethyl (cyclohexanecarbonyl)phosphonate (88)

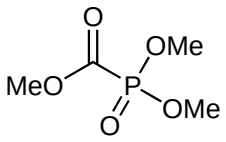


88

Reaction yielded the product **88** as a colourless oil (9.83 g, 44.6 mmol, 89 %). ν_{max}(film)/cm^{-1} 2928, (CH), 1685 (CO), 1450 (CP), 1256 (PO); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.20 (5H,

m, C*H*$_2$), 1.59 (1H, d, *J* 12.7 Hz, C*H*$_2$), 1.71 (2H, m, C*H*$_2$), 1.90 (2H, m, C*H*$_2$), 2.84 (1H, m, C*H*), 3.77 (6H, d, *J* 10.8 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 25.2 (*C*H$_2$), 25.5 (*C*H$_2$), 27.1 (*C*H$_2$), 50.9 (d, *J* 53.1 Hz, *C*H), 53.8 (d, *J* 7.7 Hz, O*C*H$_3$), 213.0 (d, *J* 115.8 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -0.84; HRMS m/z (ESI+) found 243.0760 $[MNa]^+$ (220.0862) $C_9H_{17}O_4P$ requires 220.0864.

Dimethyl (methoxyacetyl)phosphonate (84)



84

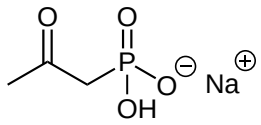
Reaction yielded the product **84** as a colourless oil (6.21 g, 36.9 mmol, 74 %). ν_{max}(film)/cm^{-1} 2961 (CH), 1722 (CO), 1451 (CP), 1230 (P=O), 1183 (P-O); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 3.80 (3H, d, *J* 1.1 Hz, OC*H*$_3$), 3.85 (6H, d, *J* 11.2 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 52.4 (d, *J* 5.3 Hz, O*C*H$_3$), 54.5 (d, *J* 6.3 Hz, O*C*H$_3$), 166.6 (d, *J* 269.7 Hz, *C*O); ^{31}P NMR (162 MHz, $CDCl_3$) δ_P -2.54; HRMS m/z (ESI+) found 169.0263 $[MH]^+$ (168.0185) $C_4H_{10}O_5P$ requires 168.0188.

Phosphonic acid inhibitors

General procedure

The corresponding phosphonate (2 mmol), allyltrimethylsilane (0.47 mL, 3 mmol) and bromotrimethylsilane (1.33 mL, 10 mmol) in 1,2-dichloroethane (10 mL) were stirred at r.t. for 2 h under argon before concentration *in vacuo* The residue was re-dissolved in 1,2-dichloroethane (5 mL) and acetonitrile (5 mL) and concentrated *in vacuo*. The residue was dissolved in ethanol:water (1:1, 10 mL). Following stirring for 15 min, the mixture was concentrated *in vacuo* before dissolving in water (10 mL). The pH of the solution was adjusted to 7.0 by addition of sodium hydroxide (1 M). The solution was lyophilised to yield the product as a powder.

(2-Oxopropyl)phosphonic acid sodium salt (105)

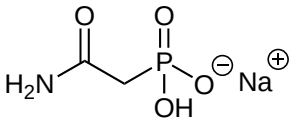


105

Dimethyl (2-oxopropyl)phosphonate **68** (80 mg, 0.48 mmol) was used and other reagents scaled accordingly. Reaction yielded the product **105** as a white powder (72 mg, 0.45 mmol, 94 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2920 (CH), 1683 (CO), 1457 (CP), 1247 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 2.23

(1H, s, C*H*$_3$), 2.89 (2H, d, *J* 20.2 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 30.5 (*C*H_3), 48.4 (d, *J* 106.8, *C*H_2), 211.8 (*C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 10.7; HRMS m/z (ESI-) found 137.0004 $[M]^-$ $C_3H_6O_4P^-$ requires 137.0009.

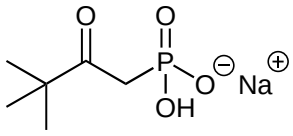
(2-Amino-2-oxoethyl)phosphonic acid sodium salt (106)



106

Dimethyl (2-amino-2-oxoethyl)phosphonate **69** (48 mg, 0.29 mmol) was used and other reagents scaled accordingly. Reaction yielded the product **106** as a white powder (43 mg, 0.27 mmol, 93 %). ^{1}H NMR (400 MHz, D_2O) δ_H 2.45 (2H, d, *J* 18.8 Hz, C*H*$_2$) ^{13}C NMR (100 MHz, D_2O) δ_C 38.3 (d, *J* 111 Hz, *C*H_2), 176.3 (*C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 11.4.

(3,3-Dimethyl-2-oxobutyl)phosphonic acid sodium salt (107)

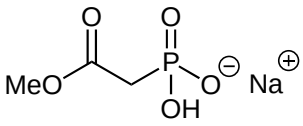


107

Dimethyl (3,3-dimethyl-2-oxobutyl)phosphonate **71** (71 mg, 0.34 mmol) was used and other reagents scaled accordingly. Reaction yielded the product **107** as a white powder (61 mg,

0.31 mmol, 91 %). M.p. >240 °C. ^{1}H NMR (400 MHz, D_2O) δ_H 1.12 (9H, s, ^{t}Bu), 2.98 (2H, d, *J* 20.2 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 25.8 (^{t}Bu); ^{31}P NMR (162 MHz, D_2O) δ_P 11.7. HRMS m/z (ESI-) found 179.0473 $[M]^-$ $C_8H_{12}O_4P^-$requires 179.0479.

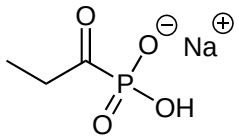
(2-Methoxy-2-oxoethyl)phosphonic acid sodium salt (110)



110

Reaction yielded the product **110** as a white powder (310 mg, 1.76 mmol, 88 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1197 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 2.72 (2H, d, *J* 20.4 Hz, C*H*$_2$), 3.63 (3H, s, OCH_3); ^{13}C NMR (100 MHz, D_2O) δ_C 37.0 (d, *J* 115.3 Hz, *C*H_2), 52.4 (O*C*H_3), 172.2 (d, *J* 6.4 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 11.9; HRMS m/z (ESI-) found 152.9954 $[M]^-$ $C_3H_7O_5P^-$ requires 152.9958.

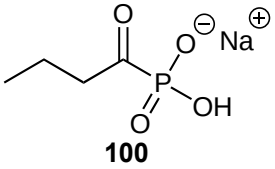
Propionylphosphonic acid sodium salt (97)



97

The reaction yielded the product **97** as a white powder (311 mg, 1.94 mmol, 97 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1673 (CO), 1459 (CP); ^{1}H NMR (400 MHz, D_2O) δ_H 0.94 (3H, t, *J* 7.3 Hz, C*H*$_3$), 2.77 (2H, q, *J* 7.3 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 6.31 (d, *J* 3.5 Hz, *C*H$_3$), 36.1 (d, *J* 46.0 Hz, *C*H$_2$), 225.9 (d, *J* 159.5 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -1.49.

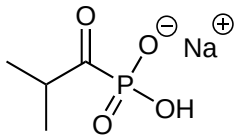
Butyrylphosphonic acid sodium salt (100)



100

The reaction yielded the product **100** as a white powder (302 mg, 1.74 mmol, 87 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1674 (CO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.81 (3H, t, *J* 7.5 Hz, C*H*$_3$), 1.47 (2H, sex, *J* 7.3 Hz, C*H*$_2$), 2.70 (2H, t, *J* 7.2 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.1 (*C*H$_3$), 16.2 (d, *J* 2.5 Hz, *C*H$_2$), 44.7 (d, *J* 40.9 Hz, *C*H$_2$), 229.5 (d, *J* 155.9 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.53; HRMS m/z (ESI-) found 151.0165 [M]$^-$ $C_4H_9O_4P^-$ requires 151.0166.

