Characterisation of Anaerococcal enzymes required for production of volatile fatty acids in human body odour

Matthew Rose

PhD Thesis

University of York

Biology

March 2023

Abstract

The human axilla hosts a diverse number of microorganisms that act upon various host-derived secretions, with microbial biotransformation of these secretions resulting in the production of body odour (BO). A major contributor to axilla BO is the generation of volatile fatty acids (VFAs), particularly two C_7 VFAs which are produced from metabolism of their respective glutamine conjugated precursors. Release of these VFAs has been attributed to species of *Corynebacterium* but has more recently been linked to bacteria in the *Anaerococcus* genus.

To understand the role of *Anaerococcus* in BO the enzymes responsible for VFA generation in two related species, *A. octavius* and *A. nagyae* were discovered using a biochemical screen and named Oct3 and Nag3, respectively. *In vitro* characterisation reveals they are zinc-dependent aminoacylases which hydrolyse their substrate to release glutamine and the VFA. These orthologous enzymes display differing activity with the two odour precursors, with Oct3 exhibiting high selectivity for a single precursor while Nag3 has more promiscuous activity with both. Site directed mutagenesis and structural determination by X-ray diffraction revealed key interactions within both binding sites that ensure enzyme specificity for glutamine conjugated substrates while also being able to accommodate a large acyl chain in the binding pocket. Structural comparison to a ligand bound corynebacterial VFA-liberating aminoacylase revealed extensive differences in protein conformation which is proposed to occur upon substrate binding. Screening of *Anaerococcus* species for VFA activity *in vivo* showed several species are capable of VFA release, suggesting a prominent role for the genus in BO production. *Escherichia coli* could be engineered to produce VFAs by expression of Nag3 and is proposed to be an experimental system for identification of the transporter responsible for VFA precursor uptake. This work provides new insight into the molecular basis of human BO formation by bacterial components of the underarm microbiome.

Contents

Abstract 2
Contents 3
List of tables 7
List of figures 8
Acknowledgments 10
Authors declaration 11
List of abbreviations 12
1 Introduction 14
1.1 Human skin and sweating 14
1.1.1 Skin structure 14
1.1.2 Thermoregulation and sweating 15
1.1.3 The eccrine gland 16
1.1.4 The apocrine gland 17
1.1.5 The apoeccrine gland: a controversial third sweat gland 18
1.1.6 The sebaceous gland 19
1.2 The axilla, microbiota and body odour 20
1.2.1 Axilla malodour is generated by bacterial action 20
1.2.2 Human perception of body odour 21
1.2.3 The axilla microbiota and links to malodour 23
1.3 Biochemical basis of axilla malodour 27
1.3.1 Odorous steroids 27
1.3.2 Thioalcohols 30
1.3.3 Volatile fatty acids 37
1.4 M20 aminoacylases and N-AGA 46
1.4.1 The M20 family 46
1.4.2 The structure of N-AGA 48
1.5 Gram-positive anaerobic cocci and *Anaerococcus* 52
1.6 Project aims 53
2 Materials and Methods 55
2.1 Chemicals and strains 55
2.1.1 Chemicals 55
2.1.2 Volatile fatty acid precursors 55
2.1.3 Antibiotics 55

2.1.4 Plasmid expression inducing reagents 55
2.1.5 Media 55
2.1.6 Strains used 57
2.2 DNA Manipulation and molecular biology techniques 59
2.2.1 Plasmids used 59
2.2.2 Synthesised genes used 60
2.2.3 Primers used 61
2.2.4 Standard PCR protocol 63
2.2.5 Inverse PCR site-directed mutagenesis protocol 63
2.2.6 Homologous recombination cloning 64
2.2.7 Agarose gel electrophoresis 64
2.2.8 DNA extraction from agarose gel 64
2.2.9 Plasmid extraction 65
2.2.10 PCR product purification 65
2.2.11 Creation of chemically competent *E. coli* cells 65
2.2.12 Plasmid transformation of competent *E. coli* cells 65
2.2.13 Sanger sequencing of DNA 66
2.3 Microbiological techniques 67
2.3.1 Culturing of *E. coli* 67
2.3.2 Culturing GPAC strains 67
2.3.3 Optical density measurements 67
2.3.4 Cell viability spot assay 67
2.3.5 Growth and harvesting of recombinant protein expression cultures 67
2.4 Recombinant protein purification 69
2.4.1 Proteins used 69
2.4.2 Sonication and isolation of cell lysate 69
2.4.3 Immobilised metal affinity chromatography (IMAC) 69
2.4.4 Size exclusion chromatography (SEC) 70
2.4.5 On-column protein unfolding and refolding 70
2.4.6 Concentrating protein 70
2.4.7 Protein quantification 70
2.4.8 Purified recombinant protein storage 71
2.5 *In vitro* protein biophysical characterisation 72
2.5.1 Sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS PAGE) 72
2.5.2 Coomassie blue staining 72
2.5.3 Size exclusion chromatography Multi angle laser light scattering (SEC-MALLS) 72

2.5.4 Electrospray ionisation mass spectrometry (ESI-MS) 73
2.5.5 Peptide fingerprint mass spectrometry 73
2.5.6 Circular dichroism (CD) 73
2.6 Structural characterisation techniques 74
2.6.1 Protein crystallisation 74
2.6.2 X-ray diffraction and structural determination 74
2.6.3 X-ray fluorescence (XRF) spectroscopy 74
2.7 *In vitro* protein biochemical characterisation 75
2.7.1 Glutamine detection kit assay 75
2.7.2 Ninhydrin reaction screening assay 75
2.7.3 Ninhydrin reaction kinetic assay 76
2.7.4 Metal reconstitution assay 76
2.8 Mass spectrometry detection of VFAs 77
2.8.1 *In vivo* whole cell VFA biotransformation 77
2.8.2 *In vitro* VFA biotransformation 77
2.8.3 Unilever LC-MS VFA detection 77
2.8.4 York GC-MS VFA detection 78
2.9 Bioinformatic techniques 79
2.9.1 Multiple sequence alignment (MSA) 79
2.9.2 Maximum likelihood Phylogenetic Tree 79
3 Identification of 3-methyl-2-hexenoic acid (3M2H) and 3-hydroxy-3-methylhexanoic acid (HMHA) generating aminoacylases within *Anaerococcus* 81
3.1 Introduction 81
3.2 Bioinformatic analysis of *Anaerococcus* aminoacylase proteins 83
3.3 Purification of putative *A. octavius* aminoacylases 87
3.4 Verification of the identity of purified protein by mass spectrometry 91
3.5 Identification of VFA generating *A. octavius* aminoacylase using a qualitative glutamine detection assay 94
3.6 Cloning of *A. nagyae* aminoacylases candidates 98
3.7 Purification and MS verification of *A. nagyae* aminoacylases 100
3.8 Screening of Nag3 activity using a qualitative glutamine detection assay 104
3.9 Ninhydrin assay development for monitoring VFA generating activity 106
3.10 Ninhydrin assay screen of purified aminoacylases for VFA activity 112
3.11 Discussion 115
4 Biochemical and biophysical characterisation of volatile fatty acid liberating aminoacylases 120
4.1 Introduction 120

4.2 Activity screen with acetylated amino acid 121
4.3 Michaelis-Menten kinetics determination of active VFA liberating aminoacylases 125
4.4 Oligomeric state of Oct3 129
4.5 Metal reconstitution assays 131
4.6 Circular dichroism of refolded VFA liberating aminoacylase proteins 137
4.7 Assessing folded state of Oct1 and Oct2 140
4.8 Analysis of metal content of N-AGA structure 142
4.9 Discussion 145
5 Mechanistic characterisation of Oct3 and Nag3 153
5.1 Introduction 153
5.2 Selection and design of mutagenesis variants 154
5.3 Nag3 and Oct3 are M20D aminoacylases 158
5.4 Characterising glutamine moiety substrate interactions in Nag3 165
5.5 Exploring Oct3 and Nag3 substrate selectivity 170
5.6 X-ray crystallisation and structural determination of Nag3 173
5.7 X-ray crystallisation and structural determination Oct3 184
5.8 The acyl binding pocket of N-AGA, Oct3 and Nag3 191
5.9 VFA precursor binding 195
5.10 Discussion 197
6 VFA production *in vivo* 205
6.1 Introduction 205
6.2 Predicting VFA liberating activity via a bioinformatic approach 206
6.3 Procurement and preparation of strains for screening VFA biotransformation activity 208
6.4 *Anaerococcus in vivo* VFA biotransformation screen 210
6.5 Peptoniphilaceae *in vivo* VFA biotransformation screen 216
6.6 MS detection from *in vitro* reactions 218
6.7 *E. coli* VFA generation gain of function experiment 220
6.8 Screen for *E. coli* VFA transporter 222
6.9 Discussion 227
7 Summary and future directions 234
8 Appendices 237
8.1 Appendix A – Plasmid maps 237
8.2 Appendix B – Acetylated amino acid structures 241
8.3 Appendix C – Mutant size exclusion chromatography traces 243
9 References 245

List of tables

Table 1.2.1: Microbial ecology profile of the axillary microbiota and its associations with malodour 25
Table 1.3.1: Major odorous steroid detected in axilla sweat 29
Table 1.3.2: Major odorous thioalcohols detected in axilla sweat 31
Table 1.3.3: Major odorous VFAs detected in axilla sweat 39
Table 2.1.1: List of strains used in this work 57
Table 2.2.1: List of plasmids used in this work 59
Table 2.2.2: List of primers used in this work 61
Table 2.4.1: Protein sequences used in this work 69
Table 3.7.1: List of purified aminoacylase proteins 103
Table 4.3.1: Michaelis-Menten kinetics with VFA precursors 128
Table 4.8.1: B-factors of modelled metal atoms in N-AGA 144
Table 5.2.1: Mutagenesis targets and primer sequences 157
Table 5.3.1: Michaelis-Menten kinetics for VFA precursors by mutants 164
Table 5.6.1: Nag3 data collections and refinement statistics 175
Table 6.3.1: Peptoniphilaceae VFA biotransformation assay strains 209
Table 6.4.1: Peptoniphilaceae cell viability spot test 214
Table 6.8.1: *E. coli* transport screen cell viability test 225

List of Figures

Figure 1.1.1: Skin composition in high and low odour environments 15
Figure 1.3.1: Biotransformation of Cys-Gly-3M3SH 33
Figure 1.3.2: Biotransformation of S-Cys-Gly-3M3SH by *S. hominis* 36
Figure 1.3.3: Chemical reactions catalysed by *Nα*-acylglutamine aminoacylase 42
Figure 1.3.4: Biotransformation of 3M2H-Gln and HMHA-Gln 45
Figure 1.4.1: Schematic of M20D peptidase active site. 47
Figure 1.4.2: Structure of N-AGA and bound inhibitor 49
Figure 1.4.3: N-AGA binding of competitive inhibitor 51
Figure 3.2.1: Phylogeny of putative aminoacylases within Anaerococcus 85
Figure 3.2.2: Multiple sequence alignment of N-AGA with *A. octavius and A. nagyae* aminoacylases 86
Figure 3.3.1: Purification of *Anaerococcus octavius* aminoacylase, Oct3 88
Figure 3.3.2: *A. octavius* and N-AGA protein purification 90
Figure 3.4.1: ESI-MS spectrum of Oct3 92
Figure 3.5.1: *In vitro* glutamine detection assay standards 95
Figure 3.5.2: Glutamine detection assay screen of candidate *A. octavius* aminoacylases 97
Figure 3.7.1: Purification of *Anaerococcus nagyae* aminoacylase, Nag3 101
Figure 3.7.2: Purified aminoacylase SDS PAGE and Nag protein gel filtration comparison 102
Figure 3.8.1: Nag3 glutamine detection assay screen 105
Figure 3.9.1: Ninhydrin assay development using a cuvette format 109
Figure 3.9.2: Nag3 time course activity assay with Ac-Met 111
Figure 3.10.1: L-Glutamine calibration curve 113
Figure 3.10.2: Aminoacylase VFA activity ninhydrin screen 114
Figure 4.2.1: Amino acid standard calibration curves 123
Figure 4.2.2: Acetylated amino acid activity screen 124
Figure 4.3.1: Michaelis-Menten kinetics with VFA precursors 127
Figure 4.4.1: SEC-MALLS of Oct3 130
Figure 4.5.1: Nag3 metal reconstitution assays 133
Figure 4.5.2: N-AGA metal reconstitution assay 135
Figure 4.5.3: Nag3 and N-AGA activity assay varying zinc concentration 136
Figure 4.6.1: Circular dichroism of refolded aminoacylases 139
Figure 4.7.1: Circular dichroism of Oct1 and Oct2 141
Figure 4.8.1: Electron density map of N-AGA with varied metal cofactors 144
Figure 5.2.1: Structural views of N-AGA binding site and opening 155
Figure 5.2.2: Multiple sequence alignment of N-AGA with Oct3 and Nag3 156
Figure 5.3.1: Close view of N-AGA monomer binding site 160
Figure 5.3.2: Endpoint activity of *Anaerococcus* metal interacting and catalytic residue mutants with VFA precursors 161
Figure 5.3.3: Substrate velocity plots for biotransformation of 3M2H-Gln by mutants 162
Figure 5.3.4: Substrate velocity plots for biotransformation of HMHA-Gln by mutants 163
Figure 5.4.1: Binding site of N-AGA 167
Figure 5.4.2: Endpoint activity of Nag3 glutamine interacting residue mutants 168
Figure 5.5.1: Endpoint activity of Oct3 and Nag3 substrate specificity mutants 172

Figure 5.6.1: AlphaFold model of Nag3 176
Figure 5.6.2: Structure of Nag3E138A homotetramer 178
Figure 5.6.3: X-ray fluorescence spectra of Nag3E138A crystals 179
Figure 5.6.4: Superimposition of Nag3E138A and N-AGA. 180
Figure 5.6.5: Nag3E138A and N-AGA substrate binding 182
Figure 5.6.6: Surface comparison of Nag3E138A and N-AGA 183
Figure 5.7.1: Structure of Oct3 monomer 186
Figure 5.7.2: Electron density of unknown ligand within Oct3 binding site 187
Figure 5.7.3: Comparison of Oct3 monomer and Nag3E138A 189
Figure 5.7.4: Oct3 binding site in monomer and tetrameric state 190
Figure 5.8.1: Electrostatic potential of VFA liberating aminoacylase binding sites 193
Figure 5.8.2: Acyl binding pocket of VFA releasing aminoacylases 194
Figure 5.9.1: Building VFA precursors into N-AGA binding site 196
Figure 5.10.1: Proposed reaction mechanism for 3M2H-gln hydrolysis 203
Figure 6.2.1: Aminoacylase phylogeny within Peptoniphilaceae 207
Figure 6.4.1: VFA HILIC ARC-QDA MS calibration curves 212
Figure 6.4.2: *Anaerococcus in vivo* VFA biotransformation screen 213
Figure 6.5.1: *Peptoniphilus* and *Finegoldia in vivo* VFA biotransformation screen 217
Figure 6.6.1: *In vitro* aminoacylase VFA MS detection 219
Figure 6.7.1: *E. coli* VFA malodour gain of function 221
Figure 6.8.1: *In vitro* Nag3 VFA GC-MS detection 226
Figure 6.8.2: *E. coli* transporter mutant screen 226

Acknowledgements

Firstly, I would like to thank my supervisors Prof. Gavin Thomas, Prof. Anthony Wilkinson, Dr Gordon James and Dr Barry Murphy for the opportunity to undertake this PhD and all their support, guidance and encouragement throughout my time on this project. Special thanks go to my TAP members, Prof. Gideon Grogan and Dr Marjan Van der Woude for all their helpful suggestions and support over the years.

I would also like to thank all the Unilever team who have helped me greatly at various stages of the project. An especially big thanks to Dr Rob Cornmell who was incredibly helpful during my placement at Unilever R&D, Port Sunlight, making the time I spent there both productive and enjoyable.

I'm grateful to past and present members of the Thomas group who have made my entire time a fun and interesting experience that I have enjoyed immensely. A special mention must be made to Reyme Herman and Dr Michelle Rudden who were fantastic in helping show me the ropes when I first joined and are still helping me learn new things now! I couldn't have managed any of this without your help or had as entertaining a time while doing it.

I'd like to thank my friends and family for their support and interest in my work over the years. Although I'm afraid I haven't quite managed to cure BO as some of you might have hoped! Finally, a massive thanks to Lauren for all the love, support and random chats about science throughout my PhD. I promise someday I'll understand a bit more about human genetics!

Authors declaration

I declare that all work presented in this thesis, except where otherwise stated, is my own original work. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as references.

List of abbreviations

3M2H	3-methyl-2-hexanoic acid
3M2H-Gln	3-methyl-2-hexanoic acid – glutamine
ABC	ATP-binding cassette
BeStSel	Beta Structure Selection
BO	Body odour
BSA	Bovine serum albumin
CD	Circular dichroism
CV	Column volumes
DLS	Diamond Light Source
ESI-MS	Electrospray ionisation mass spectrometry
GC-FID	Gas chromatography- flame ionisation detection
GC-MS	Gas chromatography- mass spectrometry
GPAC	Gram-positive anaerobic cocci
GSH	Glutathione
HMHA	3-hydroxy-3-methlyhexanoic acid
HMHA-Gln	3-hydroxy-3-methlyhexanoic acid – glutamine
IM	Inner membrane
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1 thiogalactopyranoside
LB	Luria- Bertani
LC-MS	Liquid chromatography- mass spectrometry
MALDI	Matrix assisted laser desorption/ionisation
MCS	Multiple cloning site
MHC	Major histocompatibility complex
MSA	Multiple sequence alignment
N-AGA	N^{α}-acyl-glutamine aminoacylase
OM	Outer membrane
ORF	Open reading frame
POT	Proton-dependent oligopeptide transporter
RP	Ruhemann's purple
SC	Stratum corneum
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEC-MALLS	Size exclusion chromatography – multi angle laser light scattering
VFA(s)	Volatile fatty acid(s)
XRF	X-ray fluorescence

Chapter 1 – Introduction

1 Introduction

1.1 Human skin and sweating

1.1.1 Skin structure

One of the primary functions of human skin is to act as a barrier, protecting vulnerable inner tissues and organs from external harm and disruption from physical, chemical and biological sources. Skin constitutes a large surface area with typical estimations of 2 m^2 per adult human, although more recent estimations incorporating the inner area of skin appendages such as hair follicles massively increases this number to that of at least 25 m^2 (Gallo, 2017). This creates a very large area in which skin interacts with external factors and provides an enormous and diverse environment for microbes to colonise.

Skin consists of stratified groupings of cells and can be separated into three main layers: the epidermis, the dermis and the subcutaneous layer (also known as the hypodermis) (Fig 1.1.1). The subcutaneous layer is the innermost strata of skin and is comprised principally of adipose tissue which acts as insulation and an energy store as well as providing structural stability and support to blood vessels, nerves and lymphatics which supply the other major skin layers. The dermis is situated above the hypodermis and forms a relatively thick layer (0.5 - 5 mm, depending on body site) of connective tissue, principally consisting of collagen and elastin, within which can be found several different cell types including fibroblasts and various immune cells (Lai-Cheong and McGrath, 2013). Blood vessels, lymphatic channels and nerve endings also reach into this layer to supply the surrounding tissue and outermost skin layer, the epidermis. Several other appendages also originate from within the dermis and continue into and effect the epidermis such as hair follicles, sebaceous glands and sweat glands. The epidermis refers to the outer layer of skin which interacts with the environment and is responsible for the primary barrier function of skin. This layer can be further stratified several times from the dermal-epidermal membrane to the outermost layer of the stratum corneum (SC). Each of these strata is distinguished by its level of keratinocyte differentiation, the major cell type of the epidermis, with proliferating keratinocytes occurring close to the dermis at the basement membrane which migrate towards the SC to eventually become flattened dead keratinised cells known as corneocytes. Corneocyte development in part involves production of the intracellular structural proteins loricrin, involucrin and filaggrin which form a dense crosslinking with keratin fibres (Candi et al., 2005; Sandilands et al., 2009). This is coupled with the production of an extracellular lipid matrix, principally consisting of ceramides, cholesterol and free fatty acids, which fills the gaps between the corneocyte to create a structure referred to as “bricks in mortar” (Elias, 1983). It is this structure which creates

the permeability barrier of skin and facilitates the primary function of skin. Only hair follicles and eccrine sweat gland ducts traverse into the epidermis and have direct contact with the skin surface.

The basic structure of skin is found across the body but displays variation across body sites with a major part of this being the localisation of hair follicles and sweat glands to specific sites. Such differences in skin environments supports a diverse array of microorganisms to survive upon the skin surface.

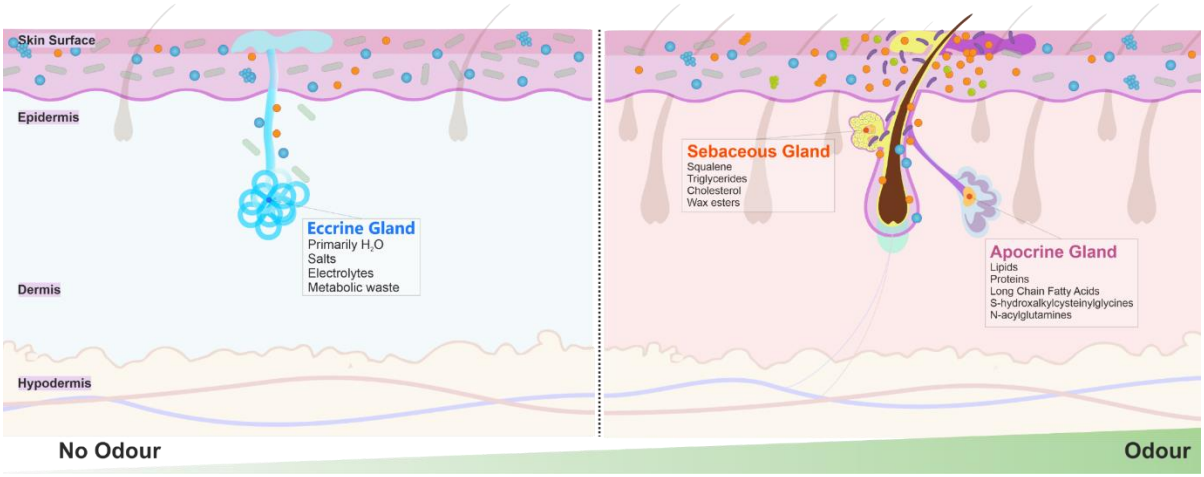


Figure 1.1.1: Skin composition in high and low odour environments

Skin environments such as the forearms are characterised as dry with the presence of just eccrine sweat glands and little odour generation by the microbiota of this environment. High odour environments such as the underarm (axilla) have a concentration of eccrine sweat glands as well as sebaceous and apocrine glands which secrete onto hair follicles. The concentration of glands and hair follicles creates a moist rich environment which supports a dense and rich microbiota, with certain odourless secretions of the apocrine gland being utilised by specific bacteria to generate odorous molecules that contribute to axilla body odour.

1.1.2 Thermoregulation and sweating

Thermal homeostasis is paramount for cellular function and human survival (Hochachka, 1986; Lepock, 2003). The large surface area of skin means skin plays a vital role in thermoregulation alongside the cardiovascular system and metabolic activity (Kurz, 2008). In humans thermal homeostasis is tightly controlled, predominantly by the hypothalamus (Satinoff, 1978) which responds to numerous factors such as core temperature fluctuation, stress and hormonal changes. The low density of hair follicles across most human skin prevents any appreciable effect during cold responses, unlike haired primates and mammals (Chaplin et al., 2014), with vasoconstriction and shivering instead being the primary effectors. During warm responses, skin contributes a major role to body cooling with perspiration. Sweating leads to the release of thermal energy upon evaporation of the aqueous component of sweat from the skin surface in a process known as evaporative cooling.

This mechanism lead to the lowering of skin temperature and consequently internal body temperature and enables cooling even if external temperatures exceeds those internal (Cramer and Jay, 2016). Sweat is released onto the skin surface by specialised exocrine sweat glands that reside within the dermis of skin. There are three types of glands present in humans that differ in their openings onto the skin, their secretions and consequently their perceived physiological roles. These glands are known as eccrine sweat glands, apocrine sweat glands and sebaceous glands which are described in greater detail in later sections (Fig 1.1.1).

Alongside its thermoregulation role, sweating has also been observed in response to other stimuli such as stress, fear, anxiety and pain in a process known as emotional sweating (Chalmers and Keele, 1952). This form of sweating has shown to be independent of temperature regulation and is restricted to the forehead, palms, soles of feet and the underarm (axilla) (Asahina et al., 2015). Sweating of the palms and soles is considered a preliminary response to danger and is commonly observed in primates to aid running and climbing with improvements in fine hand movement (Adelman et al., 1975). Forehead sweating is likewise assumed to aid responses to danger by cooling the head (Asahina et al., 2015) whereas axillary emotional sweat has no obvious function in fight/flight responses, being instead linked to human body odour (Shelley and Hurley, 1953; Shelley et al., 1953).

1.1.3 The eccrine gland

Eccrine sweat glands reside within the dermis of skin and release sweat directly onto the surface of skin through tubular appendages (Fig 1.1.1). This gland is the most common exocrine organ and distributed throughout skin although is particularly concentrated within the palm and soles (200 – 700/ cm^2). The sweat released from these glands is mainly involved in thermoregulatory sweating (alongside the response to emotional stimuli in the palms and soles), with its aqueous secretions principally consisting of water that cause evaporative cooling (Bovell, 2015). Water accounts for 99 % eccrine sweat, with the remainder including several ionic salts of which sodium, chloride and potassium represent the main solutes (Shirreffs and Maughan, 1997). Several other organic and inorganic compounds reside within eccrine sweat (Berglund and McNall, 1973; Yokoyama et al., 1991) and show significant variation in concentration between individuals although not based upon gender (Harker and Harding, 2013; Harker et al., 2006). One significant organic compound found within sweat is lactate (Patterson et al., 2000) which contributes to the slightly acidic pH of eccrine sweat (pH 4 – 6) (Shelley and Hurley, 1953). Another notable constituent is amino acids, with eccrine sweat showing a striking similarity in amino acid composition to that of the epidermal protein profilaggrin (Harker and Harding, 2013). Profilaggrin is the precursor protein to the functional filaggrin monomer present in keratin cross linking within corneocytes (Sandilands et al., 2009) and when hydrolysed forms a major

part of the natural moisturising factor within the SC that maintains the barrier integrity. This finding led to the proposal that during the passage of eccrine sweat along the gland duct to the skin surface its equilibrates its amino acid compositions with the surrounding filaggrin containing tissues although the exact significance of this contribution to natural moisturising factor is unknown (Harker and Harding, 2013). Small amounts glucose, urea and ammonia (independent of urea breakdown) can also be detected within eccrine sweat (Ono et al., 2018; Sato et al., 1989). This combined with the trace presence of several transition metals (Baker and Wolfe, 2020) creates an environment able to support microbial growth. Interestingly eccrine sweat has also been noted to contain several antimicrobial peptides such as dermicidin (Schittek et al., 2001) which is constitutively expressed and secreted from eccrine glands irrespective of immune response, forming a part of the innate defence of skin (Schittek, 2012). As a result, microorganisms present on skin must contend with both nutrient competition with neighbouring species and the antimicrobial products of the skin, sweat and other microorganisms.

1.1.4 The apocrine gland

The apocrine gland also resides within the dermis but differs morphologically to eccrine glands and in its localisation across the body. Secretions travel via the apocrine duct directly onto hair follicles rather than the skin surface like eccrine glands (Fig 1.1.1). Apocrine glands are located primarily within the axilla but are also found within the ear canal, areolae of breasts and the perineum (Collins, 1989). This specific localisation is uncommon in most primates and mammals with only humans, chimpanzee and gorillas displaying such spatial distribution (Folk and Semken, 1991). Interestingly, human fetuses initially develop apocrine glands across the entire skin before undergoing atrophy after 5 months gestation unless located in the aforementioned areas (Hashimoto, 1965; Hurley and Shelley, 1960). These remaining glands are dormant until development at the onset of puberty where apocrine secretions can then be detected in axilla sweat (Shelley and Hurley, 1953; Shelley et al., 1953). While apocrine gland sweat does provide a thermoregulatory role within haired mammals such as donkeys and camels (Bullard et al., 1970; Folk and Semken, 1991; Montagna and Parakkal, 1974), this is not the norm for mammals with most utilising eccrine gland secretion. Instead, apocrine glands have long been linked to human body odour (BO) (Shelley and Hurley, 1953; Shelley et al., 1953). Apocrine gland secretion in humans has also been shown to only respond to emotional stimuli, unlike eccrine glands, with fear and pain having the greatest effect upon secretion (Shelley and Hurley, 1953). The association of secretion with odour and apocrine development occurring during puberty, combined with the observation that apocrine glands in specialised sites such as the perineum are associated with scent glands in several mammals (Collins, 1989) has led to speculation that they may be part of a, now vestigial, pheromone system. The fact that axilla odorants which contribute to BO are often highly

perceptible to the human nose would lend credence to apocrine glands having played an important role in human evolutionary history.

Apocrine gland secretions are notably different to that of the eccrine gland; producing a viscous milky fluid which has not been completely chemically defined (Folk and Semken, 1991; Leyden et al., 1981). Notable levels of lipids have been reported with the majority being cholesterol (76 %) alongside triglyceride and fatty acids (19.2 %) (Leyden et al., 1981). Significant amounts of protein have also been detected with a multitude of different protein being observed, many at extremely low levels that together contribute a sizable amount of total apocrine secretion (Farkaš, 2015). One such protein linked to apocrine secretion in lysozyme which forms part of the skins innate immune system (Campbell et al., 1982). Apocrine sweat pH is slightly more basic than that of eccrine sweat with a range of 5 to 6.5 and is attributed to the absence of lactate (Shelley and Hurley, 1953). Apocrine sweat also contains numerous small molecule and metabolites, with some of these being of particular significance in the generation of human body odour in the axilla. Those specific molecules linked to BO will be discussed in detail in later sections.

1.1.5 The apoeccrine gland: a controversial third sweat gland

Eccrine and apocrine glands were initially assumed to be the only two types of sweat glands present on the human body although evidence in 1987 suggested the presence of a third type, located primarily within the axilla (Sato et al., 1987). This gland was noted to share morphological features with both eccrine and apocrine glands, leading to the term apoeccrine glands. These glands were purported to secrete directly onto the skin surface and produce a watery like secretion similar to that of eccrine glands but developed during puberty like apocrine glands where they would eventually represent 45 % of all axillary glands present (Sato and Sato, 1987; Sato et al., 1987). The average sweat rate of apoeccrine glands was estimated at 7-fold higher than that of eccrine glands (Sato and Sato, 1987) and was later linked as a possible cause of axilla hyperhidrosis (Lonsdale-Eccles et al., 2003). However more recent histology and immunofluorescence studies upon axilla sweat glands failed to find any apoeccrine glands present in human volunteers (Beer et al., 2006; Bovell et al., 2007) making the existence of this gland controversial. The lack of any more recent supporting evidence for apoeccrine gland presence make it highly doubtful there is a third type of sweat gland.

1.1.6 The sebaceous gland

Sebaceous glands form the third major type of gland found on the skin and are found across most of the body, with notable absence on the palms and soles of the feet. The glands are particularly concentrated on the face, forehead and scalp where between 400-900 glands/cm^2 have been reported (Folk and Semken, 1991; Thody and Shuster, 1989). Sebaceous glands secrete directly onto hair follicles (Fig 1.1.1), coating each hair with an oily substance termed sebum and classifying them as oil glands (Folk and Semken, 1991). Sebum consists of a mixture of lipids, mainly consisting of triglyceride, squalene, wax esters, fatty acids with a small level of cholesterol (Downing et al., 1989; Greene et al., 1970). This lipid mixture has been shown to be distinct in composition to that present in apocrine sweat , most notably based on differing cholesterol content as well as the characteristic presence of squalene and wax esters in sebum (Leyden et al., 1981). Sebum is not involved in thermoregulation with its exact function in humans not being fully understood, although roles in skin photoprotection, maintaining skin integrity and antimicrobial activity have been linked (Picardo et al., 1992; Zouboulis et al., 2008). As sebaceous glands are associated with hair follicles this commonly locates them close to apocrine glands in body sites that harbour them. The secretion of both apocrine and sebaceous glands onto the hair follicles in the dermis leads to a mixing of the two secretions by the time they appear on the skin surface and dilutes apocrine secretion viscosity from a glue like consistency to a more liquid-like, lather (Folk and Semken, 1991; Thody and Shuster, 1989).

1.2 The axilla, microbiota and body odour

1.2.1 Axilla malodour is generated by bacterial action

The axilla contains a dense concentration of hair follicles and glands whose secretions of sebum, eccrine and apocrine sweat create a hairy, moist and nutrient rich environment which has been likened to a tropical rainforest on the skin (Marples, 1969). This environment is in stark contrast to much of the rest of human skin which is typically drier and much more nutrient scarce, due to the absence of apocrine glands and an open environment promoting liquid evaporation (Fredrich et al., 2013; Noël et al., 2012). The axilla has long been known as a source of body odour, with references made throughout history including as far back as Aristotle in his Problemata asking:

"Why has the armpit a more unpleasant odour than any other part of the body?" (translated from the Greek) (Classen et al., 1994).

However it was not until the seminal work of Shelley, Hurley and Nichols that axilla malodour was shown to be caused by bacterial action upon apocrine secretions, with sterile apocrine sweat, and in fact all sweat, being generally odourless at the moment of secretion (Shelley et al., 1953). This work definitively showed that apocrine sweat remained odourless after two weeks incubation if sterile while non-sterile apocrine sweat developed the characteristic armpit odour within hours of incubation. Samples of eccrine sweat were also tested and noted to be odourless regardless of sterility after incubation. Development of apocrine odour was shown to be largely inhibited by application of an antibacterial wash to the axilla prior to apocrine sweat extraction and led to the conclusion that bacterial biotransformation of compounds within axilla sweat must be responsible for axilla malodour. Work three years later by Strauss and Kligman suggested that both Gram-positive and Gram-negative bacteria could produce malodour from incubation of a range of bacterial cultures with sterile apocrine sweat (Strauss and Kligman, 1956). This work also showed that sebum did not contribute to axilla malodour after incubation of sebum (taken from the scalp) with the identified malodour producing strains failed to generate any odour. Notably this work performed all incubations under aerobic conditions with aerobic bacteria, preventing any investigation of the contribution that anaerobic and microaerophilic bacteria may have. Later work by Shehadeh and Kligman amended this position to demonstrate only Gram-positive microorganisms were capable of malodour biotransformation (Shehadeh and Kligman, 1963). This study of topical antibacterial effects on the axilla microbiota noted the axilla is predominated by Gram-positive bacteria which could be almost entirely eliminated with topical treatment of neomycin and aluminium chlorhydroxide, allowing resistant Gram-negative bacteria to now dominate the microbial population. The drastic change in microbial ecology coincided with an effective odour suppression, even when Gram-negative numbers exceeded pre-treatment

levels of Gram-positives. Based on this it was concluded that axilla Gram-negatives are not responsible for axilla malodour, and an explanation of the previous erroneous assertions in Strauss and Kligman, (1956) was given to be caused by the putrid odour which the tested Gram-negatives displayed in their own right without addition of apocrine sweat. Shehadeh and Kligman observed a high prevalence of both *Staphylococcus* and *Corynebacterium* species although no statements on which group, or both, were responsible for malodour generation.

1.2.2 Human perception of body odour

Earlier use of the term axilla malodour to describe the effect of axilla odorants was chosen due to the widespread consensus in western society that underarm odour is a negative and unpleasant smell (Craig Roberts et al., 2020). This is not unique to the bacterial action in the axilla since bacterial metabolism of organic matter is responsible for several other offensive odours, such as faecal matter and rotting plant and animal tissue. The particular sensitivity of the human nose to sulphur and nitrogen containing molecules (Laska et al., 2007; Nagata, 2003) is a clear example of this given their emanation from rotting tissue by bacterial degradation of amino acids (Janzen, 1977). Thiols in particular display the highest sensitivity in the human nose with detection levels as low as 0.3 parts per billion in air (Devos et al., 1990). Human olfactory detection of several sulphur odorants originating from fire has also been noted, having clear potential benefit both in the past and present to allow fleeing from danger (Nagata, 2003; Williams and Ringsdorf, 2020). Changes in human odour due to sickness have also been shown to be noticeable to other individuals (Olsson et al., 2014). Human sensitivity and often disgust to these odours is not surprising given the implications in human evolutionary history to allow a warning of decay, disease and danger.

Human malodour is not unique to the axilla with both breath and foot odours being well known sources of unpleasant smells. Oral malodour is also associated with the bacterial breakdown of proteinaceous and other organic matter by anaerobic bacteria (De Boever et al., 1994; Delanghe et al., 1999) leading to the production of several odorant such as putrescine and cadaverine with the most significant being volatile sulphur compounds and particularly hydrogen sulphide (Tonzetich and Richter, 1964). Foot malodour has been shown to be primarily caused by generation of the volatile fatty acid (VFA) isovaleric acid, mainly due to action of *Staphylococcus* species on L-Leucine, as well as other branched aliphatic amino acids (James et al., 2013a). The high concentration of eccrine sweat glands on the sole of the foot and the presence of amino acids in eccrine secretion does allow speculation that isovaleric acid generation is dependent on eccrine sweat on the foot, with leucine being relatively significant part of the amino acid composition within eccrine sweat. However, amino acids represent only a small amount of overall eccrine sweat composition (Harker and Harding, 2013).

While high levels of eccrine sweat in the foot would certainly contribute to isovaleric acid production, the leucine utilised in foot odour was proposed to also come from the breakdown of foot calluses by other bacteria (James et al., 2013a), although more recent reviews on this topic emphasise eccrine sweat as the key source (James, 2020). Both mouth and foot odour are relatively common occurrences within the human population, but neither are considered as deliberate functions of those body sites, instead being viewed as opportunistic bacterial action on present nutrients.

In contrast, axilla odour is viewed as a deliberate function of the underarm with specialised production of odour precursors specific to that body site which generate unique odorous molecules (James et al., 2013b; Natsch and Emter, 2020). While generally perceived as a negative effect, apocrine gland derived scent has been shown to convey several signals in animals, such as scent marking, danger indications and fertility (Collins, 1989; Levy and Yizhar, 2018; Stowers et al., 2013; Vaglio et al., 2021). The sensing of stress and fear is of note given that apocrine secretion rates can be induced by those exact emotions in humans (Chalmers and Keele, 1952) and can chemosensory transmission of both has been shown in humans, interestingly with a heightened sensitivity present in women (de Groot et al., 2014; Pause et al., 2020). Chemoreception by olfaction has been noted in humans in all stages of human development, with nasal detection of odorous compounds within amniotic fluid being observed by fetuses and neonates (Schaal et al., 1995, 1998). Infants show clear olfactory attraction to their mother's mammary odour, including formula fed neonates (Macfarlane, 1975; Porter et al., 1991), and initial offense to axilla odour (Makin and Porter, 1989), suggesting an evolutionary encoded preference for mammary odour and breasts to encourage feeding. At two weeks post birth, infants were noted to lose their offense to axilla odour and instead show signs of maternal recognition based on this odour (Cernoch and Porter, 1985). The presence of apocrine glands in both the breast areola and axilla would suggest a role in odour generated from their secretions, although the exact compounds eliciting these responses in infants has yet to be identified and the likely role of milk odours should not be discounted (Schaal et al., 2020).

Human body odour has also been linked in mate selection. Partner activeness has been shown to be effected by body odour, particularly in women's choice of partners (Herz and Inzlicht, 2002), although this effect has recently been ascribed mostly to offense from axilla odours rather than attraction (Ferdenzi et al., 2020). An interesting area of study emerged in the 1970s suggesting a link between mating preference and odour in mice, which was influenced by the major histocompatibility complex (MHC) with MHC dissimilar mice showing increased attraction to one another (Yamazaki et al., 1976). The mating of MHC dissimilar individuals leads to an increased likelihood of MHC heterozygous offspring which have been shown to provide a better immune response upon pathogenic challenge (Penn et al., 2002). This preference for MHC dissimilar mates via body odour was later reported in

humans (Wedekind et al., 1995), although subsequent studies on this effect have suggested no significant link between MHC similarity and mate selection occurs (Havlíček et al., 2020). Furthermore, no key odorant in mice or humans has ever been identified to be the potential cause of this odour difference (Kwak et al., 2010; Natsch et al., 2010).

Clearly, body odour plays an important role in olfactory sensing within humans and has implications on our evolutionary past to aid development, alert to danger in a social setting and warn of disease. In fact, the role of olfactory communication in humans has been recently gaining increasing focus suggesting the discovery of further roles for body odour in the future (Craig Roberts et al., 2020). The dominant role the axilla plays in the generation of odorants and its links already in several areas of chemosensory communication would suggest future discoveries will likely originate from this scent producing organ.

1.2.3 The axilla microbiota and links to malodour

Since Shelley and Hurley (1953) showed that it was bacterial action upon odourless apocrine sweat that generates axilla malodour and Shehadeh and Kligman (1963) observed this was due to Gram-positive bacteria, numerous studies have attempted to characterise the axilla microbiota and the species responsible for malodour generation. The moist, nutrient rich environment of the axilla creates a particularly bountiful ecological niche for microorganism which allows the existence of a dense and diverse microbiome. Such is the difference of this environment to most others found on the body led to this apt analogy by Mary J Marples (1969):

“The skin is an ecosystem, with a microscopic flora and fauna and diverse ecological niches: the desert of the forearm, the cool woods of the scalp and the tropical forest of the armpit” (Marples, 1969).

Notable work on the axillary microbiota and odorant production was undertaken by Leyden *et al* in 1981, who observed the high bacterial density present in the underarm reaching up to 10^6 colony forming units (CFU) per cm^2 of skin (Leyden et al., 1981). In contrast, studies of total microbiota abundance upon dry skin sites such as the forearms estimated around 10^4 CFU per cm^2 of skin (Bibel and Lovell, 1976; Leyden et al., 1987). Aerobic and anaerobic culture-based methods were used to quantify and classify bacteria present in the axilla and correlate this with subjective determination of odour type and intensity. Data was collected from 128 males and 77 females and showed a stable population that showed little population density variation day-to-day, between left and right axilla or between sexes. Leyden *et al* showed that the microbiota contains a significant density of Micrococcaceae, a family which used to have *Staphylococcus* species misclassified within, which consisted of 10 % *Staphylococcus aureus*, 51 % *S. epidermidis,* 29 % *S. saprophyticu*s and 10 %

Micrococcus species. Alongside Micrococcaceae, high densities of lipophilic and large colony aerobic diptheroids were noted, a now defunct term which predominantly refers to *Corynebacterium* species, although these were not classified beyond identification as *Corynebacterium* species. *Propionibacterium*, now known as *Cutibacterium* (Scholz and Kilian, 2016), represented the third largest grouping found, although at two orders of magnitude lower than Micrococcaceae and diphtheroids, and contained *Propionibacterium acnes* (now *Cutibacterium acnes*), *P. avidum* (now *C. avidum*) and *P. granulosum* (now *C. granulosum*) highlighting a major presence of anaerobic bacteria within the axilla. Upon relating this to axilla odour descriptors, a "pungent apocrine" odour was shown to be present in individuals with higher bacterial densities and correlated with *Corynebacterium* numbers. A "faint acid, non-apocrine" odour description was instead associated with high *Staphylococcus* densities. Incubation of apocrine sweat with cultured axilla bacteria was then shown to only generate "pungent apocrine" odour with *Corynebacterium,* while *Staphylococcus* incubations generated "acidy" and "sweaty" odours which were noted to be similar to descriptors given to isovaleric acid. The dominance and correlation of *Corynebacteria* numbers to odour intensity was also shown in another study two years later by Jackman and Noble (1983) who characterised the axilla microbiota by a similar culture dependent method of 163 male and female volunteers (Jackman and Noble, 1983).

A later study by Taylor *et al* (2003) also used culture-based methods to investigate the axilla microbiota and its relationship with odour using a wide range of selective and nonselective medias to try overcome any culture biases inherent in other studies (Taylor et al., 2003). The approach was largely similar to that undertaken by Leyden *et al*, with axilla samples taken from a total of 36 individuals (5 females and 31 males) for microbiological culturing before investigation of relationships between the microbiota and odour descriptors. The results of this study broadly matched those detailed earlier (Leyden et al., 1981) with a high prevalence and density of *Staphylococcus* and aerobic coryneforms (a term historically used almost synonymously with *Corynebacterium*)(Table 1.2.1). A high variability was observed in genera densities between individuals, with some individuals having *Cutibacterium* as the dominant genus cultured rather than the more common *Staphylococcus* or aerobic coryneform dominance. *Micrococcus* was again identified in this study, at a lower prevalence and density than the dominant three. The yeast genus Malassezia was also identified to be present is some individuals although at a much lower average density. Malassezia is a known skin commensal but is particularly prevalent on the human scalp and linked with dandruff (Gemmer et al., 2002). Assessment of genus density to subjective axilla malodour intensity measurements a highly significant association was observed for aerobic coryneforms, again implicating *Corynebacterium* as key malodour producers. A significant result was also observed for *Micrococcus*, although later

investigations suggested the low prevalence and density of this genus prevents it from having a major role in malodour generation (James et al., 2004a, 2013b). With mounting evidence implicating *Corynebacterium* in axilla BO, Taylor *et al* used 16S rRNA gene sequencing to identify the isolated aerobic coryneforms. A total of 61 isolates were analysed, with 41 isolates having the highest similarity score to *Corynebacterium* sp. G-2 which was later shown to be *C. tuberculostearicum* (James et al., 2013b). The next most dominant species identified was *C. mucifaciens*, followed by *C. afermentans*, *C. amycolatum, C. genitalum, C. riegelii* and *C. striatum*. This study, combined with the findings of Leyden *et al* (1981), strongly suggested that *Corynebacterium* were the likely producers of apocrine axilla malodour and that *C. tuberculostearicum* was the most prevalent species present on axilla skin.