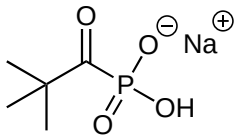
Isobutyrylphosphonic acid sodium salt (98)



98

The reaction yielded the product **98** as a as a white powder (284 mg, 1.63 mmol, 82 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2926 (CH), 1665 (CO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.00 (6H, d, *J* 7.0 Hz, C*H*$_3$), 3.20 (1H, sep, *J* 7.1 Hz, C*H*); ^{13}C NMR (100 MHz, D_2O) δ_C 17.2 (*C*H$_3$), 39.8 (d, *J* 43.2 Hz, *C*H), 230.6 (d, *J* 153.8 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -1.41; HRMS m/z (ESI-) found 151.0160 $[M]^-$ $C_4H_9O_4P$ requires 151.0166.

Trimethylacetylphosponic acid sodium salt (99)

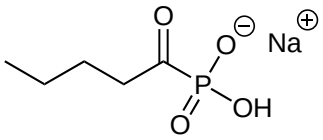


99

Dimethyl (trimethylacetyl)phosphonate **83** (139 mg, 0.72 mmol) was used and other reagents scaled accordingly The reaction yielded the product **99** as a white powder (118 mg, 0.63 mmol, 87 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 1658 (CO), 1248 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.21 (9H, s, C*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 25.4 (*C*H$_3$), 50.0 (d, *J* 41.6 Hz, *C*^{t}Bu), 231.6 (d, *J* 145.4 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -1.51;

HRMS m/z (ESI-) found 165.0316 $[M]^-$ $C_5H_{10}O_4P^-$ requires 165.0322.

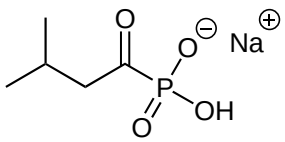
Valerylphosphonic acid sodium salt (101)



101

The reaction yielded the product **101** as a white powder (294 mg, 1.56 mmol, 78 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2920 (CH), 1664 (CO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.80 (3H, t, *J* 7.5 Hz, C*H*$_3$), 1.22 (2H, sex, *J* 7.5 Hz, C*H*$_2$), 1.42 (2H, quin, *J* 7.5 Hz, C*H*$_2$), 2.72 (2H, t, *J* 7.3 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.2 (*C*H$_3$), 21.8 (*C*H$_2$), 24.8 (d, *J* 2.6 Hz, *C*H$_2$), 42.5 (d, *J* 40.3 Hz, *C*H$_2$), 230.5 (d, *J* 154.8 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -0.20; HRMS m/z (ESI-) found 165.0323 $[M]^-$ $C_5H_{11}O_4P^-$ requires 165.0322.

Isovalerylphosphonic acid sodium salt (109)

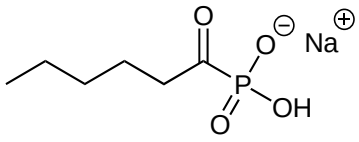


109

The reaction yielded the product **109** as a white powder (310 mg, 1.65 mmol, 83 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2927 (CH), 1663 (CO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.81 (6H, d, *J*

6.7 Hz, C*H*$_3$), 2.05 (1H, sep, *J* 6.7 Hz, C*H*), 2.60 (2H, d, *J* 6.7 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 21.9 (*C*H$_3$), 23.4 (d, *J* 2.9 Hz, *C*H), 51.7 (d, *J* 39.7 Hz, *C*H$_2$), 229.7 (d, *J* 155.4 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -0.54; HRMS m/z (ESI+) found 165.0314 $[M]^-$ $C_5H_{10}O_4P^-$ requires 165.0322.

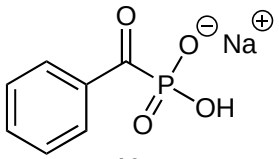
Hexanoylphosphonic acid sodium salt (102)



102

The reaction yielded the product **102** as a white powder (353 mg, 1.75 mmol, 88 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 3421 (OH), 1634 (CO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.79 (3H, t, *J* 7.0 Hz, C*H*$_3$), 1.15 – 1.28 (4H, m, C*H*$_2$), 1.47 (2H, quin, *J* 7.1 Hz, C*H*$_2$), 2.73 (2H, t, *J* 7.4 Hz, C*H*$_2$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.2 (*C*H$_3$), 21.8 (*C*H$_2$), 22.2 (d, *J* 2.9 Hz, *C*H$_2$), 30.7 (*C*H$_2$), 42.7 (d, *J* 42.0 Hz, *C*H$_2$), 228.3 (d, *J* 156.9 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P -0.88; HRMS m/z (ESI-) found 179.0484 $[M]^-$ $C_6H_{13}O_4P^-$ requires 179.0479.

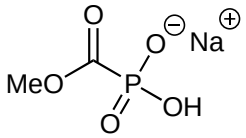
Benzoylphosphonic acid sodium salt (104)



104

The reaction yielded the product **104** as a white powder (252 mg, 1.21 mmol, 61 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1633 (CO), 1456 (CP), 1224 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 7.46 (2H, t, *J* 7.7 Hz, ArC*H*), 7.57 (1H, t, *J* 7.4 Hz, ArC*H*), 8.14 (2H, d, *J* 7.9 Hz, ArC*H*); ^{13}C NMR (100 MHz, D_2O) δ_C 128.4 (Ar*C*H), 129.2 (Ar*C*H), 133.4 (Ar*C*H), 137.4 (d, *J* 43.1 Hz, Ar*C*), 216.0 (d, *J* 160.1 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.56; HRMS m/z (ESI-) found 185.0010 $[M]^-$ $C_7H_7O_4P$ requires 185.0009.

Methoxyacetylphosphonic acid sodium salt (103)

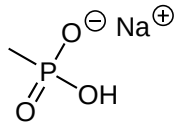


103

The reaction yielded the product **103** as a white powder (294 mg, 1.81 mmol, 91 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1656 (CO), 1251 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 3.60 (3H, s, OCH_3); ^{13}C NMR (100 MHz, D_2O) δ_C 50.6 (d, *J* 4.0 Hz, OCH_3), 179.1 (d, *J* 215.4 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P

-3.47; HRMS m/z (ESI-) found 138.9798 $[M]^-$ $C_2H_5O_5P^-$ requires 138.9802.

Methylphosphonic acid sodium salt (95)



95

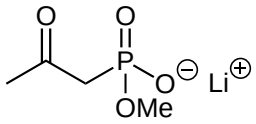
The reaction yielded the product **95** as a white powder (201 mg, 1.72 mmol, 86 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2928 (CH); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.14 (3H, d, *J* 16.1 Hz, C*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 13.3 (d, *J* 133.4 Hz, *C*H$_3$); ^{31}P NMR (160 MHz, $CDCl_3$) δ_P 23.0.

Monomethyl ester inhibitor synthesis

General procedure

The corresponding phosphonate (2 mmol) was dissolved in 2-hexanone (10 mL), to which lithium bromide (124 mg, 2.4 mmol) was added. This was stirred at 80 °C for 1 h. The resulting precipitate was filtered *in vacuo* and washed with diethyl ether before collection of the precipitate which required no further purification.

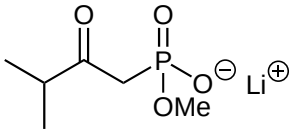
Lithium methyl (2-oxopropyl)phosphonate (244)



124

Reaction yielded the product **124** as a white powder (157 mg, 0.99 mmol, 50 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2974 (CH), 1690 (CO), 1463 (CP), 1260 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 2.25 (3H, s, C*H*$_3$), 3.01 (2H, dd, *J* 21.8, 0.5 Hz, C*H*$_2$), 3.52 (d, *J* 10.9 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 30.7 (*C*H$_3$), 43.8 (d, *J* 113.7 Hz, *C*H$_2$), 51.8 (d, *J* 5.5 Hz, *C*H$_2$), 209.3 (d, *J* 5.5 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 15.8; HRMS m/z (ESI-) found 151.0167 [M]$^-$ $C_4H_9O_4P$ requires 151.0166.