Genus/type	Prevalence (%)	Density range (log_{10} CFU cm^{-2})	Mean density (log_{10} CFU cm^{-2})	Association with malodour
Staphylococcus	98.4	1.72–7.00	5.76	No
Aerobic coryneforms	93.4	0.72–7.51	6.07	Yes ($P < 0.0001$)
Micrococcus	45.0	0.72–4.72	3.19	Yes ($P < 0.05$)
Propionibacterium	87.5	0.72–7.07	5.36	No
Malassezia	21.7	0.72–2.67	1.20	No

Table 1.2.1: Microbial ecology profile of the axillary microbiota and its associations with malodour

Taken from (James et al., 2013b), which was an adapted version of that seen in (Taylor et al., 2003).

The advent and adoption of high throughput DNA sequencing technologies by the late 2000s have allowed much more in-depth studies of skin microbiome populations. Early studies exploring the skin microbiome of various body site by 16S rRNA gene phylotyping, noted the same broad genera representation and dominance in the axilla of *Staphylococcus*, *Corynebacterium* and *Cutibacterium* as that noted in earlier, culture dependent methods (Costello et al., 2009; Grice et al., 2009). The dominance of these three genera was also observed to occur across most other sites of the body despite differing skin characteristics. Alongside the dominant genera, a much greater diversity of other bacteria was also found that had been missed from previous culture based methods (Costello et al., 2009; Grice et al., 2009; Leyden et al., 1981; Taylor et al., 2003). Of particular note was the discovery of new members of Firmicutes, which mainly consisted of a heterogeneous group of anaerobic bacteria termed Gram-positive anaerobic cocci (GPAC) (Grice et al., 2009). The GPAC found were primarily composed of the genera *Anaerococcus* and *Peptoniphilus* and were noted as present across most human skin (Grice et al., 2009), and particularly prevalent within the axilla (Egert et al., 2011).

Later work by Troccaz *et al* (2015), again using a cultured-independent approach and sequencing of 16S rRNA genes, supported these findings with *Staphylococcus*, *Corynebacterium* and *Cutibacterium* being the most prevalent (accounting for 96 % total 16S sequences) while *Anaerococcus* and *Peptoniphilus* were found to be the fourth and fifth most abundant genera in the axilla (Troccaz et al., 2015). This study also showed differences in microbiomes between the human sexes with males shown to generally have higher bacterial loads in the axilla and a higher abundance of *Anaerococcus* and *Corynebacterium.* The association between genus and species level abundance and odour descriptors was also explored in this work and showed, like that observed by both Taylor and Leyden, a correlation between *Corynebacterium* and body odour. Further analysis at the species level indicated this positive association was for *C. tuberculostearicum* which has been previously observed as the most abundant *Corynebacterium* species present in axilla (James et al., 2013b; Taylor et al., 2003). Analysis of *Staphylococcus* as a genus showed no correlation to odour however a separation into species showed that *Staphylococcus hominis* positively associated with odour, while the most abundant species *S. epidermidis* showed a weak negative correlation. A correlation was also observed between *Anaerococcus* abundance and a "fatty" axilla odour descriptor. This work therefore also implicated certain species of *Staphylococcus* and *Anaerococcus* as able to effect axilla malodour generations, widening the field of bacteria to study from the previous historical focus on *Corynebacterium.*

1.3 Biochemical basis of axilla malodour

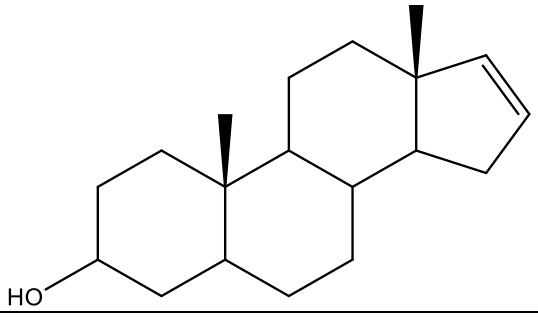
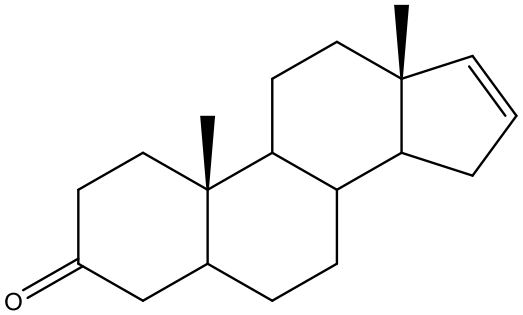
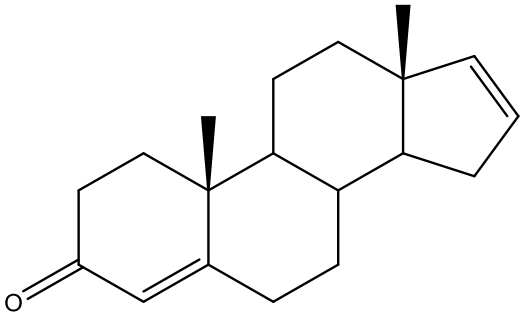
Alongside work to characterise the axilla microbiota and study associations with bacterial counts and malodour, work was also performed to identify the odorous molecules responsible for malodour. This work also involved investigation into the species capable of the biochemical mechanism to create these odorants. The identification of odorous molecules proposed to be major contributors to axilla malodour began in the 1970s and still continues to date, although focus now is more upon characterisation of the biotransformation processes in axilla bacteria rather than the discovery of novel odorants. Three classes of molecules have been proposed to be involved in axilla BO, with a general consensus now that only two of these play major roles in BO. These classes are odorous steroids (now thought to play little role), thioalcohols and volatile fatty acids (VFAs). Each of these are detailed in the sections below.

1.3.1 Odorous steroids

Odorous steroids were the first class of molecule to be proposed to contribute to malodour with this link being mentioned by Gower in 1972 with the identification of 5α-androst-16-en-3-one (androstenone) in the axilla sweat of a single male volunteer (Gower, 1972) (Table 1.3.1). The related alcohol 5α-androst-16-en-3-ol (androstenol) was later detected in the pooled axilla sweat of 12 males (Brooksbank et al., 1974). Further evidence of androstenone presence in the human axilla was published in 1976 and postulated to be a key component of axilla BO due to its known “urine” and “musky” odour (Claus and Alsing, 1976). The structurally related molecule 4, 16-androstadien-3-one (androstadienone) was later detected in axilla sweat, with the three steroids also being shown to be present in male axilla sweat and a single female (Gower et al., 1994). The potential role of these steroids in axilla BO was of particular interest due to androstenone being a prominent boar pheromone, leading to proposals that they could have formed part of a pheromone signalling system in humans that may have been of importance in human evolution (Gower et al., 1994). These steroids were also shown to have low olfactory thresholds for human detection (Amoore et al., 1977), which were below that recorded in axilla sweat by Gower *et* al (1994).

Investigation of the biochemical generation of these steroids did show that axillary *Corynebacterium* were capable but only with supplementation of precursor steroids that were structurally very similar (Austin and Ellis, 2003). Mixed populations of corynebacteria could produce low levels of both androstenone and androstenol from one precursor tested, although no single strain could perform all the biochemical reactions needed for biotransformation. Furthermore, of the mixed populations and

individual strains of *Corynebacterium* tested, only a small subset of each were capable of any biochemical action against this precursor.

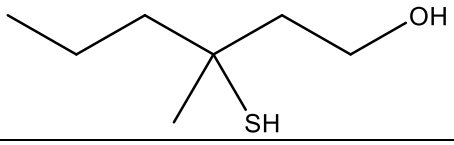
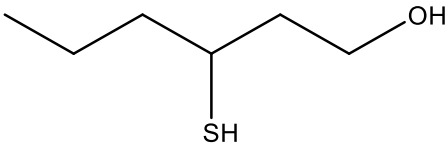
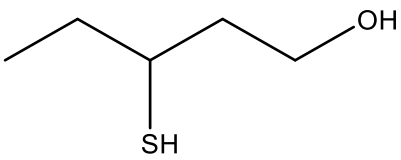
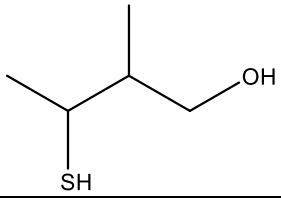
Table 1.3.1: Major odorous steroid detected in axilla sweat		
Steroid	Structure	Reference
5α-androst-16-en-3-ol (androstenol)		(Brooksbank et al., 1974)
5α-androst-16-en-3-one (androstenone)		(Gower, 1972)
4, 16-androstadien-3-one (androstadienone)		(Gower et al., 1994)

While 16-androstenes had been suggested to form part of axilla malodour, no study to date has directly linked detection of these odorous steroids to olfactory assessment of axilla odour. Study of axilla odorous molecules from axilla sweat using gas chromatography mass spectrometry (GC-MS) combined with olfactory assessment noted that while 16-androstenes could be identified by chromatography, this odour was not associated with axilla BO by any of the odour assessors. Additionally, a significant proportion of the human population displays anosmia (partial or full loss of smell) to androstenone (up to 50 %) (Amoore et al., 1977). This anosmia combined with the failure to link any odorous steroid to axilla malodour and the absence of an identified *Corynebacterium* strain capable of significant steroid production has led to the general consensus that these molecules have little significant effect on axilla malodour (James et al., 2013b).

1.3.2 Thioalcohols

The most recent class of odorous molecules to be implicated in axilla BO were the thioalcohols which were detected in axilla sweat by three independent groups in 2004 (Hasegawa et al., 2004; Natsch et al., 2004; Troccaz et al., 2004). Each group identified 3-methyl-3-sulfanylhexan-1-ol (3M3SH) (Table 1.3.2) as the primary odour contributing thioalcohol due to its abundance when isolated by GC and GC-MS methods and its association with strong odour intensity, described as “sweat and onion-like”. Hasegawa *et al.* and *Natsch et al.* also detected an additional three thioalcohols in axilla sweat of 3-sulfanylhexan-1-ol (3SH), 3-sulfanylpentan-1-ol and 2-methyl-3-sulfanylbutan-1-ol with Natsch able to link these to axilla malodour (Hasegawa et al., 2004; Natsch et al., 2004). The identified 3M3SH was also recorded to have a very low olfactory thresholds of 1 -3 parts per trillion in air (Natsch et al., 2004; Troccaz et al., 2009) making the thioalcohols the most pungent odorous molecules linked to axilla BO and likely the main contributor to malodour generation.

Natsch *et al.* (2004) also proposed that the precursor substrates for generation of odorous thioalcohols in the axilla were S-hydroxylalkyl-L-cysteine conjugates and were able to show both 3M3SH and 3SH release from their cysteine conjugates by axillary isolates of *Corynebacterium* (Natsch et al., 2004). In the same study they were also able to clone, sequence and heterologously express the gene encoding a cystathionine (C-S) β-lyase from *Corynebacterium* sp. Ax20 which was active against several synthesised L-cysteine-S-conjugates. This C-S β-lyase activity with cysteine conjugates was later shown to be relatively wide spread within axilla *Corynebacterium* strains when screened using a model substrate (S-benzyl-L-cysteine) (James et al., 2013b). Additionally, some *Staphylococcus* strains also displayed activity although only a minor subset (3 out of 17 tested) and at a lower average activity than that of the tested corynebacteria. This suggested multiple genera were capable of thioalcohol production in the axilla although *Corynebacterium* were still believed to be the primary contributors.

Table 1.3.2: Major odorous thioalcohols detected in axilla sweat		
Thioalcohol	Structure	Reference
3-methyl-3-sulfanylhexan-1-ol (3M3SH)		(Hasegawa et al., 2004; Natsch et al., 2004; Troccaz et al., 2004)
3-sulfanylhexan-1-ol (3SH)		(Hasegawa et al., 2004; Natsch et al., 2004)
3-sulfanylpentan-1-ol		(Hasegawa et al., 2004; Natsch et al., 2004)
2-methyl-3-sulfanylbutan-1-ol		(Hasegawa et al., 2004; Natsch et al., 2004)

A year after Natsch *et al.* reported S-hydroxylalkyl-L-cysteine conjugates as the precursors to thioalcohol generation, another group proposed an alternative route of thioalcohol production with precursor substrates originating from the apocrine gland as glutathione (GSH) conjugates (Starkenmann et al., 2005). These GSH-conjugates have never been detected in sampling of the axilla but are well known to breakdown to Cys-Gly conjugates (Habig et al., 1974) and S-hydroxyalkyl-L-cysteinylglycines were found within axilla sweat with the identification of S-Cys-Gly-3M3SH while detection of Cys-3M3SH was notably absent (Starkenmann et al., 2005). Synthetic homologues of these Cys-Gly-conjugates were synthesised and shown to be converted to their respective thioalcohol by incubation with a strain of *Staphylocccous haemolyticus* as well as less efficient release being observed by *Corynebacterium xerosis* (Starkenmann et al., 2005). S-Cys-Gly-3M3SH was later shown in significant levels in the axilla from a study of 49 male and female over three years, which also showed that female underarms possess significantly (up to 10-fold) more of the thioalcohol precursor than their male counterparts (Troccaz et al., 2009). The proposal of Cys-Gly-3M3SH being the primary secreted precursor in apocrine sweat was accepted by Emter and Natsch (2008) who showed that S-Cys-Gly-3M3SH was the predominant thioalcohol precursor in axilla sweat by liquid chromatography-mass spectrometry (LC-MS) (Emter and Natsch, 2008). This work did also detect Cys-3M3SH but at an 8-fold lower concentration than S-Cys-Gly-3M3SH, supporting the proposal that the dipeptide conjugate is the principal substrate for thioalcohol malodour. Alongside the acceptance of S-hydroxyalkyl-L-cysteinylglycines as the apocrine derived precursors, the biochemical mechanism of their conversion to thioalcohols was shown to be dependent on the sequential action of a dipeptidase, to generate the Cys-conjugate, and a C-S β-lyase (Fig 1.3.1). Similar to earlier work on the C-S β-lyase from *C*. sp. Ax20, the dipeptidase (TpdA) from the same organism was identified and recombinantly expressed. Characterisation of TpdA showed it to be part of the M20 family of metallopeptidases and able to hydrolyse S-Cys-Gly-3M3SH alongside several dipeptides. Enzymes of the M20 family are also noted to often have aminoacylase activity but incubations of TpdA with N^{α}-acyl-amino acids showed either no or very low activity.

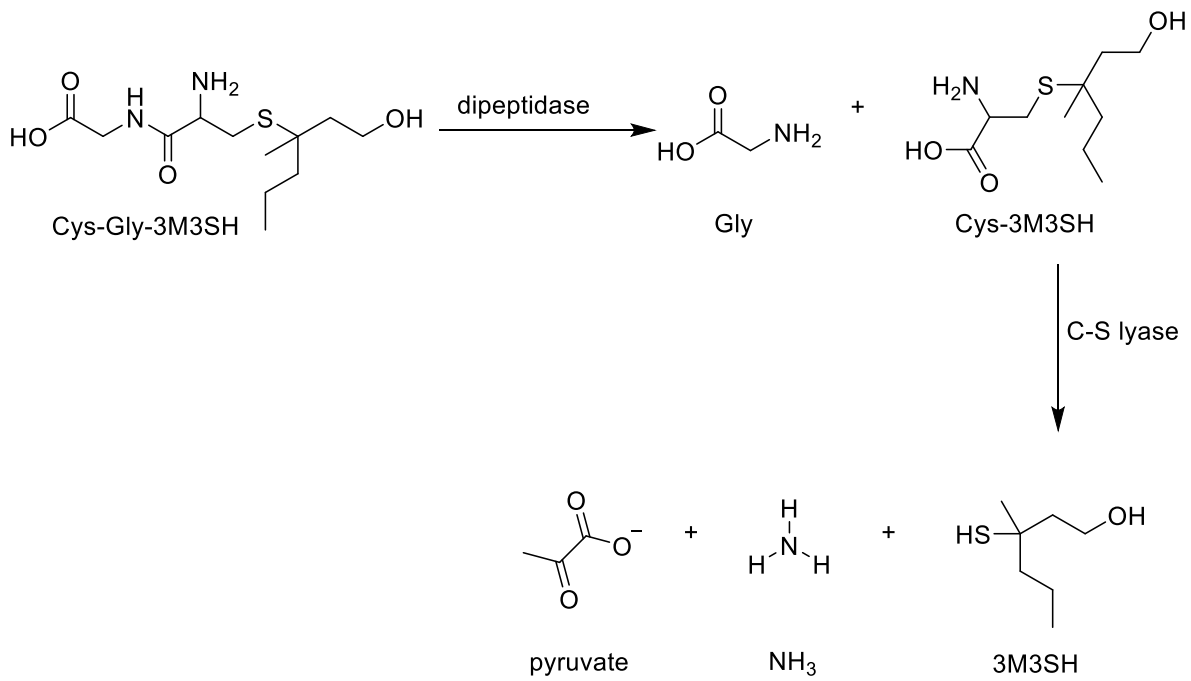


Figure 1.3.1: Biotransformation of Cys-Gly-3M3SH

While *S. haemolyticus* had been shown to generate 3M3SH from its cysteinylglycine precursor (Starkenmann et al., 2005), *Corynebacterium* were still believed to the primary causal organism responsible for thioalcohol release (James et al., 2013b). This view was changed in 2015 with the identification of three closely related species of *Staphylococcus*, *S. hominis, S. haemolyticus* and *S. lugdunensis*, that could uptake and metabolise S-Cys-Gly-3M3SH to release 3M3SH. Importantly, screening of *Corynebacterium* strains with the precursor showed most had minimal activity with only *C.* sp. Ax20 displaying noticeable 3M2SH release that was around four-fold less than the active *Staphylococcus* species. This differed to previous studies showing thioalcohol activity was prevalent within *Corynebacterium* (James et al., 2013b) but was shown to be caused by strains displaying different activity between the model substrate, S-benzyl-L-cysteinylglycine, and the physiological substrates (Bawdon et al., 2015). Other *Staphylococcus* strains were tested but showed negligible 3M3SH release including *Staphylococcus epidermidis* which is the most abundant staphylococci present on human skin (Kloos and Musselwhite, 1975), with recent evidence suggesting it is also the most prevalent species in the axilla (James, 2020). *Staphylococcus hominis* is the second most abundant species of the genera on skin and the axilla (James, 2020; Kloos and Musselwhite, 1975), and correlated to axilla malodour (Troccaz et al., 2015). Early axilla microbiota characterisation studies failed to find an association between *Staphylococcus* as a genus and malodour(Leyden et al., 1981; Taylor et al., 2003), with this now believed to be due to the dominance of *Staphylococcus* epidermidis within the microbiome which cant not generate thioalcohols In light of this evidence, *S. hominis* has been proposed as the main cause of thioalcohol malodour in the underarm (Bawdon et al., 2015; James, 2020).

The full biotransformation process of 3M3SH generation in *S. hominis* has now been almost entirely characterised (Minhas et al., 2018; Rudden et al., 2020) (Fig 1.3.2). S-Cys-Gly-3M3SH was shown to be actively transported across the membrane by the secondary transporter $PepT_{SH}$ (also known as DtpT), a member of the proton-dependent oligopeptide transporter (POT) family (Minhas et al., 2018). This transporter was co-crystalised with S-Cys-Gly-3M3SH bound, allowing deeper insight into substrate binding and revealed a large hydrophobic pocket which facilitated binding by accommodation of the thioalcohol group. $PepT_{SH}$ was also observed to transport di- and tri-Ala peptides but contained specific adaptation to enable S-Cys-Gly-3M3SH translocation. The *S. hominis* C-S β-lyase responsible for thioalcohol generation was later extensively characterised with biochemical, phylogenetic and structural techniques (Rudden et al., 2020). The identified pyridoxal phosphate (PLP) dependent enzyme, termed ShPatB, was shown to have high selectivity for the Cys-3M3SH precursor due to specific adaptations to enhance thioalcohol activity. Importantly only *Staphylococcus* strains containing a close homologue of ShPatB were able to release 3M3SH from its Cys-Gly precursors and

highlighted the same three species of *S. hominis, S. haemolyticus* and *S. lugdunensis* alongside two others of *S. devriesei* and *S. petrasii*. These species were shown to form part of a single monophyletic clade which were able to release 3M3SH and further supports the role of *S. hominis* as the causal organism of thioalcohol malodour. Structural determination (PDB 6QP1and 6QP2) and subsequent site directed mutagenesis of binding site residues revealed a unique adaptation to enable 3M3SH moiety accommodation in the binding site with a tyrosine (Tyr25) present on a flexible loop region which points into the binding site and contributes to a specific hydrophobic surface in ShPatB that orientates 3M3SH. This adaptation was only present in malodour producing staphylococci and suggested key adaptation of these microorganisms within the axilla to utilise this substrate. The same study also estimated with evolutionary phylogenetic analysis that this malodour producing *patB* gene was acquired in these staphylococci millions of years ago, suggesting that thioalcohol malodour generation is an ancient process that predates *Homo sapiens* and likely emerged at the same time of early primate divergence (Rudden et al., 2020). This estimation has interesting implications as to the evolutionary history of axilla BO and could suggest a role for thioalcohols in chemosensory communication in our past.