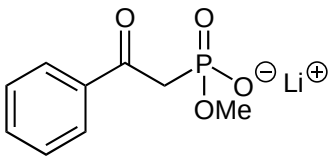
Lithium methyl (3-methyl-2-oxobutyl)phosphonate (125)



125

Dimethyl (3-methyl-2-oxobutyl)phosphonate **75** (125 mg, 0.64 mmol) was used and other reagents scaled accordingly. Reaction yielded the product **125** as a white powder (81 mg, 0.44 mmol, 68 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1682 (CO), 1229 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.99 (6H, d, *J* 7.0 Hz, C*H*$_3$), 2.85 (1H, sep, *J* 6.9 Hz, C*H*), 3.03 (2H, d, *J* 21.4 Hz, C*H*$_2$), 3.49 (3H, d, *J* 10.9 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 17.4 (*C*H$_3$), 41.1 (d, *J* 116.8 Hz, *C*H$_2$) 41.4 (*C*H), 51.8 (d, *J* 5.9 Hz, O*C*H$_3$), 215.0 (d, *J* 5.9 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 16.3; HRMS m/z (ESI-) found 179.0474 [M]$^-$ $C_6H_{12}O_4P^-$ requires 179.0479.

Lithium methyl (2-oxo-phenylethyl)phosphonate (126)

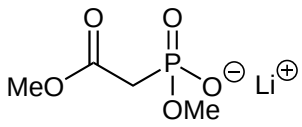


126

Dimethyl (2-oxo-phenylethyl)phosphonate **70** (246 mg, 1.08 mmol) was used and other reagents scaled accordingly.

Reaction yielded the product **126** as a white powder (121 mg, 0.55 mmol, 52 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2927 (CH), 1685 (CO), 1224 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 3.44 (3H, d, *J* 10.9 Hz, C*H*$_3$), 3.52 (2H, d, *J* 21.7 Hz, C*H*$_2$), 7.48 (2H, t, *J* 7.8 Hz, ArC*H*), 7.61 (1H, t, *J* 7.5 Hz, ArC*H*), 7.94 (2H, d, *J* 7.6 Hz, ArC*H*); ^{13}C NMR (100 MHz, D_2O) δ_C 39.3 (d, *J* 116.2 Hz, *C*H$_2$), 52.0 (d, *J* 5.9 Hz, O*C*H$_3$), 128.7 (Ar*C*H), 129.1 (Ar*C*H), 134.2 (Ar*C*H), 136.4 (Ar*C*), 199.1 (d, *J* 6.3 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 15.9; HRMS m/z (ESI-) found 213.0316 [M]$^-$ $C_9H_{10}O_4P^-$ requires 213.0322.

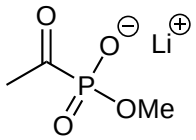
Lithium methyl (2-methoxy-2-oxoethyl)phosphonate (129)



129

Reaction yielded the product **129** as a white powder (242 mg, 1.51 mmol, 76 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2927 (CH), 1703 (CO), 1202 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 2.80 (2H, d, *J* 20.1 Hz, C*H*$_2$), 3.51 (3H, d, *J* 11.0 Hz, OC*H*$_3$), 3.63 (3H, d, *J* 0.5 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 34.2 (d, *J* 120.5 Hz, *C*H$_2$), 52.0 (d, *J* 5.7 Hz, O*C*H$_3$), 52.6 (O*C*H$_3$), 171.0 (d, *J* 6.4 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 16.2; HRMS m/z (ESI-) found 167.0109 [M]$^-$ $C_4H_8O_5P^-$ requires 167.0115.

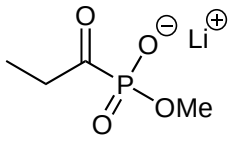
Lithium methyl (acetyl)phosphonate (114)



114

The reaction yielded the product **114** as a white powder (252 mg, 1.74 mmol, 87 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2924 (CH), 1667 (CO), 1254 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 2.34 (3H, d, *J* 4.7 Hz, C*H*$_3$), 3.53 (3H, d, *J* 10.4 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 30.1 (d, *J* 49.8 Hz, *C*H$_3$), 52.8 (d, *J* 6.0 Hz, O*C*H$_3$), 219.8 (d, *J* 164.9 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.10; HRMS m/z (ESI-) found 137.0008 $[M]^-$ $C_3H_7O_4P^-$ requires 137.0009.

Lithium methyl (propionyl)phosphonate (115)

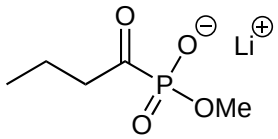


115

The reaction yielded the product **115** as a white powder (306 mg, 1.94 mmol, 97 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1627 (CO), 1255 (PO), 1223 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.96 (3H, dt, *J* 6.8, 0.4 Hz, C*H*$_3$), 2.77 (2H, q, *J* 7.2

Hz, C*H*$_2$), 3.54 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 6.0 (d, *J* 3.5 Hz, *C*H$_3$), 36.6 (d, *J* 47.9 Hz, *C*H$_2$), 52.8 (d, J 6.2 Hz, O*C*H$_3$), 222.2 (d, *J* 160.6 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.22; HRMS m/z (ESI-) found 151.0165 $[M]^-$ $C_4H_9O_4P$ requires 151.0166.

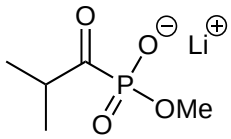
Lithium methyl (butyryl)phosphonate (118)



118

The reaction yielded the product **118** as a white powder (246 mg, 1.43 mmol, 72 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1683 (CO), 1255 (PO), 1214 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.82 (3H, t, *J* 7.5 Hz, C*H*$_3$), 1.54 (2H, appt sex, *J* 7.1 Hz, C*H*$_2$), 2.73 (2H, t, *J* 6.9 Hz, C*H*$_2$), 3.54 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 12.8 (*C*H$_3$), 15.9 (d, *J* 3.2 Hz, *C*H$_2$), 45.0 (d, *J* 45.0 Hz, *C*H$_2$), 52.8 (d, J 5.9 Hz, O*C*H$_3$), 222.2 (d, *J* 160.9 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.05; HRMS m/z (ESI+) found 189.0306 $[MNa]^+$ (166.0408) $C_5H_{11}O_4P$ requires 166.0400.

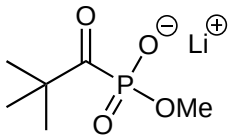
Lithium methyl (isobutyryl)phosphonate (116)



116

The reaction yielded the product **116** as a white powder (266 mg, 1.54 mmol, 77 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2923 (CH), 1657 (CO), 1253 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.04 (6H, d, *J* 6.4 Hz, C*H*$_3$), 3.11 (1H, sep, *J* 6.9 Hz, C*H*), 3.53 (3H, d, *J* 10.5 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 16.7 (*C*H$_3$), 40.7 (d, *J* 46.2 Hz, *C*H$_2$), 52.8 (d, J 6.0 Hz, O*C*H$_3$), 225.5 (d, *J* 155.2 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.05; HRMS m/z (ESI-) found 165.0323 [M]$^-$ $C_5H_{10}O_4P^-$ requires 165.0322.

Lithium methyl (pivaloyl)phosphonate (117)

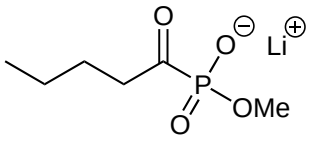


117

The reaction yielded the product **117** as a white powder (219 mg, 1.18 mmol, 59 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2915 (CH), 1655 (CO), 1253 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.16 (9H, s, C*H*$_3$), 3.52 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 24.9 (s, *C*H$_3$), 45.3 (d, *J* 44.5 Hz, *C*CH$_3$), 52.8 (d, *J* 6.4 Hz, OCH$_3$), 225.8 (d, *J* 145.6 Hz, *C*H$_2$); ^{31}P NMR (162

MHz, D_2O) δ_P -0.21; HRMS m/z (ESI-) found 179.0472 $[M]^-$ $C_6H_{12}O_4P^-$ requires 179.0479.

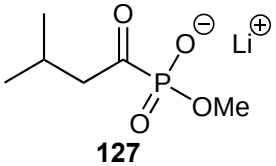
Lithium methyl (valeryl)phosphonate (119)



119

The reaction yielded the product **119** as a white powder (298 mg, 1.60 mmol, 80 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1667 (CO), 1255 (PO), 1215 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.81 (3H, t, *J* 7.3 Hz, C*H*$_3$), 1.24 (2H, appt sex, *J* 7.5 Hz, C*H*$_2$), 1.50 (2H, quin, *J* 7.5 Hz, C*H*$_2$), 2.75 (2H, t, *J* 7.3 Hz, C*H*$_2$), 3.54 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.0 (*C*H$_3$), 21.6 (*C*H$_2$), 24.4 (d, *J* 3.0 Hz, *C*H$_2$), 42.8 (d, *J* 45.8 Hz, *C*H$_2$), 52.8 (d, *J* 6.2 Hz, O*C*H$_3$), 222.3 (d, *J* 160.0 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.06; HRMS m/z (ESI-) found 179.0480 $[M]^-$ $C_6H_{12}O_4P^-$ requires 179.0479.