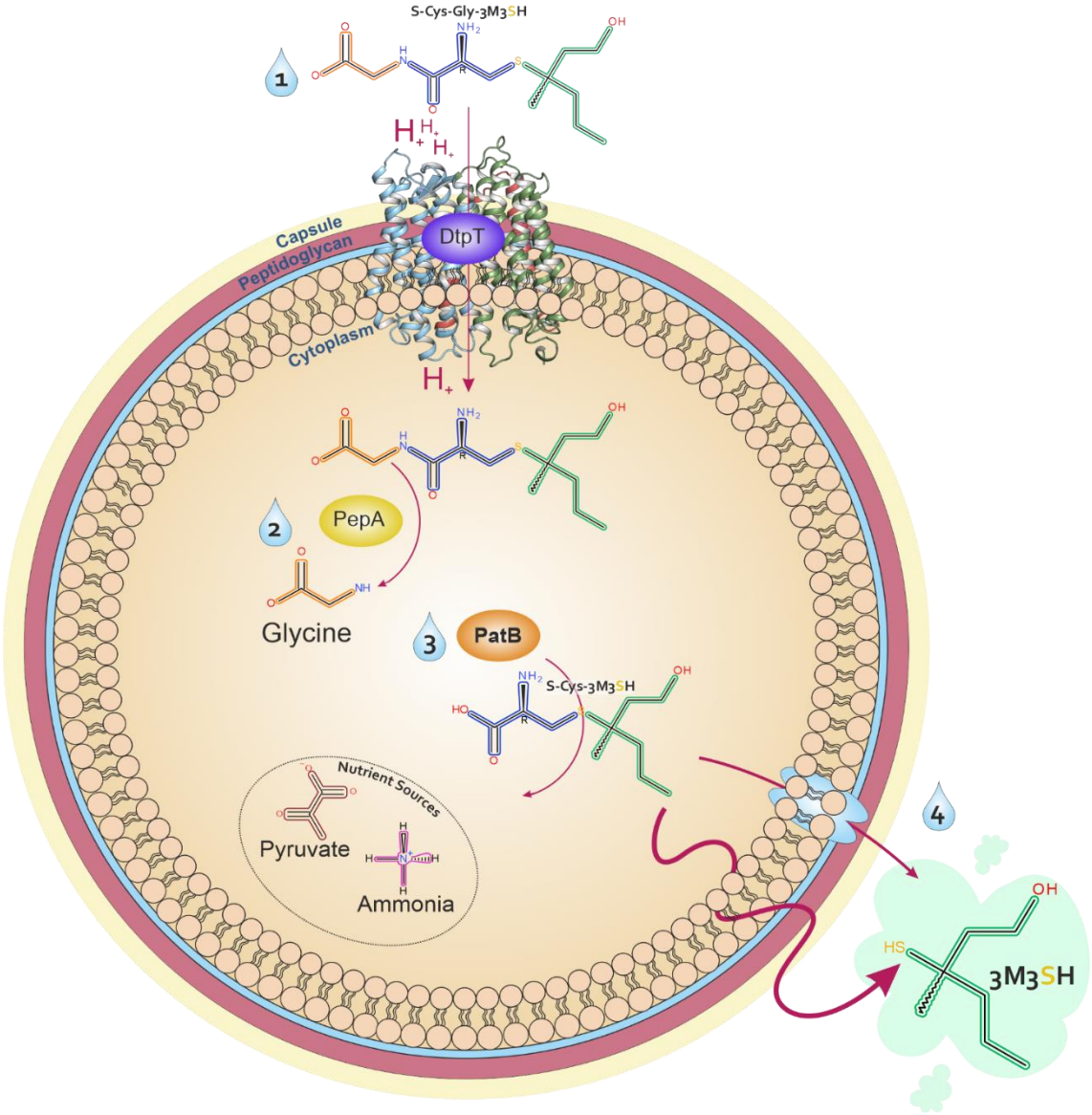


Figure 1.3.2: Biotransformation of S-Cys-Gly-3M3SH by *S. hominis*

S-Cys-Gly-3M3SH is imported across the membrane by the POT transporter DtpT (also known as $PepT_{SH}$). Once inside the terminal glycine is cleaved by a M20 peptidase (in *S. hominis* implicated to be PepA) to produce Cys-3M3SH which is acted upon by PatB to release 3M3SH and generate pyruvate and ammonia. The thioalcohol then crosses the membrane (either by passively diffusion or mediated an unknown exporter) to form part of axilla malodour. Adapted from (Rudden et al., 2020).

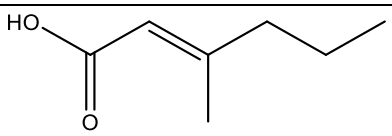
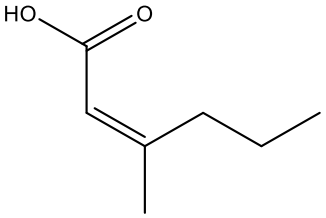
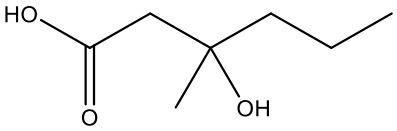
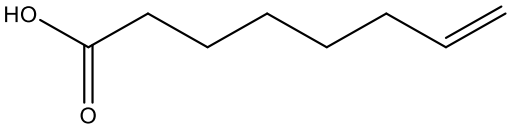
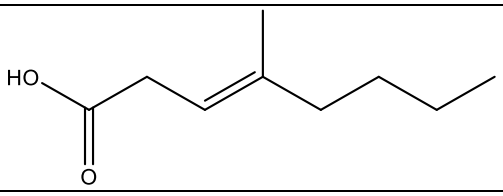
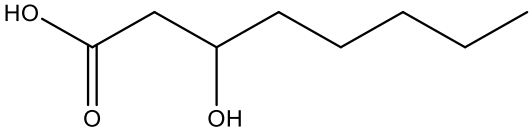
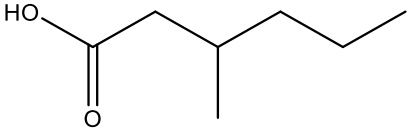
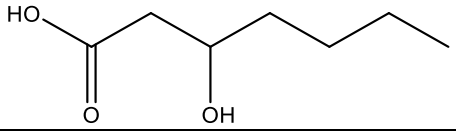
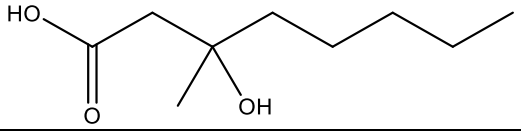
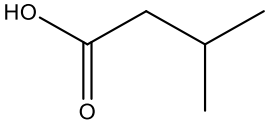
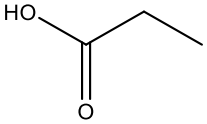
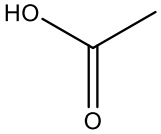
1.3.3 Volatile fatty acids

Volatile fatty acids (VFAs) have long been acknowledged to contribute to axilla malodour with early studies attributing BO to odorous steroids and short chain ($C_2 – C_5$) VFAs (Gower et al., 1988; Leyden et al., 1981). While odorous steroids have now been shown to play a very minor role in axilla BO, VFAs are still viewed as major odorants although primarily due to the presence of medium chain length ($C_6 – C_{10}$) VFAs.

Both *Cutibacterium* and *Staphylococcus* species have been shown to ferment glycerol, derived from triglyceride hydrolysis, and lactic acid to acetic and propionic acid (Table 1.3.3). However, there is little evidence to suggest either of these VFAs are major axillary odorants (James et al., 2004a). Staphylococci have been shown to produce isovaleric acid from branched aliphatic amino acids, particularly L-leucine (James et al., 2004a), which is a known odorant and has been proposed as the cause of the faint acidic odour in axilla BO (Leyden et al., 1981). Isovaleric acid is the main cause of foot odour (James et al., 2013a) and due to the presence of L-leucine across the human skin and within eccrine sweat (Harker and Harding, 2013) could be produced across the whole body rather than being localised to the axilla. Isovaleric acid was also proposed to be generated by partial degradation of unusual long chain fatty acids found in sebum (Downing et al., 1989; Greene et al., 1970) by a small subset of *Corynebacterium* species (James et al., 2004b). This position on *Corynebacterium* involvement was later revoked due to the scarcity of this subset of *Corynebacterium* in the axilla and the prevalence of staphylococci (James et al., 2013b). Though isovaleric acid generation appears wide spread within the *Staphylococcus* genus including the dominant species *S. epidermidis* (James et al., 2004a), this has never translated to a correlation at the genus level between abundance and axilla odour (Troccaz et al., 2015). In fact isovaleric acid is proposed to have only a minor contribution towards VFA based malodour with other medium chain VFAs being implicated as the causal VFA odorants (Zeng et al., 1991). The low impact of isovaleric acid on odour can be evidenced from its association with a “faint acidic odour” which was noted as distinct from typical apocrine axilla malodour (Leyden et al., 1981).

Zeng *et al.* (1991) showed that medium chain ($C_6 – C_{10}$) VFAs were present in axilla sweat using GC-MS combined with olfactory assessment. Several structurally unusual medium chain VFAs were identified in this study, with the most abundant being 3-methyl-2-hexenoic acid (3M2H) which was associated with “sweaty” and “musty” odour descriptors (Zeng et al., 1991). Both the E- and Z- isomers of 3M2H were detected in axilla sweat, with the trans isomer being 10-fold more abundant and associated with the “sweaty” odour quality rather than the more “acrid” cis isomer. Other VFAs were also identified with high odour intensity such as 7-octenoic acid as well as other unsaturated or branched fatty acids.

A small amount of the saturated form of 3M2H, 3-methylhexenoic acid but its odour impact was not described. Given the prevalence of 3M2H within axilla sweat and its clear odour impact, this molecule was proposed to have major contribution to axilla malodour.

Table 1.3.3: Major odorous VFAs detected in axilla sweat		
VFA	Structure	Reference
(E)-3-methyl-2-hexenoic acid E-(3M2H)		(Natsch et al., 2003; Zeng et al., 1991)
(Z)-3-methyl-2-hexenoic acid Z-(3M2H)		(Natsch et al., 2003; Zeng et al., 1991)
3-hydroxy-3-methylhexanoic acid (HMHA)		(Natsch et al., 2003)
7-octenoic acid		(Zeng et al., 1991)
(E)-4-methyloct-3-enoic acid		(Natsch et al., 2006)
3-hydroxyoctanoic acid		(Natsch et al., 2006)
3-methylhexanoic acid		(Zeng et al., 1991)
3-hydroxy-4-methylheptanoic acid		(Natsch et al., 2006)
3-hydroxy-3-methyloctanoic acid		(Natsch et al., 2006)
3-methylbutanoic acid (Isovaleric acid)		(James et al., 2004a)
Propionic acid		(James et al., 2004a)
Acetic acid		(James et al., 2004a)

While the odorous VFA, 3M2H, had been identified, the question remained as to what the odourless precursor of this molecule was. Four years later Spielman *et al.* (1995) proposed that 3M2H was transported to the skin surface non-covalently bound to two proteins termed apocrine secretion odour-binding proteins (ASOBs) before bacterial action, by an unknown bacterium, released the odorous VFA. One of these proteins, ASOB2, was shown to be identical to apolipoprotein D (ApoD), a member of the lipocalcin family of carrier proteins (Zeng et al., 1996). Lipocalcins have been shown to carry pheromones in rodents (Evershed et al., 1993; Singer and Macrides, 1990), which raised the interesting possibility that ASOB transport of 3M2H was part of a vestigial pheromone system in humans.

This theory was later challenged by Natsch *et al.* (2003) who identified 3M2H and the structurally related 3-hydroxy-3-methylhexenoic acid (HMHA), covalently bound to L-glutamine at the N^{α} atom in sterile axilla sweat. Both 3M2H and HMHA were detected by GC-MS in hydrolysed axilla secretions while their respective glutamine conjugated precursors (3M2H-Gln and HMHA-Gln) were identified in sterile axilla sweat by LC-MS. The chromatographically isolated precursors could also be shown to generate their respective VFA upon hydrolysis with NaOH at 100 °C, validating that these glutamine conjugates could release the odorants. The structure of 3M2H-Gln and HMHA-Gln suggested that bacterial action upon these molecules required the presence of an N^{α}-acylglutamine aminoacylase (N-AGA) (Fig 1.3.3). The VFA precursors were incubated with 44 different axilla strains of *Corynebacterium* (19 strains) and *Staphylococcus* (25 strains) for 24 h before analysis for VFA generation. This screen highlighted three species of *Corynebacterium* capable of significant 3M2H and HMHA release and were: *C.* sp. Ax20 (identified at the time as *C. striatum*), *C. jeikeium* and *C. bovis.* All tested *Staphylococcus* species were unable to hydrolyse the precursors, suggesting medium chain VFA malodour is caused by *Corynebacterium* and supporting evidence at the time which implicated corynebacteria in axilla malodour (Leyden et al., 1981; Taylor et al., 2003). All strains were also screened with another acylglutamine substrate, N^{α}-lauroyl-L-glutamine, which revealed additional strains, including some staphylococci could release lauric acid. This suggested that these other strains likely contained a homologue aminoacylase that was not specific for VFA precursors. *Corynebacterium* sp. Ax20 was chosen to characterise aminoacylase activity and the protein was observed to be active intracellularly and be a zinc dependent metallopeptidase from experiments on cell extracts. The putative N-AGA was cloned and recombinantly expressed in *E. coli*, with the purified enzyme showing high activity with both 3M2H-Gln and HMHA-Gln as well as other medium chain acylglutamine substrates (Natsch et al., 2003). The sequence of N-AGA indicated that the protein forms part of the large M20 family of metallopeptidases. The role of 3M2H and HMHA in VFA malodour generation was further supported with measurements of significant quantities of HMHA-Gln (113.5 mg/L) in the axilla

sweat which was shown to be the most abundant malodour precursor present, surpassing Cys-Gly-3M3SH by approximately 2:1 (Troccaz et al., 2009). This combined with the low olfactory threshold of HMHA and 3M2H, although not lower than that of 3M3SH (James et al., 2013b; Troccaz et al., 2009), reinforced their role as major contributors to axilla BO.

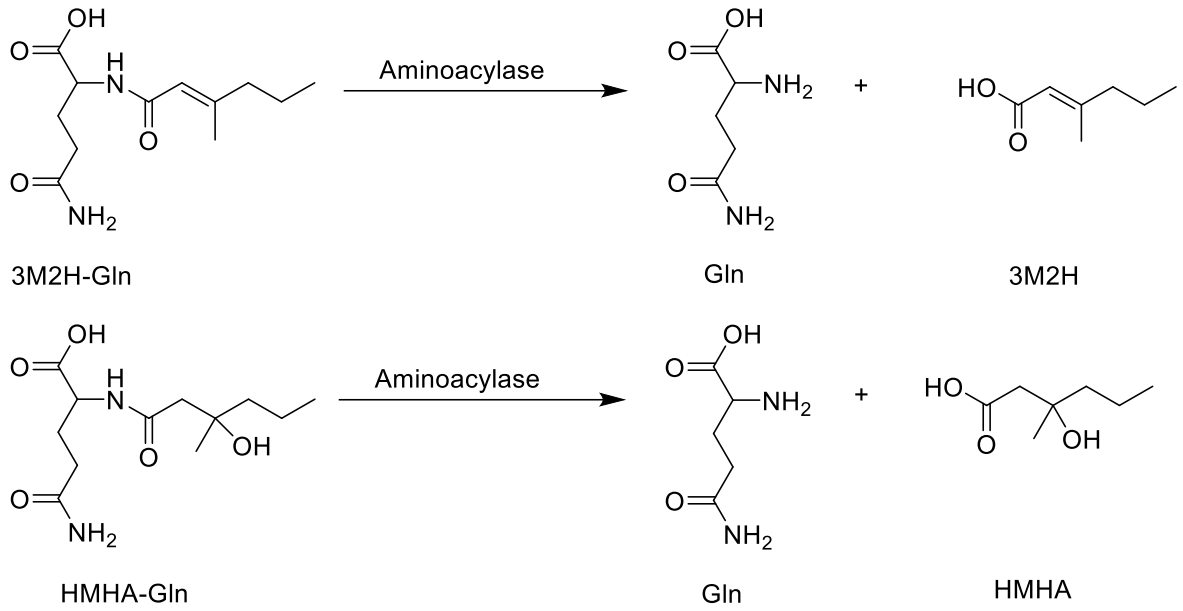


Figure 1.3.3: Chemical reactions catalysed by *Nα*-acylglutamine aminoacylase

The finding of high levels of glutamine conjugated 3M2H and HMHA disproves earlier assertions that VFAs are bound to ASOBs before release (Natsch et al., 2003; Spielman et al., 1995), a view which is now the widely accepted consensus (James, 2020; James et al., 2013b; Natsch and Emter, 2020; Troccaz et al., 2009). A theory was postulated that combined the ASOB association and glutamine conjugation to suggest that HMHA was covalently bound to the N-terminal glutamine residue of ApoD, which would be subsequently proteolytically cleaved to create HMHA-Gln in axilla sweat (Akiba et al., 2011). However, this proposal would require a one to one ratio of apoD to HMHA-Gln and given the high abundance of HMHA-Gln in the axilla (Troccaz et al., 2009) would require unrealistically high levels of protein in axilla sweat (James et al., 2013b).

The purified N-AGA protein was later shown to be able to release several other VFAs from axilla sweat that showed variation in acyl chain length and branching, and were identified as odorous (Natsch et al., 2006) (Table 1.3.3). This indicates that 3M2H and HMHA are not the only medium chain branched VFAs found in axilla sweat although HMHA was found to be by far the most abundant in all individuals tested, followed by the E-3M2H. All other VFAs isolated appeared at very minor levels of relative abundance compared to HMHA (<7 %) in all individuals, although variation was observed between the levels of these VFAs in each volunteer's axilla sweat. This led to the speculation that these low abundance VFAs may effect overall VFA malodour intensity although no work to date has conclusively shown any association except for 3M2H and HMHA (Troccaz et al., 2004, 2009; Zeng et al., 1991). While this work highlighted the presence of other VFAs within axilla sweat, those identified will have been influenced by the substrate specificity of the N-AGA enzyme, suggesting that those found do not represent the full repertoire of VFAs present in the axilla.

The work of Natsch *et al.* (2003) implicated *Corynebacterium* species as responsible for 3M2H and HMHA generation in the axilla. Later studies did show that several *Corynebacterium* strains did possess N^{α}-acylglutamine activity with the ability to hydrolyse 3M2H-Gln and a model substrate, carbobenzyloxy-L-glutamine (Z-Gln) (James et al., 2013b). However, the activity of these strains was generally low, except for the host of N-AGA, *C.* sp. Ax20, and showed insufficient 3M2H release to explain the levels detected in axilla sweat. Attempts to isolate more axilla bacteria or manipulation of growth media during VFA biotransformation assays failed to yield higher activity, suggesting that either the optimum conditions for VFA release had not been found or, more likely, that the major causal organism of VFA release had not been identified. *Corynebacterium* sp. Ax20 was initially identified as *Corynebacterium striatum* (Natsch et al., 2003) but later work with 16S rRNA showed it is more similar to *Corynebacterium glaucum* (James et al., 2013b). *Corynebacterium striatum* and *C. glaucum* have both been identified in the axilla at a relatively low abundance (James, 2020; Troccaz et al., 2015), suggesting that these species are unlikely to be the primary cause of VFA malodour. These

two species were also tested for VFA release and were shown to not produce significant levels of 3M2H comparable to *C.* sp. Ax20, making this strain appear to be unique within corynebacteria for its high activity (James et al., 2013b). More recent studies published by Natsch on VFA malodour have not implicated any other *Corynebacterium* strains (Natsch, 2015; Natsch and Emter, 2020), further highlighting the uniqueness of *C.* sp. Ax20.

In 2014, a previously understudied genus was implicated in VFA malodour with the discovery of a strain of *Anaerococcus* capable of HMHA release from its glutamine precursor (Fujii et al., 2014). This strain, *Anaerococcus* sp. A20, was isolated from the axilla and shown to hydrolyse HMHA-Gln upon incubation of the precursor with mechanically disrupted cells for 24 h, with significant HMHA release being detected by GC-flame ionisation detection (GC-FID). A total of 19 strains of anaerobic bacteria including *Cutibacterium* and *Anaerococcus* were screened for HMHA release with only *A.* sp. A20 displaying activity with the precursor. Some strains of *Anaerococcus* did display very minor levels of HMHA release including *A. octavius*, *A. murdochii* and *A. hydrogenalis*, although all were below 1.5 % of total HMHA release compared to 79.8 % for *A.* sp. A20 from incubation with 2 mM HMHA-Gln. A similar screen for 3M2H-Gln activity was not reported. Nonetheless, this study showed that a novel *Anaerococcus* species was capable to VFA malodour generation, showing that *Corynebacterium* was not the sole genus capable of medium chain VFA generation. The strain *A.* sp. A20 was later classified as *Anaerococcus nagyae* (Veloo et al., 2016) and remains the single *Anaerococcus* species to date implicated in VFA malodour. This finding of *Anaerococcus* malodour activity is supported by the association observed between *Anaerococcus* abundance in the axilla and a "fatty" odour (Troccaz et al., 2015), with more recent work also showing that *Anaerococcus* is prevalent in the axilla and linked to axilla malodour (Kim et al., 2021). Together this suggests that *Anaerococcus* likely has an important role in axilla BO that is currently not well understood.

While the intracellular enzyme responsible for 3M2H and HMHA release has been characterised in *Corynebacterium* sp. Ax20, no similar characterisation has occurred within *Anaerococcus*. Furthermore, the microbial biotransformation process of medium chain VFA generation by bacteria in the axilla has not been studied in detail, making the mechanism of glutamine precursor uptake and efflux of VFAs currently unknown (Fig 1.3.4).