Lithium methyl (isovaleryl)phosphonate (127)

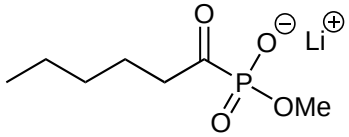


127

The reaction yielded the product **127** as a white powder (248 mg, 1.33 mmol, 67 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2927

(CH), 1667 (CO), 1255 (PO), 1214 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.84 (6H, d, *J* 6.6 Hz, C*H*$_3$), 2.12 (1H, sep, *J* 6.7 Hz, C*H*), 2.62 (2H, d, *J* 7.0 Hz, C*H*$_2$), 3.54 (3H, d, *J* 10.5 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 21.7 (*C*H$_3$), 23.6 (d, *J* 2.8 Hz, *C*H), 52.0 (d, *J* 44.4 Hz, *C*H$_2$), 52.8 (d, *J* 6.3 Hz, *C*H$_2$), 222.1 (d, *J* 160.3 Hz, *C*H$_2$); ^{31}P NMR (162 MHz, D_2O) δ_P -0.54; HRMS m/z (ESI-) found 179.0471 $[M]^-$ $C_6H_{12}O_4P^-$ requires 179.0479.

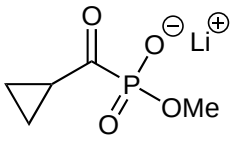
Lithium methyl (hexanoyl)phosphonate (128)



128

The reaction yielded the product **128** as a white powder (363 mg, 1.80 mmol, 90 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2923 (CH), 1667 (CO), 1254 (PO), 1216 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 0.79 (3H, t, *J* 6.9 Hz, C*H*$_3$), 1.15 – 1.30 (4H, m C*H*$_2$), 1.52 (2H, quin, *J* 7.0 Hz, C*H*$_2$), 2.74 (2H, t, *J* 7.3 Hz, C*H*$_2$), 3.53 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.1 (*C*H$_3$), 21.7 (*C*H$_2$), 21.9 (d, *J* 3.0 Hz, *C*H$_2$), 30.5 (*C*H$_2$), 43.0 (d, *J* 45.9 Hz, *C*H$_2$), 52.8 (d, *J* 6.2 Hz, O*C*H$_3$), 222.4 (d, *J* 160.5 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) 0.05; HRMS m/z (ESI-) found 193.0633 $[M]^-$ $C_7H_{14}O_4P^-$ requires 193.0635.

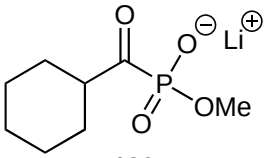
Lithium methyl (cyclopropanecarbonyl)phosphonate (120)



120

The reaction yielded the product **120** as a white powder (238 mg, 1.40 mmol, 70 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 2852 (CH), 1657 (CO), 1223 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.07 – 1.16 (4H, m, C*H*$_2$), 2.53 – 2.60 (1H, m, C*H*), 3.54 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 13.7 (*C*H$_2$), 22.2 (d, *J* 61.5 Hz, *C*H), 52.8 (d, *J* 5.9 Hz, O*C*H$_3$), 221.1 (d, *J* 168.5 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.77; HRMS m/z (ESI-) found 163.0169 [M]$^-$ $C_5H_8O_4P^-$ requires 163.0166.

Lithium methyl (cyclohexanecarbonyl)phosphonate (123)

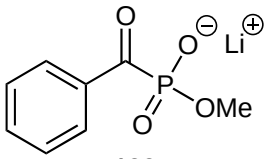


123

The reaction yielded the product **123** as a white powder (215 mg, 1.01 mmol, 51 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2928 (CH), 1661 (CO), 1255 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 1.05

– 1.33 (5H, m, C*H*$_2$), 1.55 – 1.63 (1H, m, C*H*$_2$), 1.66 – 1.75 (2H, m, C*H*$_2$), 1.80 – 1.88 (2H, m, C*H*$_2$), 2.90 (1H, appt tt, *J* 3.4, 11.1 Hz, C*H*), 3.53 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 25.1 (*C*H$_2$), 25.3 (*C*H$_2$), 27.3 (*C*H$_2$), 50.7 (d, *J* 44.3 Hz, *C*H), 52.7 (d, *J* 6.0 Hz, O*C*H$_3$), 224.8 (d, *J* 156.3 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 0.07; HRMS m/z (ESI-) found 205.0643 $[M]^-$ $C_8H_{14}O_4P^-$ requires 205.0635.

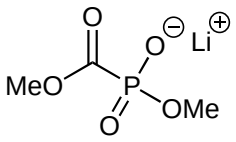
Lithium methyl (benzoyl)phosphonate (122)



122

The reaction yielded the product **122** as a white powder (293 mg, 1.48 mmol, 74 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2916 (CH), 1635 (CO), 1257 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 3.61 (3H, d, *J* 10.7 Hz, OC*H*$_3$), 7.53 (2H, tt, *J* 0.7, 7.6 Hz, ArC*H*), 7.66 (1H, tt, *J* 1.8, 7.4 Hz, ArC*H*), 8.12 (2H, dd, *J* 8.6, 1.2 Hz, ArC*H*); ^{13}C NMR (100 MHz, D_2O) δ_C 53.1 (d, *J* 6.4 Hz, O*C*H$_3$), 128.9 (Ar*C*H*)*, 129.3 (Ar*C*H), 134.8 (Ar*C*H), 208.1 (d, *J* 160.3 Hz, *C*O); ^{31}P NMR (162 MHz, D_2O) δ_P 1.05; HRMS m/z (ESI-) found 199.0169 $[M]^-$ $C_8H_8O_4P^-$ requires 199.0166.

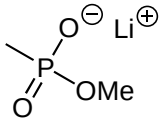
Lithium methyl (methoxycarbonyl)phosphonate (121)



121

The reaction yielded the product **121** as a white powder (291 mg, 1.82 mmol, 91 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2922 (CH), 1719 (CO), 1256 (PO); ^{1}H NMR (400 MHz, D_2O) δ_H 3.60 (3H, d, J 11.1 Hz, OC*H*$_3$), 3.72 (3H, d, J 1.0 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, D_2O) δ_C 51.9 (d, *J* 4.6 Hz, O*C*H$_3$), 53.2 (d, *J* 5.7 Hz, O*C*H$_3$); ^{31}P NMR (162 MHz, D_2O) δ_P -3.31; HRMS m/z (ESI-) found 152.9958 $[M]^-$ $C_3H_6O_5P^-$ requires 152.9958.

Lithium methyl (methyl)phosphonate (113)



113

The reaction yielded the product **113** as a white powder (239 mg, 1.80 mmol, 90 %). M.p. >240 °C. ν_{max}(nujol)/cm^{-1} 2931 (CH); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.19 (3H, d, *J* 16.4 Hz, C*H*$_3$), 3.47 (3H, d, *J* 10.6 Hz, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 10.4 (d, *J* 136.8 Hz, *C*H$_3$), 51.0 (d, *J* 5.1 Hz, O*C*H$_3$); ^{31}P NMR (160 MHz, $CDCl_3$) δ_P 28.3; HRMS m/z (ESI+) found 109.0063 $[M]^-$ $C_2H_6O_3P^-$ requires 109.0060.

4.3 Acetate kinase recombinant purification

DNA miniprep

Isolation of genomic DNA from *C. autoethanogenum* was performed using a Genomic DNA Extraction Kit (Sigma-Aldrich).

Cells (5 mL) were harvested from an overnight culture growth and centrifuged for 2 min at 12000 rpm. The supernatant was discarded and pellet re-suspended in lysozyme solution (200 µL, 107 mg mL^{-1}) and incubated for 30 min at 37 °C, followed by addition of RNase A solution (20 µL) and incubation for 2 min at r.t. Proteinase K (20 µL) and lysis solution C (200 µL) were added, vortexed for 15 sec and incubated for 10 min at 55 °C.