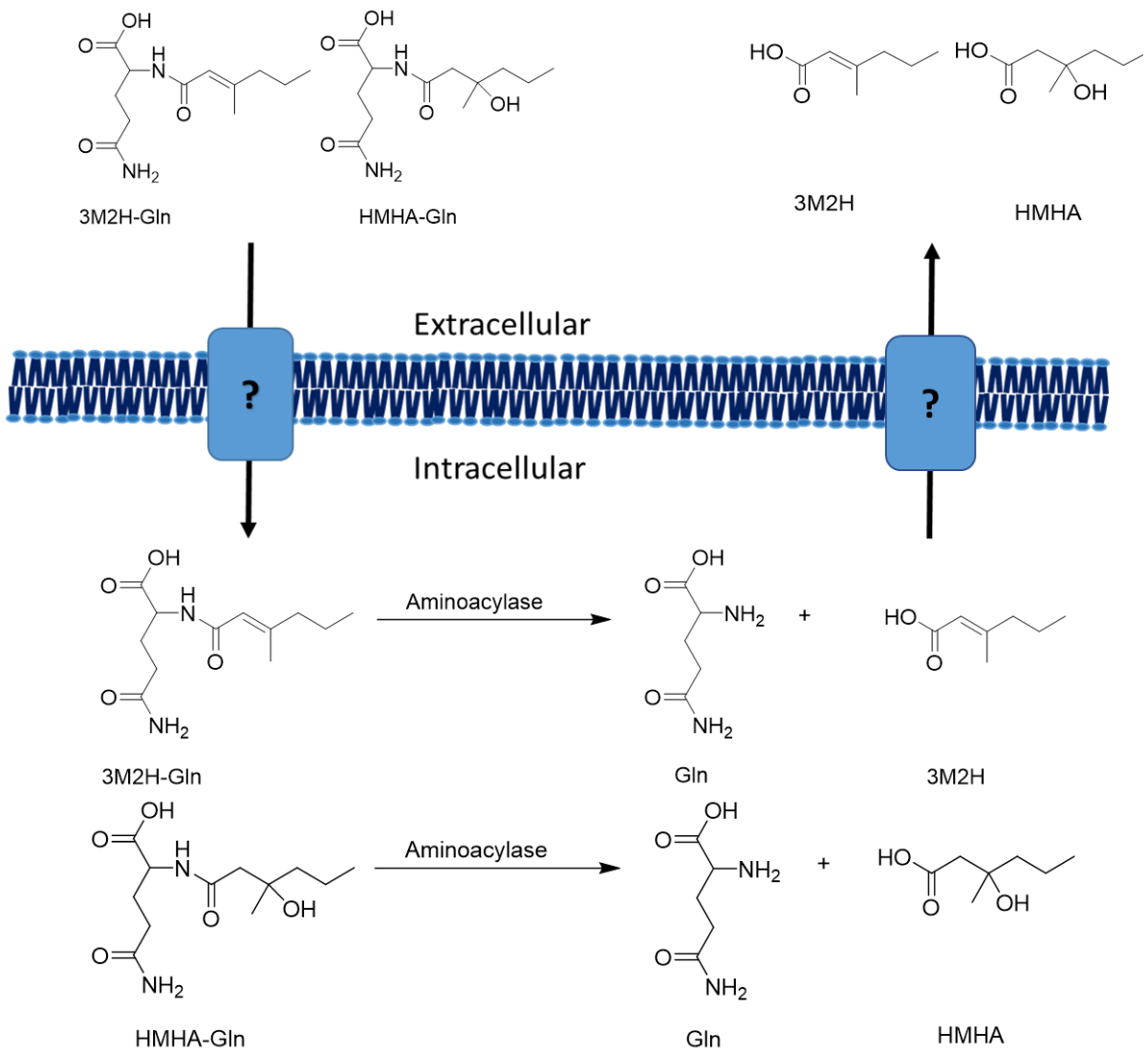


Figure 1.3.4: Biotransformation of 3M2H-Gln and HMHA-Gln

The VFA precursors, 3M2H-Glna and HMHA-Gln, are present in apocrine sweat and imported inside VFA malodour producing bacteria by an unknown transporter. Once transported the VFA precursors are both hydrolysed by an aminoacylase with N^{α}-acylglutamine activity, yielding L-glutamine for microbial metabolism and releasing the VFAs, 3M2H and HMHA. These VFAs leave the cell either by passive diffusion or a transporter protein mediated process, whereby they are volatilised and contribute to axilla BO.

1.4 M20 aminoacylases and N-AGA

1.4.1 The M20 family

The N^{α}-acylglutamine aminoacylase (N-AGA) from *C. sp. Ax20* resides within the M20 family of metallopeptidases which represents a large group of enzymes whose members have exhibited aminoacylase, aminopeptidase, carboxypeptidase and dipeptidase activity. This diversity in activity is evidenced in the breadth of research on these enzymes which has documented a wide variety of functions and substrate specificities within this family including hydrolysis of the plant auxin indole-3-acetic acid (IAA) (Bitto et al., 2009), xenobiotic metabolism (Cao et al., 2020) and enantiomer selectivity in amino acid production from racemases (Tanimoto et al., 2008).

Members of the M20 family are comprised of predominantly homodimeric but also monomeric proteins which contain divalent metal cofactors; typically using zinc or cobalt atoms for activity which are coordinated by conserved amino acid motifs within the active site (Rawlings and Barrett, 2013). A monomer has two domains, a larger catalytic domain which assumes an αβα sandwich fold containing a bimetallic binding site within the active site and a smaller oligomerisation/satellite domain. The catalytic domain contains the 5 metal coordinating residues, with one of these residues ligating to both metal ions. The oligomerisation domain is inserted within the catalytic domain sequence and forms an αβ fold that is key to quaternary structure formation in homodimeric enzymes (Palm and Röhm, 1995; Rowsell et al., 1997). This oligomerisation domain also plays a role in substrate binding with a conserved arginine that points into its own monomer's binding site where it interacts with the carboxylate group of the amino acid substrate (Jozic et al., 2002; Natsch and Emter, 2020). Other conserved residues within the oligomerisation domain, notably a histidine and asparagine, are distal to their own chain's active site and instead are close to their counterpart monomer's binding site and have been shown to have a role in catalytic function, suggesting dimerization is required for catalytic activity (Lindner et al., 2003). Evidence of a functional monomeric enzyme has been documented with PepV from *Lactobacillus delbrueckii* which contains an extended oligomerisation domain, termed a "lid" domain, that wraps over the binding site to form the binding cavity and contains the same conserved residues observed in their dimeric counterparts, thereby mimicking the interactions provided by a homodimer (Jozic et al., 2002). This larger lid domain has also been reported in other M20 peptidases with one assuming a monomeric state in solution (Girish and Gopal, 2010), while another has been observed in both homodimer and monomer states (Chang et al., 2010). The observation of a homotetramer state for some aminoacylases has also been observed and is quite commonly the noted conformation in crystallographic structures of members (Botelho et al., 2011; Girish et al., 2013; Natsch and Emter, 2020).

Due to the large size of the M20 family this grouping can also be further separated into subfamilies, namely M20 A, B, C, D, F and unclassified based on the conservation of active site residues including the metal cofactor coordinating residues. N-AGA is part of the M20D subfamily which has the noticeable distinction of using a cysteine rather than aspartate as the metal binding residue that interacts with both metal ions (Bitto et al., 2009; Girish et al., 2013; Jamdar et al., 2015; Jozic et al., 2002; Minasov et al., 2005; Natsch and Emter, 2020). A representation of the arrangement of the metal ions and their respective ligating residues can be seen in Fig 1.4.1. The two metal ions are believed to both be catalytic, being required for the activation of the catalytic water molecule which attacks the carbonyl carbon of the substrate's acyl group and forms the tetrahedral intermediate formed from this nucleophilic attack. The catalytic water molecule is activated and coordinated by a glutamate residue, that is believed to act a catalytic base and is adjacent to the metal coordinating Glu (Lindner et al., 2003). The coordination of this tetrahedral intermediate within the active site has been mimicked in N-AGA using a competitive inhibitor containing a non-cleavable phosphinic group (Natsch and Emter, 2020).

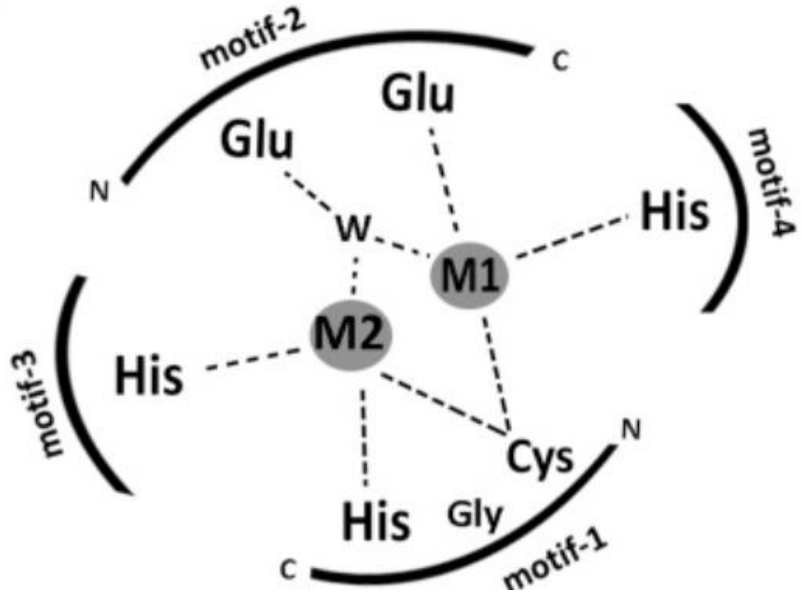


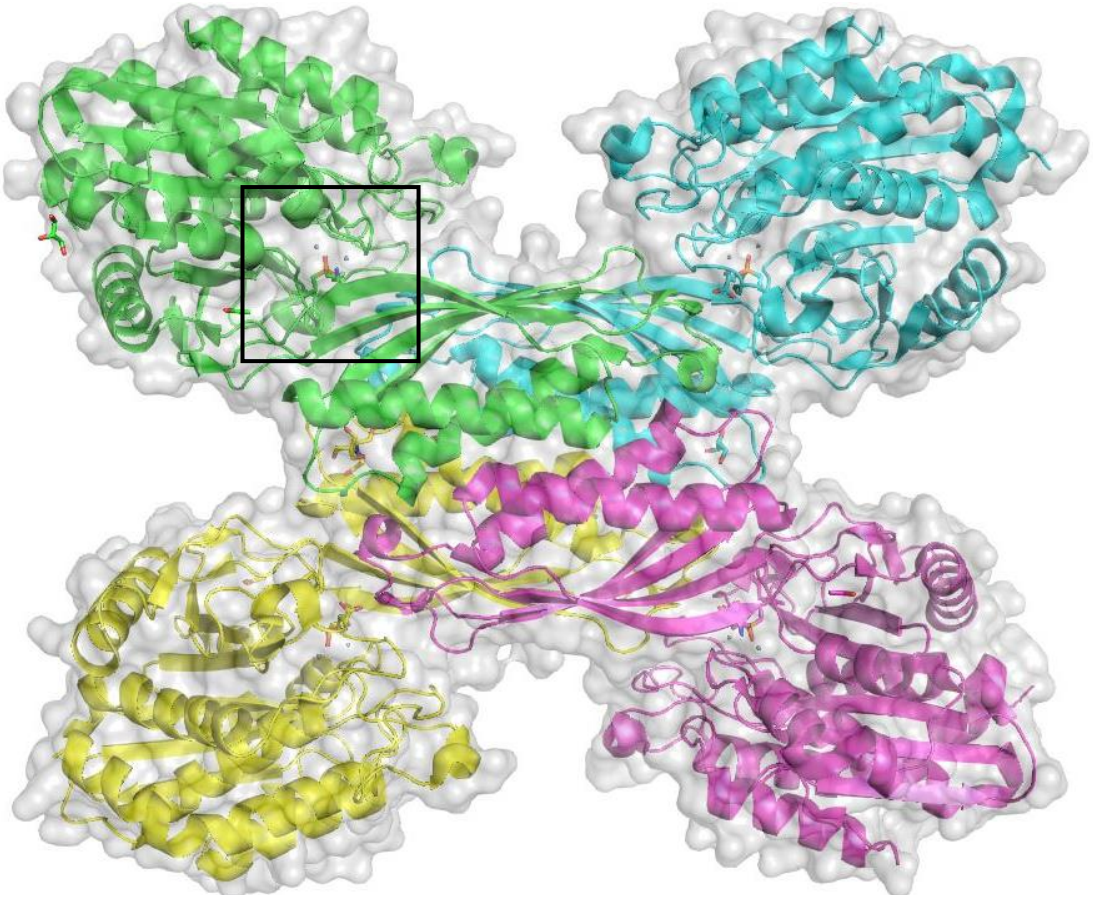
Figure 1.4.1: Schematic of M20D peptidase active site.

Representation of active site of M20D peptidase family showing conserved metal coordinating residues and other key conserved residues upon their respective structural motifs. The 5 metal binding residues are His, His, His, Glu and Cys. The catalytic water that attacks the carbonyl group of the substrate is coordinated by the two metal ions and a Glu residue and is depicted as a W. Taken from (Jamdar et al., 2015).

1.4.2 The structure of N-AGA

The structure of N^{α}-acylglutamine aminoacylase (N-AGA) from *C.* sp. Ax20 with a bound competitive inhibitor has been published (Fig 1.4.2 A) (Natsch and Emter, 2020), allowing structural insight into the 3M2H and HMHA producing enzyme and its likely binding interactions. This inhibitor, termed LJ8, contains the substructure of glutamine, a phosphinic acid group that imitates the tetrahedral transition state of a natural substrate and a large acyl group (Fig 1.4.2 B). The structure shows the key metal and substrate interacting residues within the active site. Looking at the metal binding site five residues: Cys105, His107, His168, His370 and Glu141 can be observed interacting with the two zinc atoms (Fig 1.4.3 A). The two oxygen atoms of the phosphinic acid group are also coordinating the two zinc atoms which is mimicking the interactions which would occur in the tetrahedral transition state of the natural substrates. Looking closer at the Gln moiety within the inhibitor and its interactions you can see Asn191, Asp301 interact with the glutamine side chain through water coordination while Arg264 and Asn251 interact with the carboxylate (Fig 1.4.3 B). Asn251 originates from a second monomer's oligomerisation domain, which also occurs for His202 which interacts with an oxygen atom bound to phosphorous. These residues are the same as that noted to be key for catalysis in hACY1 (Lindner et al., 2003, 2005) and other M20 proteins (Jozic et al., 2002; Nocek et al., 2018). The acyl portion of the inhibitor molecule is positioned in a large hydrophobic pocket formed by His168, Phe170, Phe185, Leu186, Leu335, Gly337 and Met354 which does not assume any specific interactions with the inhibitor (Fig 1.4.3 C). The large size and hydrophobic nature of this pocket is suggested to allow the accommodation of a variety of acyl groups while substrate specificity appears to be mainly provided by the interactions with the Gln moiety of substrates (Natsch and Emter, 2020). This structural observation supports previous biochemically characterisation of N-AGA which observed activity was limited almost exclusively to Gln containing substrates, whereas greater promiscuity was observed in the acyl group (Natsch et al., 2003). This promiscuity in acyl moiety also extended to N^{α}-acylglutamine precursor molecules present in axilla sweat with N-AGA able to release a variety of VFAs (Natsch et al., 2006).

A



B

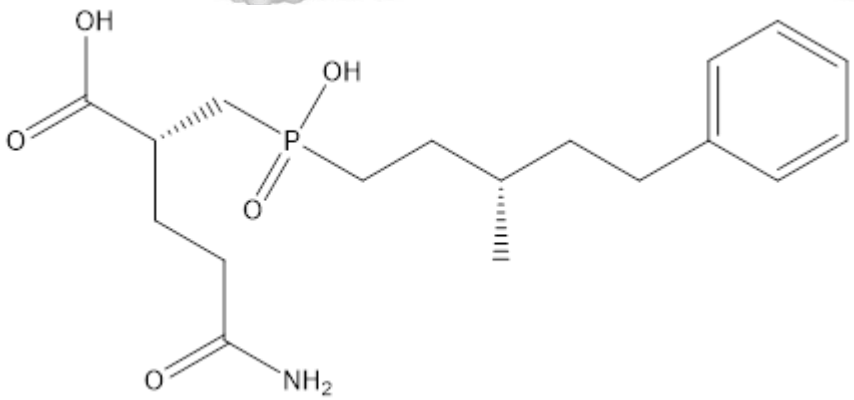
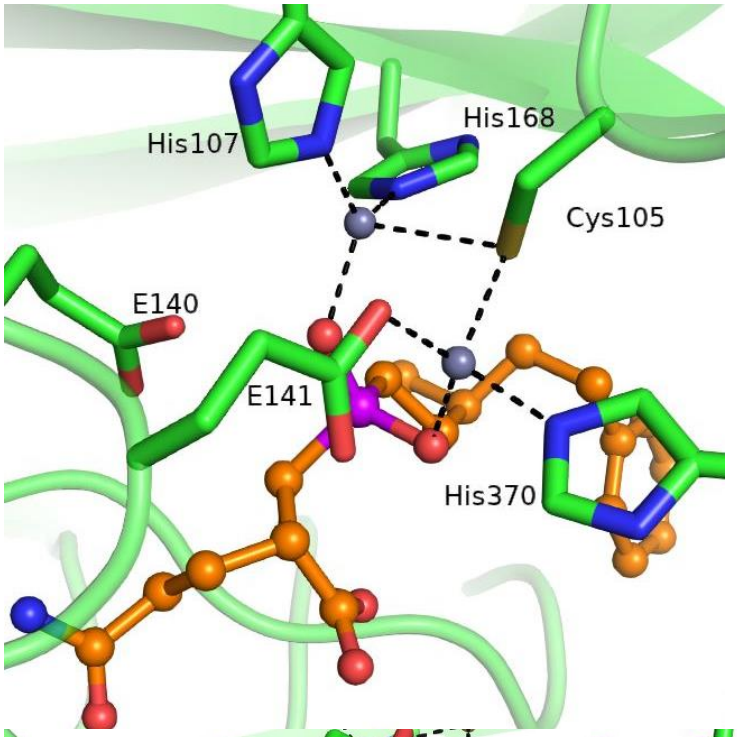


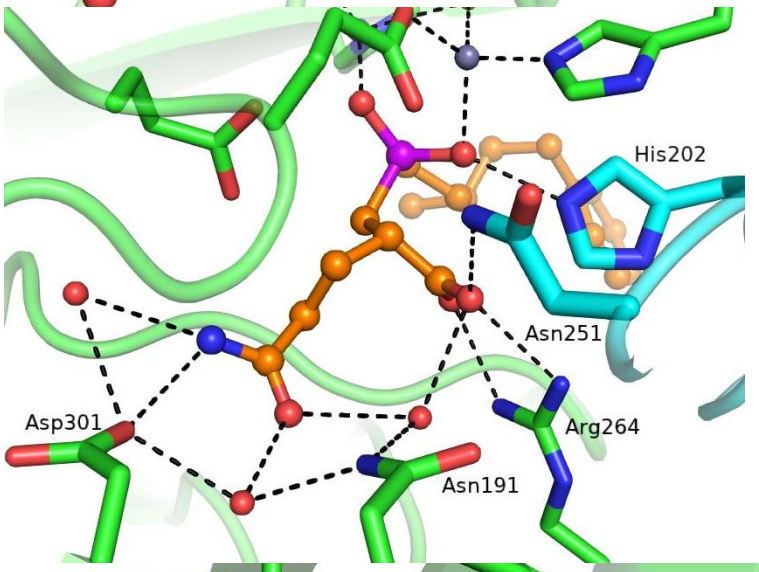
Figure 1.4.2: Structure of N-AGA and bound inhibitor

A. Homotetramer arrangement of N-AGA (PDB 6SLF) visualised with molecular surface map in transparent grey. Each monomer chain is coloured separately. Black box highlights one of the four active sites present in the tetramer with each positioned at the interface of catalytic and oligomerisation domain.
B. The chemical structure of the bound competitive inhibitor (2~{S})-5-azanyl-2-[[[(3~{S})-3-methyl-5-phenyl-pentyl]-oxidanyl-phosphoryl]methyl]-5-oxidanylidene-pentanoic acid termed LJ8.

A


His107
His168
Cys105
E140
E141
His370

B


His202
Asn251
Asp301
Arg264
Asn191

C

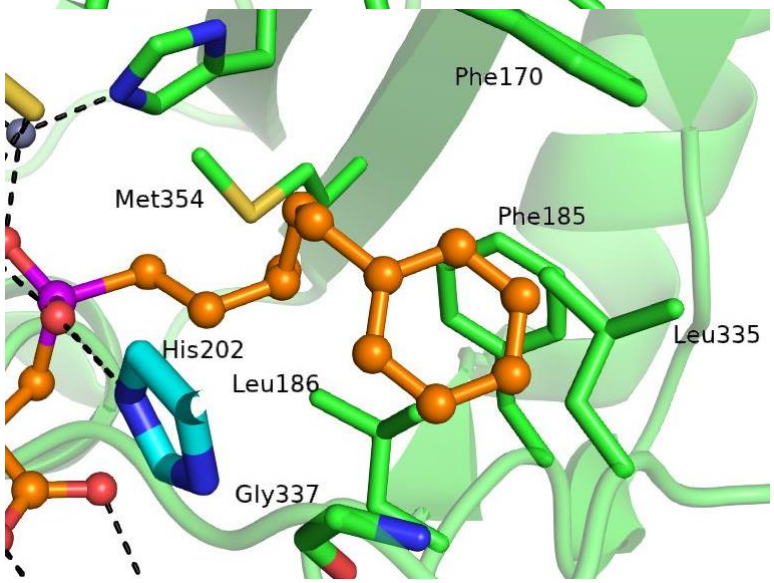

Phe170
Met354
Phe185
Leu335
His202
Leu186
Gly337

Figure 1.4.3: N-AGA binding of competitive inhibitor

A. Zinc coordination and their interaction with phosphorous bound oxygens. Zinc atoms shown as grey spheres. The competitive inhibitor, termed LJ8, is coloured orange.
B. Interaction with glutamine moiety of inhibitor. One monomer is shaded green, the second is coloured cyan to differentiate.
C. The large acyl binding pocket residues.

1.5 Gram-positive anaerobic cocci and *Anaerococcus*

The term Gram-positive anaerobic cocci (GPAC) refers to a prevalent heterologous group of anaerobic bacteria which represent approximately 25 % of all Gram-positive anaerobic bacteria isolated form clinical specimens (Murdoch, 1998). This group consists of several genera including *Anaerococcus, Peptoniphilus* and Finegoldia among others, and are commonly found within the normal human microbiota with isolation from skin and mucosal surfaces (Murdoch, 1998; Murphy and Frick, 2013). GPAC are also prevalent in the axilla with *Anaerococcus* and *Peptoniphilus* being noted as the fourth and fifth most abundant genera (Troccaz et al., 2015). GPAC strains and *Anaerococcus* are classed as opportunistic pathogens and much of the work today on these microorganisms focuses on their identification and role in clinical infections, often in polymicrobial infections (Brazier et al., 2008; Murdoch, 1998; Wildeboer-Veloo et al., 2007).

The genus *Anaerococcus* belongs within the phylum Firmicutes and is a relatively newly defined genus after extensive taxonomic revision in the past (Ezaki et al., 2001). This genus was formed after reclassification of a large and heterologous genus *Peptostreptococcus* to create three new genera of *Anaerococcus, Peptonihphilus and Gallicola*. *Anaerococcus* use peptone and amino acids as major energy sources and produce butyrate (Ezaki et al., 2001). This genus, and all GPACS, are classed as obligate anaerobes although *Anaerococcus* displays a surprising level of oxygen tolerance with two species (*A. rubiinfantis* and *A.* senegalensis) being observed to have very minor growth under microaerophilic conditions (Dahmani et al., 2018; Tidjani Alou et al., 2016). Little is known about the mechanism of colonisation by *Anaerococcus* on human skin or its spatial localisation within this environment, but the requirements for anaerobic conditions could suggest colonisation of hair follicles and glands similar to that observed for *Cutibacterium* (Leeming et al., 1984). The prominence of *Anaerococcus* in wounds during polymicrobial infections has also led to the proposal that they may exist in close proximity to aerobic colonisers that deplete local oxygen levels, facilitating anaerobic bacterial growth (Murphy and Frick, 2013). The abundance of follicles and glands in the axilla, combined with the microbial density of the site suggests that *Anaerococcus* is ideally suited to this environmental niche, as evidenced by it abundance in this environment (Troccaz et al., 2015).

1.6 Project aims

This work aims to characterise medium chain VFA malodour generation in *Anaerococcus* species. The putative aminoacylases that could generate VFAs in two species of *Anaerococcus, A. octavius* and *A. nagyae*, will be identified and recombinantly expressed in *E. coli* to discover which are responsible for VFA release.

Identified VFA liberating aminoacylases will be characterised using biochemical, biophysical and structural techniques to understand their function and inform on the ability of *Anaerococcus* to generate VFA malodour.

In detail characterisation of VFA generating aminoacylases by X-ray crystallography and site directed mutagenesis will provide further insight into their mechanism of action and substrate binding and allow comparison to the known VFA liberating enzyme, N-AGA.

Finally, the ability of GPAC species to generate 3M2H and HMHA from their glutamine conjugates will be assessed with an *in vivo* biotransformation screen to inform as to the prevalence of VFA malodour activity, particularly within *Anaerococcus* and *Peptoniphilus*. Further work using the *in vivo* biotransformation assay with *Escherichia coli* recombinantly expressing a VFA liberating aminoacylase will be performed to validate aminoacylase function in VFA malodour and act as a model to study VFA precursor uptake.

Chapter 2 -Materials and Methods

2 Materials and Methods

2.1 Chemicals and strains

2.1.1 Chemicals

All chemicals were procured at reagent grade purity unless otherwise stated and purchased from Merck (Sigma Aldrich) unless otherwise specified.

2.1.2 Volatile fatty acid precursors

Both 3-methyl-2-hexenoic acid (3M2H) and 3-hydoxy-3-methyl-hexenoic acid (HMHA) conjugated by an amide bond to the N^{α}-amino group of L-glutamine (3M2H-Gln and HMHA-Gln respectively) were custom synthesised by Concept Life Sciences and stored at 4 °C.

2.1.3 Antibiotics

Antibiotics were procured from Melford (ampicillin sodium salt) and TCI chemicals (kanamycin sulphate). Stock solutions of each antibiotics were made up at 1000x concentration (100 mg/ml ampicillin and 60 mg/ml kanamycin) in dH_2O, filter sterilised through a 0.22 µm syringe filter and stored at 4 °C. Working concentrations of antibiotics used were 100 µg/ ml ampicillin and 60 µg/ml kanamycin.

2.1.4 Plasmid expression inducing reagents

Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Melford) or arabinose was used to induce plasmid expression depending on expression vector utilised. IPTG was dissolved in dH_2O at 1000x concentration (500 mM), filter steilised through a 0.22 µM syringe filter and stored at -20 °C. Arabinose was dissolved in dH_2O at 1000x concentration (100 mg/ml), filter sterilised using a 0.22 µM syringe filter and stored at 4 °C. Working concentrations of inducer used were 0.5 mM IPTG and 100 µg/ml arabinose.

2.1.5 Media

Liquid and solid media were prepared by dissolving appropriate quantities of solid/ powdered constituents in dH_2O before autoclaving at 121 °C, 15 Psi, 15 min. For solid media such as agar plates, molten agar was cooled to approximately 60 °C before antibiotic addition if required. Roughly 20 ml

molten agar was then poured into individual petri dishes aseptically before being allowed to cool and solidify. Poured agar plates were then dried in a laminar flow hood and stored at 4 °C before use.

Luria-Bertani (LB) medium – contained (per 1 L dH_2O): 10 g Tryptone (Formedium), 5 g yeast extract, 10 g NaCl. For agar technical number 3 (Oxoid) was added to a final concentration of 1.5 %. pH 7.

Tryptic soy broth + Tween (TSBT) – contained (per 1 L dH_2O): 30 g tryptic soy broth, 10 g yeast extract and 1 g Tween 80 (Thermo Fisher). pH 7.

Blood Agar Base No. 2 with Horse blood (Oxoid) – procured as pre-poured plates. contained (per 1 L dH_2O): 15 g proteose peptone, 2.5 g liver digest, 5 g yeast extract, 5 g sodium chloride, 12 g agar and 7 % sterile horse blood. pH 7.4.

Fastidious anaerobic agar plates – contained (per 1 L dH_2O): 46 g anaerobe basal agar (Oxoid), 3 g peptone, 5 g Tween 80 and 5 % defibrinated horse blood (Oxoid). pH 7.4.

2.1.6 Strains used

All strains were stored as glycerol stocks (overnight culture + 15 % glycerol) at – 80 °C for long term storage. In the case of routine, short term use strains were streaked onto an agar plate and stored at 4 °C for up to 1 week.

Table 2.1.1: List of strains used in this work

Strain name	Genotype	Source
Escherichia coli DH5α	F-, *Δ(argF-lac)169*, ϕ80d*lacZ58*(M15), *ΔphoA8*, *glnX44*(AS), *λ⁻*, *deoR481*, *rfbC1*, *gyrA96*(NalR), *recA1*, *endA1*, *thiE1*, *hsdR17*	Invitrogen
E. coli BL21 (DE3)	F⁻*omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm* (DE3)	Invitrogen
E. coli BW25113	F-, Δ(*araD-araB*)567, Δ*lacZ*4787(::*rrnB*-3), lambda-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514	(Baba et al., 2006)
E. coli ATCC 10536	Wildtype	ATCC
E. coli JW0795-1	BW25113, *ΔglnP::kan*	(Baba et al., 2006)
E. coli JW0794-3	BW25113, *ΔglnQ::kan*	(Baba et al., 2006)
E. coli JW4038-1	BW25113, *ΔgltP::kan*	(Baba et al., 2006)
E. coli DB1	BW25113, *ΔoppC, ΔdppC*	(Bawdon, 2014)
E. coli DB50	BW25113, *ΔdtpA, ΔdtpB, ΔdtpC, ΔdtpD*	(Bawdon, 2014)
Anaerococcus octavius DSM 10024	Wildtype	DSMZ
Anaerococcus octavius DSM 11663	Wildtype	DSMZ
Anaerococcus hydrogenalis DSM 7454	Wildtype	DSMZ
Anaerococcus lactolyticus DSM 7456	Wildtype	DSMZ
Anaerococcus pacaensis DSM 26346	Wildtype	DSMZ
Anaerococcus nagyae DSM 101193	Wildtype	DSMZ
Anaerococcus vaginalis DSM 7457	Wildtype	DSMZ
Anaerococcus vaginalis DSM 25446	Wildtype	DSMZ
Anaerococcus prevotii DSM 20473	Wildtype	DSMZ
Anaerococcus prevotii DSM 20548	Wildtype	DSMZ
Anaerococcus tetradius DSM 2951	Wildtype	DSMZ
Anaerococcus tetradius DSM 20358	Wildtype	DSMZ
Anaerococcus senegalensis DSM 25366	Wildtype	DSMZ

Anaerococcus degeneri DSM 29674	Wildtype	DSMZ
Anaerococcus murdochii DSM 21462	Wildtype	DSMZ
Anaerococcus rubeinfantis DSM 101186	Wildtype	DSMZ
Peptoniphilus porci DSM 104947	Wildtype	DSMZ
Peptoniphilus lacydonensis DSM 100661	Wildtype	DSMZ
Peptoniphilus asaccharolyticus DSM 20463	Wildtype	DSMZ
Peptoniphilus harei DSM 10020	Wildtype	DSMZ
Peptoniphilus gorbachii DSM 21461	Wildtype	DSMZ
Peptoniphilus grossensis DSM 25475	Wildtype	DSMZ
Peptoniphilus pacaensis DSM 101839	Wildtype	DSMZ
Finegoldia magna DSM 20470	Wildtype	DSMZ
Peptoniphilus sp. 022	Wildtype	DSMZ
Peptoniphilus stercorisuis DSM 27563	Wildtype	DSMZ

2.2 DNA Manipulation and molecular biology techniques

2.2.1 Plasmids used

All plasmids were stored at – 20 °C in dH_2O.

Table 2.2.1: List of plasmids used in this work

Plasmid name	Description	Resistance	Source
pGM1	pET derived expression vector from GeneMill. IPTG inducible expression and C-terminal His tag. Encodes *oct1* gene codon optimised for *E. coli* expression.	Amp	GeneMill
pGM2	pET derived expression vector from GeneMill. IPTG inducible expression and C-terminal His tag. Encodes *oct2* gene codon optimised for *E. coli* expression.	Amp	GeneMill
pGM3	pET derived expression vector from GeneMill. IPTG inducible expression and C-terminal His tag. Encodes *oct3* gene codon optimised for *E. coli* expression.	Amp	GeneMill
pGM4	pET derived expression vector from GeneMill. IPTG inducible expression and C-terminal His tag. Encodes *oct4* gene codon optimised for *E. coli* expression.	Amp	GeneMill
pGMAx20	pET derived expression vector from GeneMill. IPTG inducible expression and C-terminal His tag. Encodes *n-aga* gene codon optimised for *E. coli* expression.	Amp	GeneMill
pET-21b(+)	High copy number, IPTG inducible expression vector. Contains C-terminal His tag. Empty vector codon optimised for *E. coli* expression.	Amp	Novagen
pET-21b-Nag3	High copy number, IPTG inducible expression vector. Contains C-terminal His tag. Vector encodes *nag3* gene codon optimised for *E. coli* expression.	Amp	This work
pET-21b-Nag5	High copy number, IPTG inducible expression vector. Contains C-terminal His tag. Vector encodes *nag5* gene codon optimised for *E. coli* expression.	Amp	This work
pETYSBLLIC3C	Derivative of pET28a plasmid. High copy number, IPTG inducible expression vector containing 3C protease site within linker region to N-terminal His tag. Empty vector.	Amp	(Fogg and Wilkinson, 2008)
pETYSBLLIC3c-Nag3	Derivative of pET28a plasmid. High copy number, IPTG inducible expression vector containing 3C protease site within linker region to N-terminal His tag. Vector encodes *nag3* gene codon optimised for *E. coli* expression.	Amp	This work
pBADcLIC2005	Arabinose inducible expression vector. Contains C-terminal 10x His tag.	Amp	(Geertsma et al., 2008)
pBADcLIC-Nag3	Arabinose inducible expression vector. Contains C-terminal 10x His tag. Encode *nag3* gene.	Amp	This work

2.2.2 Synthesised genes used

In the case of cloning of genes from *A. nagyae,* the desired genes were ordered by custom gene synthesis using Thermo Fisher Gene String service. The DNA sequence of the gene of interest was inputted and subsequently codon optimised for expression in *E. coli*. Additionally, a 20 bp regions of homology to the pET-21b vector was added to both the 5' and 3' ends of the gene to enable direct homologous recombination cloning into that vector. Synthesised genes were dissolved to a final concentration of 25 ng/µl in dH_2O and stored at -20 °C. This service was performed for the DNA sequence of *nag3* (protein id: WP_117521798.1) and *nag5* (protein id: WP_117521944.1).

2.2.3 Primers used

All oligonucleotides were supplied from either Sigma or Integrated DNA Technologies (IDT). Primers were dissolved to a final concentration of 100 µM in dH_2O and stored at -20 °C.

Table 2.2.2: List of primers used in this work

Primer name	Primer sequence
T7seqF	TAA TAC GAC TCA CTA TAG GG
T7seqR	GCT AGT TAT TGC TCA GCG G
pBADseqF	CAA AAG TGT CTA TAA TCA CGG CAG
pBADseqR	ATC CGC CAA AAC AGC CAA GCT G
pET-21bvectorF	ATGTATATCTCCTTCTTAAAGTTAAACA
pET-21bvectorR	GCA CTC GAG CAC CAC C
pBADvectorF	GAA AAT TTA TAC TTC CAA GGT CAT CAT CAC
pBADvectorR	AGC AAA TCC ACC ACC CAT
pETYSBLvectorF	CGC GCC TTC TCC TCA CAT ATG
pETYSBLvectorR	TGC TGG TCC CTG GAA CAG
Nag3pBADF	TAA CAG GAG GAA TTA ACC ATG AGC AAA GAG AAG CTG A
Nag3pBADR	TTG GAA GTA TAA ATT TTC CTG CTT CAG ATC TTC G
Nag3pETYSBLF	GTT CTG TTC CAG GGA CCA GCA ATG AGC AAA GAG AAG CTG AAA G
Nag3pETYSBLR	CAT ATG TGA GGA GAA GGC GCG TTA CTG CTT CAG ATC TTC GCC TTC
Inverse mutagenesis primers Underlining denotes the codon of the mutated residue. Lower case indicates a substituted nucleotide to that of the original gene sequence.	
N3CAMF	AAGCCATGCAgcgGGTCATGATAC
N3CAR	TTGCCTTCGTTTTTGCTTTTG
N3CDMF	AAGCCATGCAgaTGGTCATGATAC
N3CDR	TTGCCTTCGTTTTTGCTTTTG
N3EAMF	CAGCCTGGTGcgGAAACCGGT
N3EAR	AAACAGCAGACGAACGGTGC
O3CAMF	AAGCCATGCAgcgGGTCATGATAC
O3CAR	TTACCCGGATTCTTGCTTTTG
O3EAMF	TCAGCCTGGCGcgGAAACCGG
O3EAR	AACAGCAGACGAACGGTG
N3NAMF	TAGCGTTGGTgcgGCAGTTGC
N3NAR	CCTGCGCTAATAACACCA
N3RAF	TTTTTTGATAACGATAGCGCAG
N3RAMR	cgcAAAGCTACCGGTCAGTTTGGC
N3QAMF	TACCTATAATgcgGGTGCAGGTC

N3QAR	AATTTGGCGCTACCACCA
N3QDMF	TACCTATAATgAtGGTGCAGGTC
N3QDR	AATTTGGCGCTACCACCA
N3HAMF	TAAAAGTGCAgcgGGTAGCAC
N3HAR	CCTTTGAAGCTCACCTCG
N3FNMF	AGGTAACTATaacTTCGAGGTG
N3FNR	GCCATAACCTGATTATAACC
N3FAMF	AGGTAACTATgcgTTCGAGGTG
N3FAR	GCCATAACCTGATTATAACC
N3FHMF	CATTCATATCcaTACCAATTTCGATC
N3FHR	GCAAAACCGCTATCAACAC
O3HFMF	CATTCATATTttcACCGCATTTCC
O3HFR	GCAAAACCGCTATCAACAC

2.2.4 Standard PCR protocol

Oligonucleotide primers were prepared as described in section 2.2.3. Template DNA consisted of either synthesised gene fragment or vector purified by a plasmid miniprep kit (Thermo Fisher) from appropriate strain. A 10 mM dNTP stock solution (Thermo Fisher) was used containing 2.5 mM each of dATP, dTTP, dCTP and dGTP. All cloning used Phusion high fidelity DNA polymerase (Thermo Fisher) and Phusion Green HF buffer in a 20 µl or 50 µl reaction volume. Each 20 µl reaction was prepared in a 0.5 ml PCR tube on ice and contained the following (volumes to be multiplied by 2.5 for 50 µl reactions): 4 µl 5x Phusion Green HF buffer, 0.4 µl 10 mM dNTPs, 1 µl 10 µM forward primer, 1 µl 10 µM reverse primer, 0.2 µl Phusion polymerase, 1 ng template DNA and add to 20 µl volume dH_2O. dH_2O was added first to the tube and polymerase last.

All PCRs were performed on either a Techne Techgene thermocycler or Bio-Rad T-100 thermocycler using standard PCR cycling parameters noted in the Phusion polymerase manufacturers guidelines and detailed below.

PCR stage	**Temperature**	**Time**	
Initial Denaturation	98 °C	5 min	
Denaturation	95 °C	10 s	X 30 cycles
Annealing	50 °C – 65 °C	30 s	
Extension	72 °C	30 s/kb	
Final extension	72 °C	10 in	

2.2.5 Inverse PCR site-directed mutagenesis protocol

Site-directed mutagenesis was used to generate proteins with substitution mutations at specific residues. A single mutagenic primer containing the desired substitution mutation was paired with a reverse primer complementary to the vector-gene sequence. The primers were divergent to each other to enable inverse PCR and yield a linear vector-gene sequence. Inverse PCR amplification was performed under standard conditions detailed above. Amplified PCR products were purified with a plasmid miniprep kit (Thermo Fisher) before use in a blunt end ligation with a combined T4 PNK kinase and T4 DNA ligation reaction using the following conditions, prepared on ice up to a 20 µl volume: 1 µl T4 PNK kinase, 1 µl T4 DNA ligase, 2 µl 10x T4 DNA ligase buffer with 10 mM ATP, 500 ng linear PCR product and dH_2O. Reaction were incubated overnight, 14 °C. Circularised plasmid was transformed

into *E. coli* DH5α competent cells and successful mutation of the gene verified by Eurofins Mix2Seq Sanger sequencing of plasmid extracted and purified from transformant colonies

2.2.6 Homologous recombination cloning

Homologous recombination was used to ligate vector and gene where necessary. Vector was amplified by inverse PCR to generate linearised vector. The gene of interest was either amplified by PCR or custom synthesised with added 20 bp sequences at the 5' and 3' ends that are homologous to the vector to enable recombination cloning. Recombination reactions used Vazyme ClonExpress II kit using the following conditions for a 20 µl reaction, prepared on ice: 4 µl 5x Exnase buffer, 50 ng gene insert, 200 ng vector, 2 µl Exnase II and up to 20 µl dH_2O. Reactions were incubated 37 °C, 30 min. Successful recombination was verified after transformation of recombination reaction products into competent *E. coli* DH5α and subsequent plasmid extraction and Mix2Seq Sanger sequencing.

2.2.7 Agarose gel electrophoresis

A 1 % agarose (w/v) was made by melting agarose (Melford, analytical grade) into 1 X TAE buffer (10.8 g Tris HCl, 1.14 ml acetic acid ACS grade, 0.5 M EDTA pH 8 per 1 L dH_2O). Following addition of 10000 X dilution of SYBR Safe DNA Gel Stain (Thermo Fisher), the molten agar was poured into a pre-cast gel electrophoresis frame with a comb placed to generate desired wells. Upon solidification after 30 min at room temperature the gel was placed into a DNA gel electrophoresis tank, submerged under 1 X TAE buffer and the comb removed. Samples used contained Phusion Green HF buffer which enabled direct loading of samples into wells without the ned of additional loading dye. Samples were electrophoresed at 100 V for 40 min and visualised under the transilluminator from a Syngene GeneGenius Gel BioImaging Unit with the Syngene GeneSnap software.

2.2.8 DNA extraction from agarose gel

Agarose gel was visualised under a low-intensity UV transilluminator and the desired DNA band excised with a clean scalpel and placed inside a clean 1.5 ml microcentrifuge tube. DNA gel extraction was performed using a Sigma GenElute Gel extraction kit following the manufacturers recommended protocol and resultant DNA eluted into 20 µl dH_2O. DNA was quantified using a Nanodrop (ND-1000) spectrophotometer and stored at -20 °C.

2.2.9 Plasmid extraction

A 3 ml culture of *E. coli* containing the desired plasmid was grown overnight (37 °C, 200 RPM) in LB supplemented with the appropriate antibiotic. Plasmid DNA was extracted using a GenElute HP Plasmid Miniprep kit following the manufacturer recommended protocol. Plasmid DNA was quantified using a Nanodrop (ND-1000) spectrophotometer and stored -20 °C.

2.2.10 PCR product purification

PCR reaction products were purified using a GenElute PCR Clean-Up kit to remove PCR impurities and contaminants. Purified DNA product was eluted into 20 µl dH_2O and quantified using a Nanodrop (ND-1000) spectrophotometer before storage at -20 °C.

2.2.11 Creation of chemically competent *E. coli* cells

A 5 ml overnight of the desired strain was grown overnight (37 °C, 200 rpm) in LB with the appropriate antibiotic selection. The following day, 200 µl of the culture was used to inoculate 10 ml of LB (containing appropriate antibiotic if necessary) and incubated aerobically (37 °C, 200 rpm) until the culture reached an OD_{600} nm 0.3 - 0.4. Cells were pelleted by centrifugation (3000 RCF, 5 min, 4°C), resuspended in 1 ml of 0.1 M ice cold sterile $CaCl_2$ and incubated on ice for 30 min. Cells were pelleted by centrifugation (3000 RCF, 5 min, 4°C) and resuspended in 0.5 ml of 0.1 M ice cold $CaCl_2$. Cell suspension was aliquoted into 50 µl volumes in clean 1.5 ml centrifuge tubes and either used immediately for plasmid transformation or snap frozen in liquid N_2 and stored at -80 °C.

2.2.12 Plasmid transformation of competent *E. coli* cells

A 50 µl aliquot of chemically competent *E. coli* was prepared fresh or thawed on ice from -80 °C storage. A 100 ng DNA was added to this aliquot and incubated on ice for 30 min. Cells were heat shocked at 42 °C for 30 s before addition of 950 µl pre-warmed (37 °C) LB and incubation at 37 °C for 1 h. A 200 µl aliquot of the cell suspension was spread onto a LB agar plate containing the appropriate antibiotic and incubated overnight (37 °C). A negative control with no added plasmid was included to validate transformation.

2.2.13 Sanger sequencing of DNA

Successful cloning and mutagenesis experiments were verified using the Mix2seq Sanger sequencing service, following the service providers stated guidelines. Sequenced DNA was checked with both forward and reverse strand sequencing results and aligned in Benchling.

2.3 Microbiological techniques

2.3.1 Culturing of *E. coli*

Escherichia coli strains were cultured aerobically on LB liquid and solid media with added antibiotic as appropriate. Cell growth was performed at 37 °C, unless otherwise stated, and liquid culture continuously shaken (200 rpm).

2.3.2 Culturing GPAC strains

Gram-positive anaerobic cocci (GPAC) were cultured anaerobically in a Don Whitely A35 anaerobic workstation filled with 5 % hydrogen, 95 % nitrogen gas and heated to 37 °C. All media and labware used in GPAC growth were equilibrated within the anaerobic environment at least 6 h prior to use. Strains were cultured in TSBT for all liquid growth and either Blood Agar Base No. 2 with Horse blood or fastidious agar plates for solid media growth. Liquid cultures were sealed under anaerobic conditions before removal from chamber for overnight growth. All cell growth was performed at 37 °C and liquid culture continuously shaken (200 rpm).

2.3.3 Optical density measurements

A 1 ml sample of cell suspension was pipetted into a 3 ml plastic cuvette and absorbance measured at 600 nm using the cuvette spectrophotometer function of a EPOCH2 microplate reader (BioTek). Samples were diluted with appropriate media where necessary.

2.3.4 Cell viability spot assay

A 1 µl volume of cell suspension was spotted onto a solid agar plate containing appropriate media and incubated 37 °C aerobically overnight for *E. coli* strains and anaerobically for 2 days with GPAC strains. Cell viability was determined from visible growth from the original spotting.