To a GenElute Miniprep Binding Column seated in a collection tube (2 mL) was added Column Preparation Solution (500 µL) and centrifuged at 12000 rpm for 1 min, before discarding the flow through. Ethanol (200 µL) was added to the cell lysate and vortex for 10 sec before transferring to the binding column and centrifuging at >6500 rpm for 1 min, discarding the flow through. Wash solution 1 (500 µL) was added to the column and centrifuged for 1 min at >6500 rpm, discarding

the flow through. Wash solution (500 µL) was added to the column before centrifuging at 12000 rpm for 3 min. The binding column was then transferred to a new collection tube (1 mL) before addition of water (200 µL), incubation for 1 min and centrifuging at >6500 rpm for 1 min to elute genomic DNA.

PCR amplification

PCR amplification of genomic DNA was performed using two step PCR. Firstly, primers coding for the forwards and reverse sequences had to be generated. The forward sequence used was 5'-CCCCGGATCCATGAAAATATTAGTAGTAAACTGT-3' and the reverse primer sequence was 5'-GGGGCCCGGGTTTATTTTCAACTATTTCTTTTG-3'. To create the PCR master mix, water (167.5 µL), 5x Phusion HF Buffer (50 µL), deoxyribose nucleotide triphosphates (5 µL), forward primer (10 µL) and reverse primer (10 µL) were combined. In five tubes, master mix (48.5 µL), extracted genomic DNA (1 µL) and polymerase (0.5 µL) were added and underwent PCR with the programme below to amplify genomic DNA (**Table 9**):

Temperature / °C	Time / s	Number of repeats
98	30	
98	10	11
x-3	30	
72	45	
98	10	26
x	30	
72	45	
72	600	
15	∞	

Table 9 – PCR programme where x is the annealing temperature, in this case 50.2 °C.

Following successful PCR, the pThydA plasmid underwent restriction digest to generate the pT plasmid (provided by Girbal *et al.*).[83] To plasmid DNA amplified by PCR (15 µL) was added Cfr9I restriction enzyme (5 µL), 10x Cfr9I buffer (5 µL) and water (20 µL). This was incubated for 16 h at 37 °C before addition of BamHI restriction enzyme (5 µL) and further incubation for 2 h at r.t. to yield linearized plasmid DNA.

To the linearized plasmid DNA (100 ng) was added genomic insert DNA (300 ng), 10x T4 ligase buffer (2 µL), T4 DNA ligase (1 µL) and water (20 µL). This is incubated at 22 °C for

20 min. Chemically competent *E. coli* BL21 DE3 cells were thawed and ligation mix (5 µL) was added to competent cells (50 µL) before mixing by tube inversion. The cells were incubated on ice for 20 min before heatshock treatment for 90 s at 42 °C, followed by addition of LB broth (500 µL) and incubation for 60 min at 37 °C with shaking. Samples (50 µL, 100 µL and 200 µL) were plated on plates containing ampicillin and incubated for 16 h at 37 °C.

Plasmid miniprep

Isolation of Plasmid DNA from *E. coli* was performed using a GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich).

Colonies from these sample plates were picked and incubated in liquid media for 16 h at 37 °C. Cells were harvested by centrifugation for 1 min at 12000 rpm before resuspension in Resuspension Solution (200 µL) and lysed using Lysis Solution (200 µL). To mix, the tube was inverted until mixture became clear and viscous whereupon Neutralisation Solution (350 µL) was added and mixed by inversion. The mixture was centrifuged at 12000 rpm for 10 min. The GenMiniprep Binding Column was loaded into microcentrifuge tube (2 mL) and Column Preparation Solution (500 µL) was added before centrifugation at 12000 rpm for 1 min. Flowthrough was

discarded before addition of lysate and centrifugation at 12000 rpm for 1 min. Flowthrough was discarded before addition of Washing Solution (750 µL) and centrifugation at 12000 rpm for 1 min. Flowthrough was discarded before further centrifugation at 12000 rpm for 2 min. Water (50 µL) was added before incubating for 5 min, followed by centrifugation at 12000 rpm for 1 min to elute plasmid.

Following plasmid miniprep, plasmid (7 µL) was added to BamHI FastDigest enzyme (0.5 µL), SmaI FastDigest enzyme (0.5 µL), Buffer Green FastDigest (2 µL) and water (10 µL). This was incubated at 37 °C for 20 min on agarose gel, showing the successful ligation of plasmid insert (**Figure 14**), which was sent for sequencing to confirm ligation.

Recombinant protein expression

To purify acetate kinase, affinity chromatography was used using the native Strep-Tag® system in the pT plasmid system. A fresh colony of *E. coli* BL21 DE3 containing the plasmid with the correctly ligated insert was used to inoculate LB media (10 mL) which contained ampicillin (20 µL, 100 µg ml^{-1}). Following aerobic growth at 37 °C for 16 h with shaking at 180 rpm, LB media (250 mL) containing ampicillin (500 µL, 100 µg ml^{-1}) was inoculated and growth continued until an OD of at least

1.0 at 600 nm was achieved (approximately 16 h). Cells were harvested and purified according to manufacturer instructions of Strep-Tactin-Sepharose. To disrupt cells, BugBuster® protein extraction reagent (10 mL), DNase (1 mg), lysozyme (1 mg) and avidin (1 mg) were added and incubated for 20 min at r.t. with shaking at 15 rpm. The mixture was centrifuged and supernatant removed, to be used for affinity chromatography using the Strep-Tactin-Sepharose column, with a sample being taken for analysis.

ACK purification

A column was loaded with Strep-Tactin-Sepharose (50 % suspension, 2 mL) and washed with washing buffer (5 x 1 mL, Tris-HCl pH 8.0 100 mM & NaCl 150 mM) before loading the cell lysate. The flowthrough was collected before washing the column with washing buffer (5 x 1 mL, Tris-HCl pH 8.0 100 mM & NaCl 150 mM). The protein was eluted with elution buffer (6 x 0.5 mL, Tris-HCl pH 8.0 100 mM, NaCl 150 mM & desthiobiotin 2.5 mM) to elute acetate kinase (approximately 0.5 mg/mL by Bradford Assay).

SDS & Western Blot analysis

To confirm the successful expression and purification of acetate kinase, SDS-PAGE and Western Blot analysis were

performed as follows. The wells were loaded with samples taken from supernatant, flowthrough, washing fractions and elution fractions.

For SDS-PAGE and Western Blot analysis, SDS loading dye containing 4x NuPage buffer (100 µL), DTT (32 µL, 1 M) and water (68 µL) was used. The protein ladder used was Precision Plus Protein™ Kaleidoscope™ Prestained Standards provided by Bio-Rad. Ladder and cell samples (containing 20 µL sample and 20 µL loading dye) were incubated at 99 °C for 10 min. Ladder (7 µL) was loaded into the left well. To each of the other wells of the two polyacrylamide gels was added cell sample (20 µL). The SDS-PAGE gel was stained with Coomassie Blue (50 mL) for 16 h (**Figure 15**).

For Western Blot analysis, the protein bands on the polyacrylamide gel was transferred to a nitrocellulose membrane using transfer buffer from a stock solution containing 20x NuPage transfer buffer (50 mL), methanol (100 mL) & water (1 L). This was run for 2 h at 30 V. Following successful blotting, the membrane was added to blocking solution containing 1x TBST buffer (50 mL), Bovine Serum Albumin (0.5 g), milk powder (2.5 g) and avidin (1 mg) and

shaken at 15 rpm overnight. TBST buffer was diluted from 10x TBST stock solution containing Tris (1.21 g), sodium chloride (0.55 g), Tween 20 (0.5 mL) and water (100 mL). The membrane was washed with 1x TBST (3 x 50 mL) for 5 min before incubation with TBST (50 mL) and Strep-Tactin conjugate (5 µL) at r.t. for 45 min. The membrane was then washed with 1x TBST (3 x 10 mL) for 10 min before colourisation with NBT/BCIP (45 µL) in 1x AP buffer (5 mL) diluted from a stock 10x AP buffer containing Tris (1.21 g), sodium chloride (5.7 g), magnesium chloride (0.47 g) and water (100 mL). The membrane is then washed with water and visualized (**Figure 15**).