2.3.5 Growth and harvesting of recombinant protein expression cultures

A 20 ml culture of *E. coli* harbouring the desired expression plasmid was grown overnight (37 °C, 200 rpm) in LB media supplemented with ampicillin. The following day this culture was used to inoculate 1 L prewarmed (37 °C) LB containing ampicillin in a 2.5 L expression flask to a starting OD 0.05 and incubated (37 °C, 120 rpm). Upon reaching OD 0.4, 1 ml 0.5 M IPTG was added and the culture incubated overnight (20 °C, 120 rpm). After overnight expression, cells were pelleted by centrifugation

(10000 RCF, 20 min) and resuspended in 35 ml resuspension buffer (50 mM KPi pH 7.8, 200 mM NaCl, 40 mM imidazole and 20 % glycerol) and either used immediately for subsequent sonication and protein purification or stored at -80 °C.

2.4 Recombinant protein purification

2.4.1 Proteins used

Table 2.4.1: Protein sequences used in this work

Protein	Host species	Protein ID (NCBI accession)
N-AGA	*Corynebacterium* sp. Ax20	Q8GGD4.1
Oct1	*Anaerococcus octavius* DSM 11663	WP_115595702.1
Oct2	*Anaerococcus octavius* DSM 11663	WP_115594916.1
Oct3	*Anaerococcus octavius* DSM 11663	WP_115595530.1
Oct4	*Anaerococcus octavius* DSM 11663	WP_115595954.1
Nag3	*Anaerococcus nagyae*	WP_117521798.1
Nag5	*Anaerococcus nagyae*	WP_117521944.1

2.4.2 Sonication and isolation of cell lysate

Resuspended expression cells were allowed to thaw if frozen and supplemented with 1 mM AEBSF protease inhibitor. Cells were lysed on ice by sonication (3 min total, 3 s on, 7 s off, 50 % amp). Cell lysate was centrifuged (25000 RCF, 30 min, 4 °C) and the supernatant isolated for immediate protein purification.

2.4.3 Immobilised metal affinity chromatography (IMAC)

Expression culture supernatant was loaded onto a 5 ml Ni-NTA His-Trap column and nickel- affinity purified on an AKTA Start using a manufacturer's recommended protocol. The protein bound Ni-NTA column was washed with 30 ml resuspension buffer followed by protein elution fractionated into 10 X 3 ml aliquots of elution buffer (50 mM Kpi pH 7.8, 200 mM NaCl, 500 mM imidazole, 20 % glycerol). Fractions containing the highest protein purity were determined by continuous A_{280} nm detection throughout purification, with the highest absorbance fractions been pooled for further purification. Samples collected from the waste valve during flow and wash through IMAC purification steps were retained for later SDS PAGE analysis. All protein samples collected were stored at 4 °C until later purification.

2.4.4 Size exclusion chromatography (SEC)

Pooled fractions from IMAC purification were loaded onto a HiPep 16/60 sephacryl S-200 HR column equilibrated with 20 mM Tris HCl, 50 mM NaCl pH 7.4. Protein was purified on an AKTA Start using the manufacturers recommended protocol. Elution fraction containing the highest A_{280} nm values were retained and pooled for concentration.

2.4.5 On-column protein unfolding and refolding

Purified protein was loaded onto a 1 ml Ni Sepharose HisExcel EDTA resistant column and underwent protein unfolding and refolding on column on an AKTA Start using a manufacturer recommended protocol. Bound protein was unfolded with 15 ml unfolding buffer (50 mM KPi pH 7.8, 200 mM NaCl, 20 mM imidazole, 20 % glycerol, 2 M guanidine HCl, 25 mM EDTA). The unfolded bound protein was refolded on column with buffer gradient exchange into 30 ml refolding buffer (50 mM KPi pH 7.8, 200 mM NaCl, 20 mM imidazole, 20 % glycerol). Refolded protein was eluted with addition of elution buffer as described in IMAC chromatography. Highest purity fractions were pooled, concentrated and buffer exchanged into 20 mM Tris HCl pH 7.4, 50 mM NaCl.

2.4.6 Concentrating protein

Dilute recombinant protein solutions were concentrated using Vivaspin protein concentrator spin column (Cytiva). Spin column size and membrane molecular weight cut off chosen were appropriate volume of solution and protein size. Proteins were concentrated using the manufacturers recommended centrifugation concentrating protocol.

2.4.7 Protein quantification

Protein concentration was determined by Bradford assay using known standards of bovine-serum albumin as a calibration curve. Bradford assay was performed using the microplate format stated for the manufacturers recommended protocol and absorbance values measured on an EPOCH2 microplate reader (BioTek).

2.4.8 Purified recombinant protein storage

All generated protein stocks were stored at 4 °C for short term use. For long term storage protein stocks were buffer exchanged into 20 mM Tris HCl pH 7.4, 50 mM NaCl and 10 % glycerol, snap frozen with liquid N_2 and stored at -80 °C.

2.5 *In vitro* protein biophysical characterisation

2.5.1 Sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS PAGE)

Polyacrylamide gels were made using the following protocol. Each resolving polyacrylamide gel was cast containing 12 % w/v acrylamide, 0.32 % w/v bisacrylamide, 370 mM Tris HCl pH 8.8, 0.1 % w/v SDS, 0.1 % w/v ammonium persulphate (APS), 0.01 % w/v tetramethylethylenediamine (TEMED) and up to required volume with dH_2O. Gel polymerisation occurred upon addition of APS and TEMED last. Upon solidification of resolving gel, the stacking gel was cast containing 4 % w/v acrylamide, 0.11 % w/v bisacrylamide, 120 mM Tris HCl pH 6.8, 0.1 % w/v SDS, 0.1 % w/v APS, 0.01 % w/v TEMED and up to necessary volume with dH_2O. A comb was placed in the stacking gel to create appropriately sized wells and stacking gel left to solidify. Running buffer was composed of 3 g/L Tris, 14 g/L glycine and 1 g/L SDS, pH 8.8.

A 4 X loading dye was prepared containing 10 % glycerol, 8 % SDS, 70 mM Tris pH 7.2, 0.4 % bromophenol blue and 3 % β-mercaptoethanol. A 5 µl aliquot of loading dye was added to 15 µl protein samples and heated (90 °C, 10 min). 15 µl of each sample was pipetted into separate wells with one well containing 3 µl of PageRuler prestained protein ladder (Thermo). SDS PAGE was conducted (200 V, 50 min).

2.5.2 Coomassie blue staining

Following SDS PAGE, the gel was placed in a container and submerged in Coomassie brilliant blue dye which contains 2.5 g Remazol Brilliant Blue R dissolved in 450 ml methanol, 100 ml acetic acid and 450 ml dH_2O. Gels were left to gently shake on an orbital shaker for 30 min before removal of the dye and submersion in destaining solution (10 % ethanol, 10 % acetone) for 6 h. Destaining solution was removed and the gel visualised under white light.

2.5.3 Size exclusion chromatography Multi angle laser light scattering (SEC-MALLS)

All SEC-MALLS experiments were performed with assistance from Dr Andrew Leech in the York Technology facility. Purified protein was loaded onto a Superdex 200 HR10/30 column equilibrated with 20 mM Tris HCl pH 7.4, 50 mM NaCl. Protein was eluted with continuous measurements of eluent UV A_{280} nm, light scattering of laser light passed through the solution and refractive index. Protein was run at two concentrations to check for concentration dependent effects. A control protein sample of 2 mg/ml BSA was performed to check instrument accuracy.

2.5.4 Electrospray ionisation mass spectrometry (ESI-MS)

ESI-MS was performed by Dr Adam Dowle in the York Technology Facility. Purified protein was buffer exchanged into 20 mM ammonium acetate and diluted to 10 µM with aqueous 50 % acetonitrile containing 1 % formic acid. The sample was then infused into a maXis qTOF mass spectrometer using electrospray ionisation. Positive ESI-MS spectra were averaged over 1 min before maximum entropy deconvolution to average masses. Detected intact and truncated protein was compared against the primary sequence of the respective protein.

2.5.5 Peptide fingerprint mass spectrometry

All peptide fingerprint mass spectrometry used either MALDI-MS/MS or LC-MS/MS and was performed by Dr Adam Dowle or Dr Chris Taylor in the York technology Facility. Both techniques used protein gel bands excised after SDS PAGE and Coomassie staining. For LC-MS/MS, the protein underwent in-gel tryptic digest following addition of dithioerythritol (DTE) and iodoacetamide. Extracted peptides were acquired over a 10 min LC-MS/MS acquisition with elution from a 50 cm micropillar column onto a Bruker maXis qTOF mass spectrometer, operated in DDA mode. MALDI-MS/MS protein bands underwent tryptic digest following DTE and iodacetamide addition. Extracted peptide were acquired on a MALDI-MS/MS acquisition on a Bruker ultraflexTreme. All tandem mass spectra obtained from both techniques were searched against an in-house database appended with expected primary sequence of respective proteins.

2.5.6 Circular dichroism (CD)

Purified protein was diluted to 0.2 mg/ml into 10 mM NaH_2PO_4 pH 7.4 and transferred to a 1 mm path length quartz cuvette and a CD spectrum obtained using a J-1500 CD spectrophotometer while under nitrogen purging. Data was collected as a measure of ellipticity (mdeg) over a 185 to 260 nm wavelength range at a constant 25°C temperature. Data for wavelengths displaying a high tension (HT) voltage < 600 V was omitted from spectra analysis. Protein samples were normalised against a 10 mM NaH_2PO_4 buffer used in the same cuvette and orientation. CD spectra data were analysed to give an estimation of secondary structure content using the online service BeStSel (Beta Structure Selection) (Micsonai et al., 2018).

2.6 Structural characterisation techniques

2.6.1 Protein crystallisation

SEC purified protein was used at two concentrations (20 mg/ml and 10 mg/ml unless otherwise stated) in protein crystallisation trials using several commercially available crystal screen trays in a 96 well-plate sitting drop format. Crystal trays used were JCSG+ (Molecular Dimensions), Crystal screen HT (Hampton Research), Index (Hampton Research) and PEG/ION (Hampton Research). Protein was aliquoted into each condition present on crystallisation plates using a Mosquito LCP (SPT Labtech) before the plates were sealed and stored in a controlled climate room (20 °C). Crystal formation was checked periodically with a light microscope and, if necessary, crystal optimisation was performed for viable crystallisation conditions. Optimisation used 24 well-plates using the hanging-drop vapour diffusion method with variations in reservoir conditions and protein concentration. Plates were sealed and stored in a climate controlled room (20 °C).

Viable crystals from 96 and 24 well-plates were harvested with nylon loops of various size. Proteins were resuspended with cryoprotectant solution containing mother liquor and 20 % glycerol and cryocooled in liquid N_2.

2.6.2 X-ray diffraction and structural determination

Harvested crystals were stored in liquid N_2 and sent to Diamond Light Source, UK for X-ray diffraction data collection on either the i03 or i04 beamlines. Viable protein diffraction data collected was indexed and scaled using Xia2, followed by data reduction with Aimless in CCP4i2 (Potterton et al., 2018). Molecular replacement performed with MOLREP was used to obtain initial phase information with the AlphaFold model of Nag3 being used as the candidate for Nag3E138A structural determination and that structural model being later used for Oct3 molecular replacement. Structural solutions were refined with iterative rounds of REFMAC5 and model building using COOT with assistance from Reyme Herman and Eleanor Dodson.

2.6.3 X-ray fluorescence (XRF) spectroscopy

Diffracting protein crystals of Nag3E138A were used for XRF data collection on the i24 beamline up to 18500 eV for 1 s exposure. Data was collected and auto processed by ISPyB to label elements responsible for significant fluorescence peaks in the spectra.

2.7 *In vitro* protein biochemical characterisation

2.7.1 Glutamine detection kit assay

Each reaction (200 µl) contained stated concentrations of VFA precursors, HEPES buffer (25 mM HEPES, pH 7.2) and purified recombinant enzyme as stated. Reactions were incubated 37 °C for 1 h before two 10 µl samples were taken from each reaction mixture for use in the Megazyme L-glutamine/ammonia (Rapid) glutamine detection assay to quantify enzyme activity. The assay is based on quantitative endpoint reading and is performed using the standard protocol stated for a microplate reader format and utilises the following two reactions.

1) L-glutamine + H_2O $\xrightarrow{\text{(glutaminase)}}$ L-glutamate + NH_4^+
2) NH_4^+ + 2-oxoglutarate + NADPH $\xrightarrow{\text{(GDH)}}$ L-glutamate + $NADP^+$ + H_2O

$NADP^+$ formation is stoichiometric with the amount of glutamine and ammonia, enabling quantification of enzyme activity by measuring depletion of NADPH can be measured as a decrease in A_{340} nm. Effect of ammonia upon A_{340} nm was determined by one of the two samples assayed omitting step 1 of the reaction, thereby only measuring the effect of ammonia not ammonia and glutamine. This change in absorbance is subtracted from the second measurements which determines the Abs change of ammonia and glutamine. The concentration of glutamine can then be calculated as follows.

$$c = \frac{V \times MW}{\varepsilon \times d \times v} \times \Delta A$$

C = L-glutamine concentration (g/L)

$\Delta A = \Delta A_{(\text{glutamine + ammonia})} - \Delta A_{(\text{ammonia})}$

V	=	final volume [mL]
MW	=	molecular weight of analyte [g/mol]
ε	=	extinction coefficient of NADPH at 340 nm
	=	6300 [l x mol^{-1} x cm^{-1}]
d	=	light path [cm]
v	=	sample volume [mL]

2.7.2 Ninhydrin reaction screening assay

Each reaction (200 µl) contained purified recombinant enzyme as noted, HEPES buffer (25 mM HEPES, 8.5 mM NaCl pH 7.2) and indicated acylated substrate at concentrations detailed. Reactions were incubated (static) at 37 °C for the noted period of time for the respective assay. Enzyme activity was quantified by measurement of the formation of the free amino acid from its acylated substrate by a ninhydrin colourimetric assay, with ninhydrin reacting with the primary α-amino groups released enzymatic activity to form Ruhemann's purple which was detected on an EPOCH2 microplate reader (BioTek) at A_{570} nm. A 20 µl aliquot from each sample was taken and mixed with 20 µl ice cold 25 mM

EDTA to quench the reaction. A 50 µl volume of 2 % ninhydrin reagent solution (2 % ninhydrin and 0.5 % hydrindantin, lithium acetate buffer pH 5.2) was added to each quenched reaction sample and incubated (70 °C, 10 min) in a Techne Techgene thermocycler. Samples were cooled to 4°C before the addition of 100 µl dH2O and transferred to a microplate for A_{570} nm measurement on an EPOCH2 microplate reader. Samples were diluted with dH2O where necessary. The concentration of free amino acid product formed was determined from calibration curves created from amino acid standards for each respective amino acid.

2.7.3 Ninhydrin reaction kinetic assay

Kinetic experiments used the same protocol as stated for ninhydrin reaction screening assay with a stated purified enzyme and a range of six substrate concentrations in time course reactions. Four time points (t = 0, 5, 10 and 15 min unless otherwise stated) were taken in each experiment with quenching into 12.5 mM (final concentration) ice cold EDTA and initial rates of glutamine formation were determined from the linear slope at the start of reactions. Michaelis-Menten kinetic values were obtained using nonlinear regression analysis in GraphPad Prism9 after plotting enzyme initial velocity against substrate concentration.

2.7.4 Metal reconstitution assay

Reactions were performed as stated in ninhydrin reactions screening assay with enzyme incubation reactions amended with metal at stated concentrations.

2.8 Mass spectrometry detection of VFAs

2.8.1 *In vivo* whole cell VFA biotransformation

Six colonies of each desired strains were picked from solid media plates and grown in separate 20 ml liquid media 37 °C, 200 rpm for 1 day aerobically (*E. coli*) or 2 days anaerobically (GPAC). Cultures were pelleted by centrifugation (2500 RCF, 5 min) and every two pellets combined and resuspended in incubation buffer (3 g KH_2PO_4, 1.9 g K_2HPO_4, 0.2 g $MgSO_4$. $7H_2O$, 1.4 g NaCl, 1 g NH_4Cl, 0.2 g yeast extract per 1 L dH_2O) to yield three resuspended cell suspensions per strain. Each cell suspension was added to three VFA biotransformation reactions (final OD 2.6 unless otherwise stated) containing 2 mM 3M2H-Gln, HMHA-Gln or no substrate up to a 2 ml reaction volume of incubation buffer. Biotransformation reactions were incubated 37 °C, 200 rpm for 40 h either aerobically (*E. coli*) or anaerobically (GPAC). After incubation, released VFAs were isolated by solvent extraction. 0.9 ml sample was added to 1.8 ml ethyl acetate before acidification with 50 µl 20 % HCl and incubation for 1 h, static, room temperature. Remaining biotransformation reactions were retained for cell viability spot assays. After 1 h, samples were vortexed and centrifuged (2000 RCF, 2 min) before pipetting off 0.8 ml of top organic layer into glass GC vials. Extracted samples were stored at 4 °C before mass spectrometry analysis.

2.8.2 *In vitro* VFA biotransformation

Enzyme substrate incubations were prepared in triplicate in 2 ml reaction volumes containing: 2 mM VFA precursor, HEPES buffer and 1.65 µM purified protein. Reactions were initiated upon protein addition and incubated (1 h, 37 °C, static). After incubation, released VFAs were isolated by solvent extraction in the same manner as detailed for *in vivo* whole cell VFA biotransformation.

2.8.3 Unilever LC-MS VFA detection

Detection and quantification of solvent extracted VFA products was performed at the Unilever facilities within Material Innovation Factory, Liverpool. Experimental work was performed by Viv Perchard and Emily Nicholson using a Unilever proprietary LC-MS system. Samples underwent hydrophilic interaction chromatography and mass spectrometry detection on an ACQUITY ARC-QDA using electrospray ionisation. Positive ESI spectra were collected and personally analysed in Skylines, with released VFA quantification achieved using calibration curves of 3M2H-Gln and HMHA-Gln standards.

2.8.4 York GC-MS VFA detection

Detection of solvent extracted VFA products from *E. coli* KEIO transport mutant experiments was performed in the Metabolomics and Proteomics Lab in the York Biology Technology Facility. Experimental work and subsequent data analysis was performed by Swen Langer using a developed GC-MS system. A 2 µl sample was injected onto a Zebron ZB-5HT Inferno column before select ion mode detection on an Agilent 5975C MSD for 3M2H and HMHA. Data was processed as peak areas from intensity of CFA detection in each sample.

2.9 Bioinformatic techniques

2.9.1 Multiple sequence alignment (MSA)

Protein sequences with homology to a query sequence were searched for with BlastP and collected. A MSA of these sequences was created using the online resource Clustal Omega (Madeira et al., 2022; Sievers et al., 2011). The created alignment was visualised and annotated with Jalview (Waterhouse et al., 2009).

2.9.2 Maximum likelihood Phylogenetic Tree

Protein MSAs generated by Clustal Omega were uploaded to the online resource IQTree to produce a maximum likelihood tree phylogenetic tree using the best fit substitution model according to Bayesian information criterion (in all instances this was model: LG+F+I+G4). The resultant phylogeny tree was visualised, annotated and pruned using the online website iTOL (Letunic and Bork, 2016).

Chapter 3 – Identification of 3-methyl-2-hexenoic acid (3M2H) and 3-hydroxy-3-methylhexanoic acid (HMHA) generating aminoacylases within *Anaerococcus*

3 Identification of 3-methyl-2-hexenoic acid (3M2H) and 3-hydroxy-3-methylhexanoic acid (HMHA) generating aminoacylases within *Anaerococcus*

3.1 Introduction

Volatile fatty acids (VFAs) have long been known to be involved in body odour (BO) (Zeng et al., 1991), with further work identifying two medium chain length C_7 VFAs, 3-methyl-2-hexanoic acid (3M2H) and the structurally similar 3-hydroxy-3-methylhexanoic acid (HMHA), as being highly abundant in axilla sweat and major contributors to human BO (Natsch et al., 2003, 2006). Shelley *et al.'*s seminal work in 1953 showed microbial action is required for BO production and this theory extends to VFAs with the identification of a corynebacterial strain, *Corynebacterium sp.* Ax20, capable of generating both 3M2H and HMHA from their respective odourless glutamine-conjugated precursor molecules present in apocrine sweat (Natsch et al., 2003). Natsch and colleagues also isolated and characterised *in vitro,* the aminoacylase enzyme responsible for release of both VFAs, identifying it as N^{α}-acyl-glutamine aminoacylase (N-AGA). This coupled with axilla malodour studies since the 1980s that correlated axilla malodour intensity to corynebacterial numbers (Jackman and Noble, 1983; Leyden et al., 1981; Taylor et al., 2003), suggested that species of *Corynebacterium* were major contributors to BO and VFA generation. Despite this initial evidence for corynebacterial dominance in VFA generation, all other *Corynebacterium* species tested have displayed either undetectable or very low aminoacylase activity (James et al., 2013b) making the high VFA generating activity of *C. sp Ax20* appear to be unique to that species rather than a norm for the genus. The activity of *C. sp Ax20* alone is also not sufficient to match the levels of 3M2H and HMHA present in axilla sweat (James et al., 2013b) indicating that other microorganisms must also contribute to VFA generation.

The advancement of culture-independent studies of the axilla microbiome revealed previously unidentified prominent taxa in the underarm with a significant finding being the prevalence of Gram-positive anaerobic cocci (GPAC), of which the *Anaerococcus* genus is a part (Egert et al., 2011; Troccaz et al., 2015). *Anaerococcus* was identified as the fourth most abundant genera in the underarm and correlated to a "fatty" malodour, hinting at a role in human BO (Troccaz et al., 2015). In 2014, a species of *Anaerococcus,* later classified as *A. nagyae* (Veloo et al., 2016)*,* was observed to generate HMHA from incubation of bacterial cells with HMHA-Gln providing clear evidence that medium chain VFA production is not limited to strains of Corynebacteria (Fujii et al., 2014). Preliminary internal work within Unilever had also observed *Anaerococcus octavius* as able to produce VFAs at significant enough levels to contribute to VFA malodour. The presence of two active VFA generating strains within

the *Anaerococci* implicates that this genus is a potentially major contributor to human body odour and warranting further study of the biotransformation process that generates these VFAs.

Initial work focused on characterisation of the enzymatic process which generates VFAs within *A. octavius*, to identify the enzyme(s) responsible for VFA generation from odourless precursors. Candidate enzymes from *A. octavius* were identified using a bioinformatics approach, recombinantly expressed and purified to allow screening of activity with VFA precursors. This work was then followed with the identification of orthologous enzymes within *A. nagyae* and subsequent purification and characterisation.