Enzyme kinetic parameter determination

For determination of the kinetics of acetate kinase, cuvette and well plate assays were performed in the physiological direction, where the enzyme converts acetyl phosphate to acetate.

For cuvette assays, to a black cuvette (1 mL) was added imidazole hydrochloride (50 µL, 400 mM, pH 7.3), glucose (50 µL, 80 mM), magnesium chloride (50 µL, 120 mM), $NADP^+$ (50 µL, 12.8 mM), ADP (50 µL, 40 mM), glucose-6-phosphate dehydrogenase (20 µL, 2 mg mL^{-1}), hexokinase (20 µL, 1 mg mL^{-1}), freshly purified acetate kinase (1:10 dilution, 40 µL) and distilled water (640 - 664 µL). The cuvette was placed in the spectrophotometer and the assay was initiated with the addition of acetyl phosphate in various concentrations (1 – 25 µL, 200 mM) which allowed kinetic parameters to be determined.

For well plate assays, the same concentrations of substrates and enzymes were used, scaled down to 200 µL for use in 96-well plates. However, to reduce error in pipetting, a master mix was made and added in one portion to the wells. This master mix (8.35 mL) consisted of imidazole hydrochloride

(0.5 mL, 400 mM, pH 7.3), glucose (0.5 mL, 80 mM), magnesium chloride (0.5 mL, 120 mM), $NADP^+$ (0.5 mL, 12.8 mM), ADP (0.5 mL, 40 mM), glucose-6-phosphate dehydrogenase (0.2 mL, 2 mg mL^{-1}), hexokinase (0.2 mL, 1 mg mL^{-1}) and distilled water (5.45 mL). To each well was added master mix (167 µL), acetate kinase (1:10 dilution, 8 µL) and distilled water (0 – 24 µL) dependent on the concentration of acetyl phosphate required for the well. The assay was initiated with the addition of acetyl phosphate via injection in various concentrations (1 – 25 µL, 20 mM) which allowed kinetic parameters to be determined.

Acetate kinase inhibition

Assays for inhibition of *C. autoethanogenum* acetate kinase were performed in a 96-well plate for 200 µL assays. For inhibition studies, stock samples of inhibitors were made up to 100 mM and diluted to the required concentration. As in the well plate kinetics experiments, a master mix (8.25 mL, sufficient for 50 well assays) containing imidazole hydrochloride (0.5 mL, 400 mM, pH 7.3), glucose (0.5 mL, 80 mM), magnesium chloride (0.5 mL, 120 mM), $NADP^+$ (0.5 mL, 12.8 mM), ADP (0.5 mL, 40 mM), glucose-6-phosphate dehydrogenase (0.2 mL, 2 mg mL^{-1}), hexokinase (0.2 mL, 1 mg mL^{-1}) and distilled water (5.35 mL). Master mix (165 µL) was added to each well, with inhibitor (2 µL, 100 mM) and acetate kinase from *C. autoethanogenum* (1:10 dilution, 8 µL). The assay was initiated with the addition of acetyl phosphate via injection (25 µL, 20 mM) which allowed 1 mM inhibition versus the non-inhibited enzyme to be measured.

For IC_{50} determination, the assay contents were of the same concentration as previously, with the exception of the volume of water used and the concentration of inhibitor. Inhibitor was added to final concentrations of: 0.01 mM, 0.05 mM, 0.10 mM,

0.25 mM, 0.50 mM, 0.75 mM, 1.0 mM, 2.5 mM, 5 mM and 10 mM and the volume of water adjusted accordingly.

For determination of K_i, the assay was consistent with the method used for the well plate kinetics experiments, except inclusion of the inhibitor compound to a corresponding final concentration of 1 or 2 mM, as well as a kinetics experiment containing no inhibitor for comparison.

5.0 References

1. S. Shafiee and E. Topal, *Energ. Policy*, 2009, **37**, 181-189.
2. A. M. Henstra, J. Sipma, A. Rinzema and A. J. Stams, *Curr. Opin. Biotechnol.*, 2007, **18**, 200-206.
3. R. A. Berner, *Nature*, 2003, **426**, 323-326.
4. K. A. Kvenvolden, *Org. Geochem.*, 2006, **37**, 1-11.
5. A. T. Van Huisen and P. E. Griffin, *Geothermal Petroleum Refining*, 1981.
6. http://ths.talawanda.org/~BrambleN/classroom/Pictures/fractional distillation.JPG.
7. T. F. Yen and G. V. Chilingarian, *Oil Shale*, 1976.
8. F. Birol, *International Energy Agency*, 2010.
9. W. Youngquist, *Hubbert Center Newsletter*, 1998, **4**.
10. S. Naik, V. V. Goud, P. K. Rout and A. K. Dalai, *Renew. Sust. Energ. Rev.*, 2010, **14**, 578-597.
11. E. Ghaly, D. Dave, M. Brooks and S. Budge, *Am. J. Biochem. Biotechnol.*, 2010, **6**.
12. W. Gerhartz, *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-Vch, 1988.
13. M. E. Dry, *Catal. Today*, 2002, **71**, 227-241.
14. G. Marland, *J. Ind. Ecol.*, 2008, **12**, 136.

15. J. Fargione, J. Hill, D. Tilman, S. Polasky and P. Hawthorne, *Science*, 2008, **319**, 1235-1238.

16. J. M. Melillo, J. M. Reilly, D. W. Kicklighter, A. C. Gurgel, T. W. Cronin, S. Paltsev, B. S. Felzer, X. Wang, A. P. Sokolov and C. A. Schlosser, *Science*, 2009, **326**, 1397-1399.

17. J. Zaldivar, J. Nielsen and L. Olsson, *Appl. Microbiol. Biotechnol.*, 2001, **56**, 17-34.

18. M. Köpke, C. Held, S. Hujer, H. Liesegang, A. Wiezer, A. Wollherr, A. Ehrenreich, W. Liebl, G. Gottschalk and P. Dürre, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 13087-13092.

19. J. Daniell, M. Köpke and S. Simpson, *Energies*, 2012, **5**, 5372-5417.

20. M. R. Wilkins and H. K. Atiyeh, *Curr. Opin. Biotechnol.*, 2011, **22**, 326-330.

21. M. Kopke, C. Mihalcea, F. Liew, J. H. Tizard, M. S. Ali, J. J. Conolly, B. Al-Sinawi and S. D. Simpson, *Applied and environmental microbiology*, 2011, **77**, 5467-5475.

22. J. Abrini, H. Naveau and E.-J. Nyns, *Arch. Microbiol.*, 1994, **161**, 345-351.

23. M. Köpke, C. Mihalcea, J. C. Bromley and S. D. Simpson, *Curr. Opin. Biotechnol.*, 2011, **22**, 320-325.

24. P. C. Munasinghe and S. K. Khanal, *Bioresour. Technol.*, 2010, **101**, 5013-5022.

25. H. N. Abubackar, M. C. Veiga and C. Kennes, *Biofuels, Bioprod. Bioref.*, 2011, **5**, 93-114.

26. H. L. Drake, K. Küsel and C. Matthies, in *Prokaryotes*, Springer, 2006, pp. 354-420.

27. R. Prins and A. Lankhorst, *FEMS Microbiol. Lett.*, 1977, **1**, 255-258.

28. S. D. Brown, S. Nagaraju, S. Utturkar, S. De Tissera, S. Segovia, W. Mitchell, M. L. Land, A. Dassanayake and M. Köpke, *Biotechnol. Biofuels*, 2014, **7**, 40.

29. C. M. Humphreys, S. McLean, S. Schatschneider, T. Millat, A. M. Henstra, F. J. Annan, R. Breitkopf, B. Pander, P. Piatek, P. Rowe, A. T. Wichlacz, C. Woods, R. Norman, J. Blom, D. A. Barrett, N. R. Thomas, K. Winzer and N. P. Minton, *BMC Genomics*, 2015, **16**, 1.

30. R. S. Tanner, L. M. Miller and D. Yang, *Int. J. Syst. Bacteriol.*, 1993, **43**, 232-236.

31. E. Stackebrandt, I. Kramer, J. Swiderski and H. Hippe, *FEMS Immunol. Med. Microbiol.*, 1999, **24**, 253-258.

32. F. Liew, A. M. Henstra, M. Köpke, K. Winzer, S. D. Simpson and N. P. Minton, *Metabolic engineering*, 2017, **40**, 104-114.

33. T. Ioannides, *J. Power Sources*, 2001, **92**, 17-25.

34. D.-g. Li, H. Zhen, L. Xingcai, Z. Wu-gao and Y. Jian-guang, *Renew. Energy*, 2005, **30**, 967-976.

35. F. M. Liew, M. Köpke and S. D. Simpson, *Gas Fermentation for Commercial Biofuels Production*, INTECH Open Access Publisher, 2013.

36. L. Girbal, C. Croux, I. Vasconcelos and P. Soucaille, *FEMS Microbiol. Rev.*, 1995, **17**, 287-297.

37. J. L. Cotter, M. S. Chinn and A. M. Grunden, *Enzyme Microb. Technol.*, 2009, **44**, 281-288.

38. T. Ezeji, N. Qureshi and H. Blaschek, *World J. Microb. Biot.*, 2003, **19**, 595-603.

39. C. M. Cooksley, Y. Zhang, H. Wang, S. Redl, K. Winzer and N. P. Minton, *Metab. Eng.*, 2012, **14**, 630-641.

40. K. Winzer, K. Lorenz and P. Dürre, *Microbiology (Reading, U. K.)*, 1997, **143**, 3279-3286.

41. S. Y. Lee, J. M. Park and T. Y. Kim, in *Methods Enzymol.*, ed. V. Christopher, Academic Press, 2011, vol. Volume 498, pp. 67-93.

42. W. Wiechert, *Metab. Eng.*, 2001, **3**, 195-206.

43. S. I. Yun, S. J. Gang, H. M. Ro, M. J. Lee, W. J. Choi, S. G. Hong and K. K. Kang, *Bioprocess and biosystems engineering*, 2013, **36**, 591-595.

44. T. P. Begley, C. Kinsland and E. Strauss, *Vitamins & Hormones*, 2001, **61**, 157-171.

45. F. C. Neidhardt, *Escherichia coli and Salmonella: cellular and molecular biology*, 1996.

46. Z. Gao, J. Wang, A. K. Norquay, K. Qiao, Y. Tang and J. C. Vederas, *J. Am. Chem. Soc.*, 2013, **135**, 1735-1738.

47. I. H. Gilbert, M. Ginty, J. A. O'Neill, T. J. Simpson, J. Staunton and C. L. Willis, *Bioorg. Med. Chem. Lett.*, 1995, **15**, 1587-1590.

48. M. Jenner, S. Frank, A. Kampa, C. Kohlhaas, P. Pöplau, G. S. Briggs, J. Piel and N. J. Oldham, *Angew. Chem.*, 2013, **125**, 1181-1185.

49. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 1997, **23**, 3-25.

50. A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, *J. Comb. Chem.*, 1999, **1**, 55-68.

51. F. J. Annan, A. M. Henstra, K. Winzer and N. P. Minton, *Manuscript in preparation*, 2017.

52. A. L. Mandel, J. J. La Clair and M. D. Burkart, *Org. Lett.*, 2004, **6**, 4801-4803.

53. D. H. G Court, E. R. Lee, D. P. J. Pearson, *J. Chem. Soc., Chem. Commun.*, 1990, 331-333.

54. A. C. Cope and H. L. Jackson, *Preparation of acyl phosphoric esters*, 1956.

55. G. Zubay, *Biochemistry*, 1998, **4th Ed.**

56. A. Gorrell, S. H. Lawrence and J. G. Ferry, *J. Biol. Chem*, 2005, **280**, 10731-10742.

57. S. Zinder, K. Sowers and J. Ferry, *Int. J. Syst. Evol. Microbiol.*, 1985, **35**, 522-523.

58. Z. Gao, J. Wang, A. K. Norquay, K. Qiao, Y. Tang and J. C. Vederas, *J. Am. Chem. Soc.*, 2013, **135**, 1735-1738.

59. A.-S. Jullien, C. Gateau, I. Kieffer, D. Testemale and P. Delangle, *Inorg. Chem.*, 2013, **52**, 9954-9961.

60. J. D. Patrone, J. Yao, N. E. Scott and G. D. Dotson, *J. Am. Chem. Soc.*, 2009, **131**, 16340-16341.

61. V. Agarwal, S. Diethelm, L. Ray, N. Garg, T. Awakawa, P. C. Dorrestein and B. S. Moore, *Org. Lett.*, 2015, **17**, 4452-4455.

62. D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat and E. Dong, *Nucleic Acids Res.*, 2012.

63. D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong and S. Bouatra, *Nucleic Acids Res.*, 2009, **37**, D603-D610.

64. D. S. Wishart, D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt and S. Sawhney, *Nucleic Acids Res.*, 2007, **35**, D521-D526.

65. R. W. Adams, C. M. Holroyd, J. A. Aguilar, M. Nilsson and G. A. Morris, *Chem. Commun.*, 2013, **49**, 358-360.

66. J. A. Aguilar, M. Nilsson, G. Bodenhausen and G. A. Morris, *Chem. Commun.*, 2012, **48**, 811-813.

67. I. Kahane and A. Muhlrad, *J. Bacteriol.*, 1979, **137**, 764-772.

68. M. T. Latimer and J. G. Ferry, *J. Bacteriol.*, 1993, **175**, 6822-6829.

69. C. Ingram-Smith, R. D. Barber and J. G. Ferry, *J. Biol. Chem*, 2000, **275**, 33765-33770.

70. Y. S. Kim and P. Chan, *Biochimica et Biophysica Acta*, 1988, **956**, 103-109.

71. R. Karaman, A. Goldblum, E. Breuer and H. Leader, *J. Chem. Soc., Perkin Trans.*, 1989, 765-774.

72. M. D. Kosobokov, I. D. Titanyuk and I. P. Beletskaya, *Mendeleev Commun.*, 2011, **21**, 142-143.

73. H. Krawczyk, *Synth. Comm.*, 1997, **27**, 3151-3161.

74. B. J. Lundy, S. Jansone-Popova and J. A. May, *Org. Lett.*, 2011, **13**, 4958-4961.

75. J. W. McGrath, F. Hammerschmidt, W. Preusser, J. P. Quinn and A. Schweifer, *Org. Biomol. Chem.*, 2009, **7**, 1944-1953.

76. V. R. Sirivolu, S. K. V. Vernekar, T. Ilina, N. S. Myshakina, M. A. Parniak and Z. Wang, *J. Med. Chem.*, 2013, **56**, 8765-8780.

77. J. T. Whitteck, W. Ni, B. M. Griffin, A. C. Eliot, P. M. Thomas, N. L. Kelleher, W. W. Metcalf and W. A. van der Donk, *Angew. Chem., Int. Ed.*, 2007, **46**, 9089-9092.

78. A. Arbuzov, *J. Russ. Phys. Chem. Soc*, 1906, **38**, 687.

79. H. Finkelstein, *Ber. Deutsch. Chem. Gesell*, 1910, **43**, 1528-1532.

80. K. Radwan-Olszewska, F. Palacios and P. Kafarski, *J. Org. Chem.*, 2011, **76**, 1170-1173.

81. T. J. Curphey, *J. Org. Chem.*, 2002, **67**, 6461-6473.

82. C. E. McKenna, M. T. Higa, N. H. Cheung and M.-C. McKenna, *Tetrahedron Lett.*, 1977, **18**, 155-158.

83. L. Girbal, G. von Abendroth, M. Winkler, P. M. Benton, I. Meynial-Salles, C. Croux, J. W. Peters, T. Happe and P. Soucaille, *Applied and environmental microbiology*, 2005, **71**, 2777-2781.

84. H. Nakajima, K. Suzuki and K. Imahori, *J. Biochem.*, 1978, **84**, 193-203.

85. R. Breitkopf, R. Uhlig, T. Drenckhan and R.-J. Fischer, *Extremophiles*, 2016, 1-9.

86. H. Jeong, V. Barbe, C. H. Lee, D. Vallenet, D. S. Yu, S.-H. Choi, A. Couloux, S.-W. Lee, S. H. Yoon and L. Cattolico, *J. Mol. Biol.*, 2009, **394**, 644-652.

87. F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode and G. F. Mayhew, *Science*, 1997, **277**, 1453-1462.

88. L. Menten and M. Michaelis, *Biochem Z*, 1913, **49**, 333-369.

89. H. Lineweaver and D. Burk, *J. Am. Chem. Soc.*, 1934, **56**, 658-666.

90. G. Eadie, *J. Biol. Chem*, 1942, **146**, 85-93.

91. B. Hofstee, *Nature*, 1959, **184**, 1296-1298.

92. C. S. Hanes, *Biochem. J.*, 1932, **26**, 1406.

93. K. A. Johnson and R. S. Goody, *Biochemistry*, 2011, **50**, 8264-8269.

94. C. Yung-Chi and W. H. Prusoff, *Biochem. Pharmacol.*, 1973, **22**, 3099-3108.

95. B. T. Burlingham and T. S. Widlanski, *J. Chem. Educ*, 2003, **80**, 214.

96. I. H. Gilbert, M. Ginty, J. A. O'Neill, T. J. Simpson, J. Staunton and C. L. Willis, *Bioorganic & Medicinal Chemistry Letters*, 1995, **5**, 1587-1590.

97. S. D. Brown, S. Nagaraju, S. Utturkar, S. De Tissera, S. Segovia, W. Mitchell, M. L. Land, A. Dassanayake and M. Köpke, *Biotechnology for biofuels*, 2014, **7**, 40.

98. U. Consortium, *Nucleic Acids Res.*, 2014, gku989.

99. S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403-410.

100. M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi and M. Tanabe, *Nucleic Acids Res.*, 2014, **42**, D199-D205.

101. M. Kanehisa and S. Goto, *Nucleic Acids Res.*, 2000, **28**, 27-30.

102. A. Bateman, L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths‐Jones, A. Khanna, M. Marshall, S. Moxon and E. L. Sonnhammer, *Nucleic Acids Res.*, 2004, **32**, D138-D141.

103. F. Servant, C. Bru, S. Carrère, E. Courcelle, J. Gouzy, D. Peyruc and D. Kahn, *Briefings Bioinf.*, 2002, **3**, 246-251.

6.0 Appendix

Acetate kinase gene and protein sequences

Gene sequence for acetate kinase:

>NC_022592.1:3613208-3614404 Clostridium autoethanogenum DSM 10061, complete genome
ATGAAAATATTAGTAGTAAACTGTGGAAGTTCATCTTTAAAATATCAACTTATTGATATGCAA
GATGAAAGTGTTGTAGCAAAGGGTCTTGTAGAAAGAATAGGAATGGACGGTTCAATTTTAACA
CACAAAGTTAATGGAGAAAAGTTTGTTACAGAGCAACCAATGGAAGACCACAAAGTTGCTATA
CAATTAGTATTAAATGCTCTTGTAGATAAAAAACATGGTGTAATAAAAGACATGTCAGAAATA
TCCGCTGTAGGACATAGAGTTTTGCACGGTGGAAAGAAATATGCAGCATCCATTCTTATTGAC
GAAAATGTAATGAAAGCAATAGAAGAATGTATCCCACTAGGACCACTACATAATCCAGCTAAT
ATAATGGGAATAGATGCTTGTAAAAAATTAATGCCAAATACTCCAATGGTAGCAGTATTTGAT
ACAGCATTTCATCAGACAATGCCAGATTATGCTTATACTTATGCAATACCTTATGATATATCT
GAAAAGTATGATATCAGAAAATATGGTTTTCATGGAACTTCTCATAGATTCGTTTCAATTGAA
GCAGCTAAATTATTAAAGAAAGATCCAAAAGATCTTAAGTTAATAACTTGTCATTTAGGAAAT
GGAGCTAGCATATGTGCAGTAAACCAAGGAAAAGCAGTAGATACAACTATGGGACTTACTCCT
CTTGCAGGACTTGTAATGGGAACTAGATGCGGTGATATAGATCCAGCTATAGTACCATTTGTA
ATGAAAAGAACAGGCATGTCTGTAGATGAAGTGGATACCTTAATGAATAAAAAGTCAGGAATA
CTTGGAGTATCAGGAGTAAGCAGTGATTTTAGAGATGTAGAAGAAGCTGCAAATTCAGGAAAT
GATAGAGCAAAACTTGCATTAAATATGTATTATCACAAAGTTAAATCTTTCATAGGAGCTTAT
GTTGCAGTTTTAAATGGAGCAGATGCTATAATATTTACGGCAGGACTTGGAGAAAATTCAGCA
ACTAGCAGATCTGCTATATGTAATGGATTAAGCTATTTTGGAATTAAAATAGATGAAGAAAAG
AATAAGAAAAGGGGAGAGGCACTAGAAATAAGCACACCTGATTCAAAGATAAAAGTATTAGTA
ATTCCTACAAATGAAGAACTTATGATAGCTAGGGATACAAAAGAAATAGTTGAAAATAAATAA

Amino acid sequence for acetate kinase:
mkilvvncgs sslkyqlidm qdesvvakgl verigmdgsi lthkvngekf
vteqpmedhk vaiqlvlnal vdkkhgvikd mseisavghr vlhggkkyaa
silidenvmk aieeciplgp lhnpanimgi dackklmpnt pmvavfdtaf
hqtmpdyayt yaipydisek ydirkygfhg tshrfvsiea akllkkdpkd
lklitchlgn gasicavnqg kavdttmglt plaglvmgtr cgdidpaivp
fvmkrtgmsv devdtlmnkk sgilgvsgvs sdfrdveeaa nsgndrakla
lnmyyhkvks figayvavln gadaiiftag lgensatsrs aicnglsyfg
ikideeknkk rgealeistp dskikvlvip tneelmiard tkeivenk

GenDB computational genome analysis

The genome of *C. autoethanogenum* was previously annotated by Brown *et al.* (2014).[97] However, this annotation was for the industrial strain used by LanzaTech, which contains genetic modifications from the original strain isolated by Abrini *et al.* (1994).[22] Within this research, the strain used and worked towards has been JA1-1, which is genetically identical to the strain used by Abrini *et al*.[22]

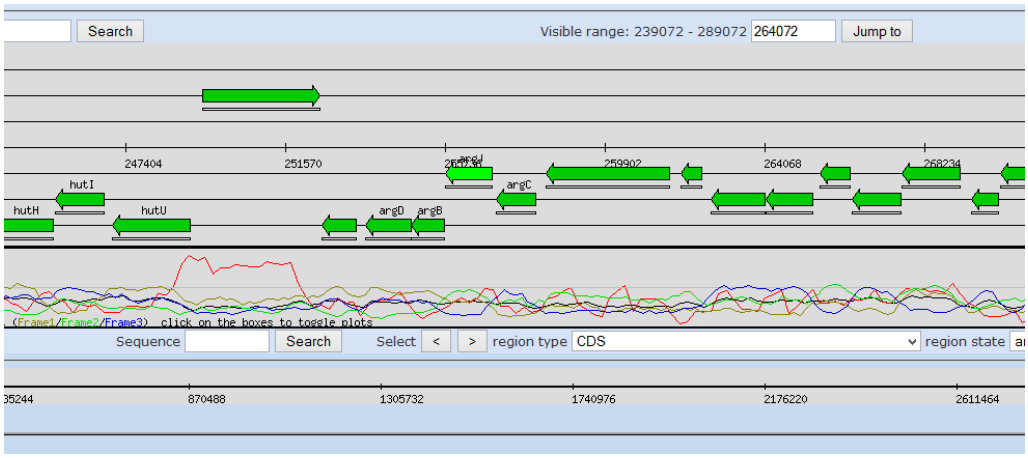


Figure 45 – The GenDB program showing successfully annotated genes within the sequence.

To annotate the genome, which was divided between eleven researchers, online databases of genes, proteins and enzymes were used. The databases used for this project were UniProt,[98] NCBI BLAST,[99] KEGG,[100, 101] PFAM[102] and PRODOM[103] using sequence searches of the amino acid sequence of the protein. This gave information on the gene function, reaction involved, enzymatic function, EC number, gene name and any homology

with other bacteria. The work was performed and published in collaboration with other members of the Synthetic Biology Research Centre (SBRC).[29]