3.2 Bioinformatic analysis of *Anaerococcus* aminoacylase proteins

With the identification of an active VFA generating enzyme, N-AGA, in *C. sp* Ax20, a search within the Anaerococcus genus for putative aminoacylases which share sequence identity with the active corynebacterial enzyme was performed. The identification of a homologue of the *C. sp* Ax20 N-AGA within the Anaerococcus genus, and specifically in *A. octavius,* would make it the most likely candidate for being the principle VFA liberating enzyme in those species. The primary sequence of N-AGA was used as the query sequence for a BlastP search limited to the Anaerococcus genus and using a single representative genome for each species. This revealed putative proteins throughout the *Anaerococcus* genus that displayed some similarity to N-AGA, although this similarity was limited to lower than 35 % identity in all cases preventing any clear candidates for VFA production being easily identified. All the sequences retrieved from this bioinformatic search (including N-AGA) were processed in a multiple sequence alignment and used to generate a maximum likelihood phylogeny tree to study the distribution of these sequences within the genus (Fig 3.2.1).

This revealed that most Anaerococcus species contain 3-5 putative aminoacylases with four being present within *A. octavius,* and each of these sequences forming their own distinct clades. The distribution of *A. octavius* gene sequences in the phylogenetic tree suggests evolution from a single common ancestor with one gene duplicating multiple times to give rise to four genes. These proteins will be hence distinguished by the arbitrary nomenclature of Oct1 – 4. None of the *A. octavius* sequences cluster close to that of N-AGA suggesting that VFA activity in Anaerococcus does not require a high identity homologue of N-AGA. A close homologue of each of the four *A. octavius* aminoacylase proteins is observed in *A. nagyae*, a species that has been previously shown to release HMHA from its glutamine conjugate (Fujii et al., 2014), and this species also contains an additional 5th protein that has no similar representative in *A. octavius*. These findings in *A. nagyae* and their implication for VFA generation in that species are explored further later in this chapter.

The protein sequences from *A. octavius* (Oct1 – 4) and *A. nagyae* (similar naming convention Nag 1-5) and N-AGA were used in a multiple sequence alignment to study the levels of conservation and similarity both to inform protein characterisation and ensure the BLAST search had correctly identified genuine candidates (Fig 3.2.2). The alignment displays areas of similarity throughout the proteins and noticeably highlights key residues known to be conserved in the M20D aminoacylase family of which N-AGA has been assigned (Natsch et al., 2003). Based on this it seems likely that the identified protein sequences are genuine aminoacylase proteins and if active with the VFA precursors will likely catalyse via a similar mechanism of action.

This bioinformatics search approach has identified four putative candidate proteins in *A. octavius* which will be studied further *in vitro* to assess if any are able to generate VFAs.

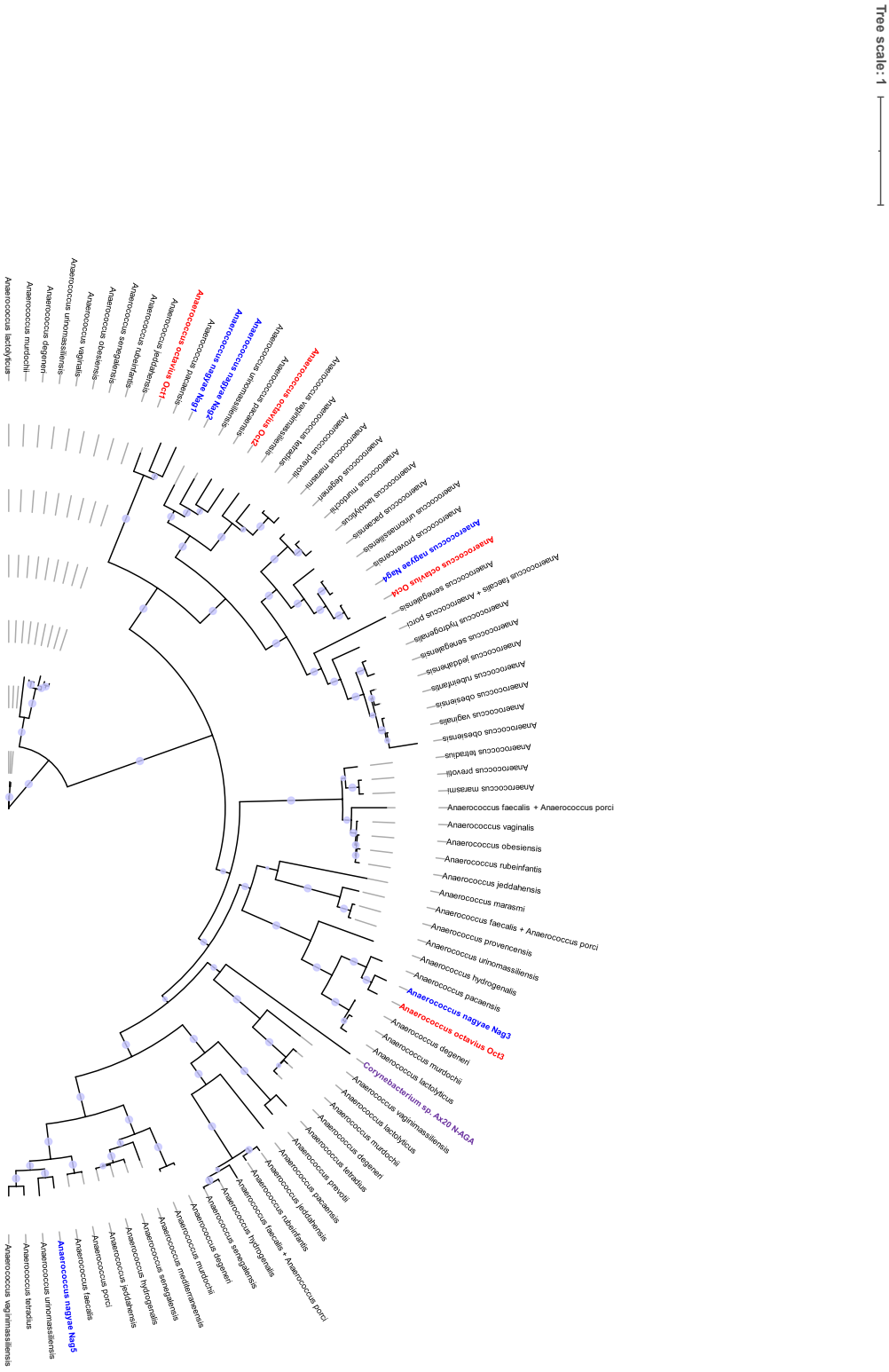


Figure 3.2.1: Phylogeny of putative aminoacylases within Anaerococcus

Protein sequences obtained from protein BLAST within Anaerococcus using N-AGA primary sequence as the query. Multiple sequence alignment generated in Clustal Omega and used in IQ-TREE for maximum likelihood phylogeny tree. Circles denote node support from 1000 ultrafast bootstrap values from IQ-TREE (>90). The 4 aminoacylase protein sequences present in *A. octavius* are indicated in red. The 5 aminoacylase proteins from *A. nagyae* are shown in blue. Purple denotes N$^{\alpha}$-acylglutamine aminoacylase from *C. sp.* Ax20.

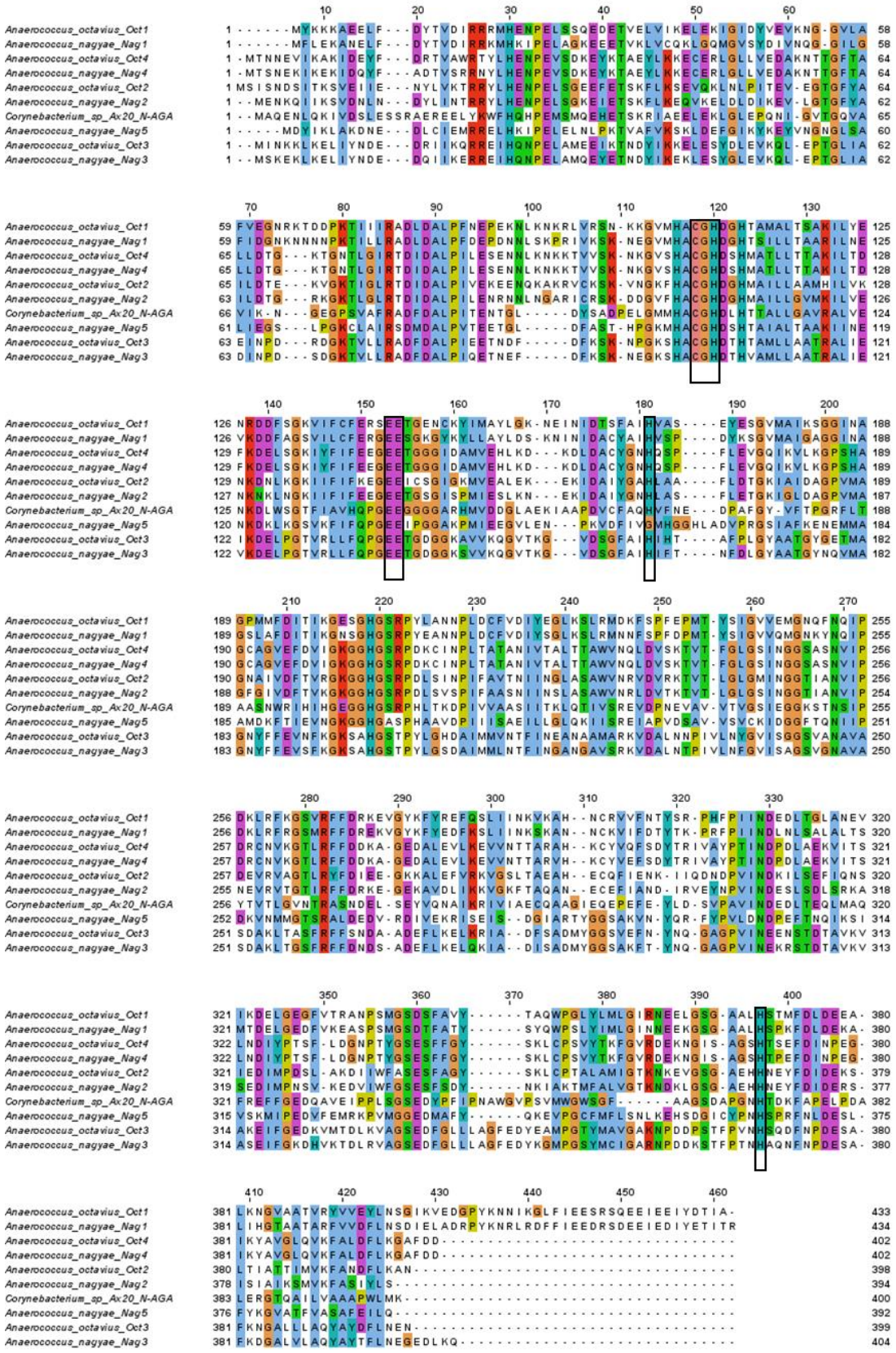


Figure 3.2.2: Multiple sequence alignment of N-AGA with *A. octavius* and *A. nagyae* aminoacylases

Alignment performed using Clustal Omega of N-AGA, Oct 1-4 and Nag 1-5. Using Clustal colouring to denote conservation. Black boxes highlight conserved metal binding and catalytic activity residues in members of the M20D aminoacylase family.

3.3 Purification of putative *A. octavius* aminoacylases

To ascertain if any of the four putative aminoacylases from *A. octavius* identified in the bioinformatics search can generate VFA malodour, the genes encoding them (Oct1 – 4) were expressed and the resulting proteins purified to enable *in vitro* characterisation. The *A. octavius* genes were each cloned into an IPTG inducible proprietary expression vector, derived from a pET vector background, by the external company GeneMill (Appendix A). This vector contains a C-terminal hexa-histidine tag for an expressed protein whose gene is cloned into the cloning site. This vector contains the pBR332 origin of replication and *rop* gene, as well as *lacI* repressor gene and T7 polymerase promoter and terminator up and downstream of the ORF which are characteristic of pET vectors. The *Corynebacterium* gene encoding N-AGA (Genbank ID Q8GGD4.1) was also cloned into the same vector to act as a positive control in later experiments.

All proteins were expressed in the cytoplasm of *E. coli* BL21 (DE3) and purified as described in Materials and Methods. The following description details the expression and purification of a single protein, Oct3, with all proteins being handled in the exact same manner. A 1 L culture of *E. coli* BL21 (DE3) containing the Oct3 expression plasmid was harvested 20 h post IPTG induction. The collected cells were lysed by sonication and the supernatant containing soluble Oct3 protein loaded onto a HisTrap HP 5ml column packed with nickel Sepharose medium for IMAC purification. The loaded column was washed with 5 column volumes (CV) of IMAC wash buffer containing 40 mM imidazole to remove non-specifically bound protein before elution of bound Oct3 by changing to 100% IMAC elution buffer containing 500 mM imidazole, outcompeting the hexa-His tag for nickel affinity (Fig 3.3.1 A). The protein presence during the purification was continuously monitored by absorbance readings at 280 nm.

The IMAC A_{280} trace shows protein eluting soon after switching to IMAC elution buffer with a high peak occurring between 83 – 89 ml solution volume during the purification. The purity of the eluted protein was checked on a 12% SDS PAGE gel, which contains a clear large band migrating at just over 40 kDa in both Oct3 expression cell lysate and IMAC elution fractions (Fig 3.3.1 C) which is close to the predicted molecular weight of an Oct3 monomer at 44.8 kDa. The absence of a similar large band in both the IMAC flow and wash through samples validates that the purification conditions were ideal with minimal protein loss throughout the run due to failure in Oct3 specific binding to the column and excessive imidazole in the IMAC wash buffer out-competing Oct3, respectively.

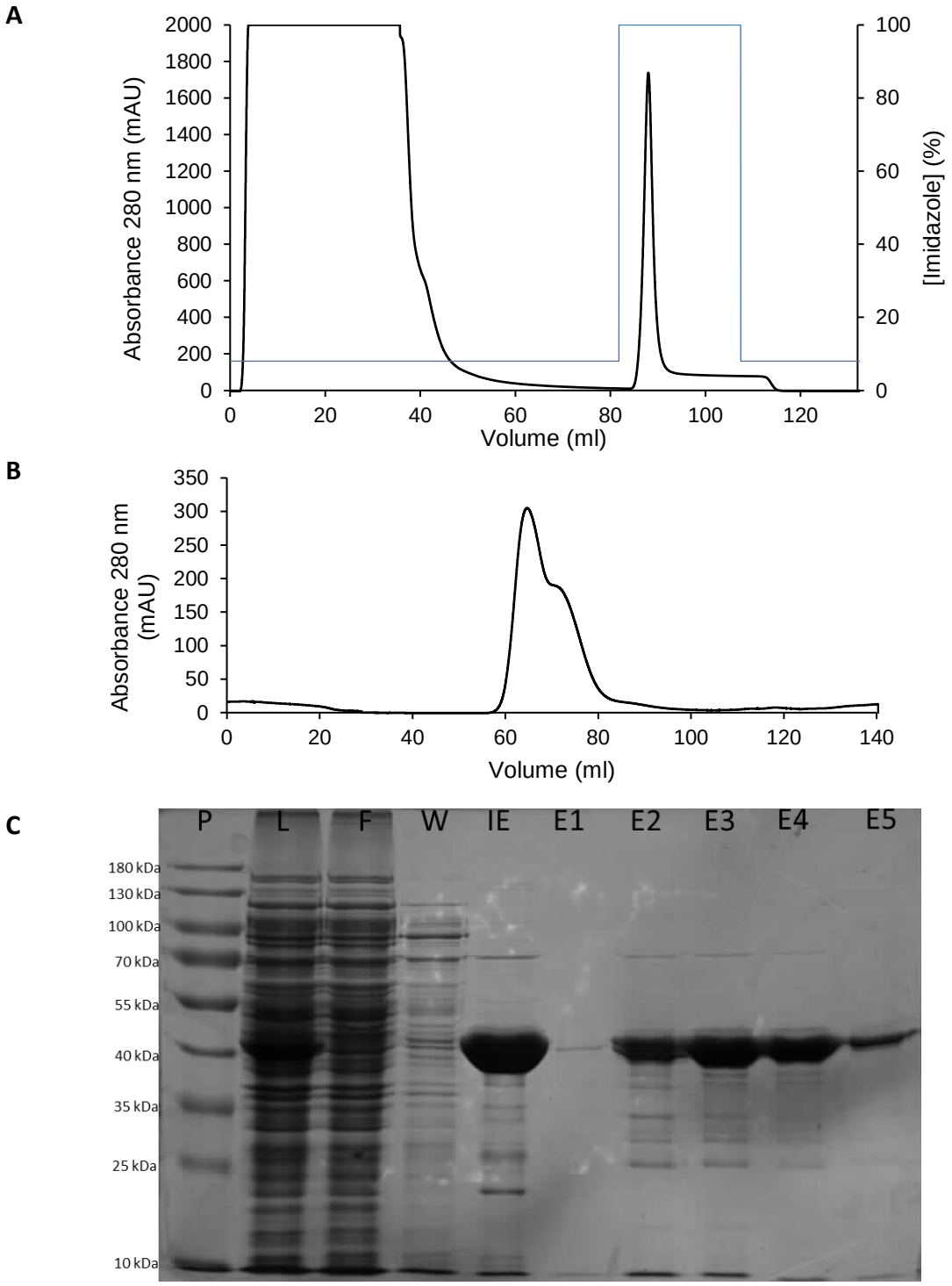


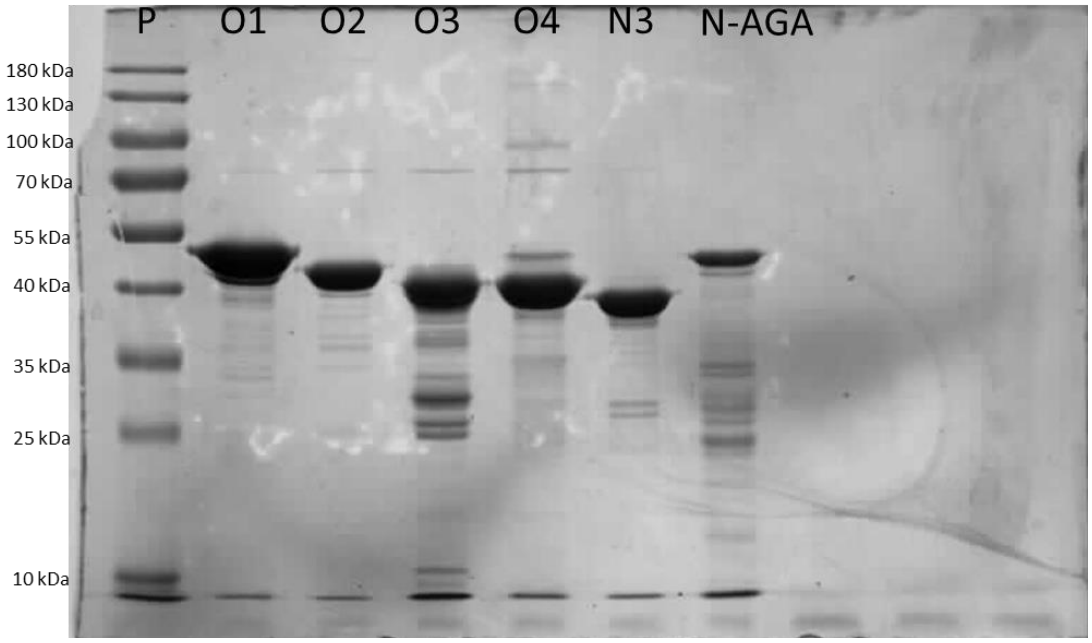
Figure 3.3.1: Purification of *Anaerococcus octavius* aminoacylase, Oct3

A. Immobile metal affinity chromatography (IMAC) trace of the wash and elution of Oct3 from cell lysate of *E. coli* BL21 (DE3) recombinantly expressing Oct3 from the proprietary expression vector pGM111_3. The black line details the A_{280} and the blue line represents the % concentration of imidazole present in the solution (percentage denotes 0 – 500 mM range of imidazole).
B. Size exclusion chromatography (SEC) of Oct3. IMAC purified elution fractions containing the highest purity of Oct3 were loaded onto a HiPrep 16/60 Sephacryl S-200 column.
C. 12 % SDS PAGE gel of Oct3 purification. Gel lanes contain (from left to right): (P) Page Ruler prestained protein ladder, (L) *E. coli* BL21 (DE3) expressing Oct3 lysate, (F) IMAC flow through, (W) IMAC wash through, (IE) IMAC pooled elution fractions of Oct3, (E1-5) Oct3 SEC fractions 9 -13.

An additional protein chromatography step was also used to try and improve the purity of the IMAC purified proteins further. The highest purity fractions from IMAC were used for size exclusion chromatography (SEC) on a HiPrep 16/60 Sephacryl S-200 HR column. Using Oct3 as a representative protein purification, the 6 ml of IMAC elution fractions that contained the highest purity of Oct3 were pooled and concentrated with Vivaspin centrifugal concentrators to approximately 4 ml volume. The Oct3 sample was then loaded onto the S-200 HR column with the protein being constantly monitored during the purification by A_{280} measurements. Oct3 eluted across the 60 – 80 ml volume forming an elongated peak that appears to be created by two peaks with an initial peak and a secondary smaller peak that overlaps with the first suggesting they can't be resolved (Fig 3.3.1 B). Examination of the fractions containing these peaks by SDS-PAGE (Fig 3.3.1 C) showed that all contain intact Oct3 indicating that the second peak is not a truncated product of Oct3 and instead likely a different oligomerisation state to that of the initial peak. All but the first SEC elution fraction, given its low amount of Oct3 relative to the other fractions, were pooled and concentrated with Vivaspin centrifugal concentrators from an initial volume of 16 ml to approximately 2 ml. Due to the absence of tryptophan residues in Oct3 and other purified proteins preventing accurate measurement at A_{280}, the protein concentration was determined with Bradford reagent and the comparison of the A_{590} of the protein sample with a calibration curve of known concentrations of bovine serum albumin (BSA).

The other *A. octavius* proteins and N-AGA were purified in the exact same manner of IMAC followed by SEC and each yielded soluble protein of high purity at the approximate molecular weight when examined by SDS-PAGE. (Fig 3.3.2 A)**.** All Oct proteins and N-AGA eluted after similar lengths of time during SEC suggesting that each likely share the same shape and oligomeric state given all have roughly similar molecular weights of between 40 kDa – 50 kDa (Fig 3.3.2 A). Each protein also displayed the same phenomena of twin peaks during elution (Fig 3.3.2 B), with some displaying better resolution into separate peaks such as Oct2. Like Oct3, SDS-PAGE analysis of elution fractions across these resolved peaks showed no clear differences in protein banding, again suggesting a difference in oligomeric state for the two peaks (data not shown).

A



B

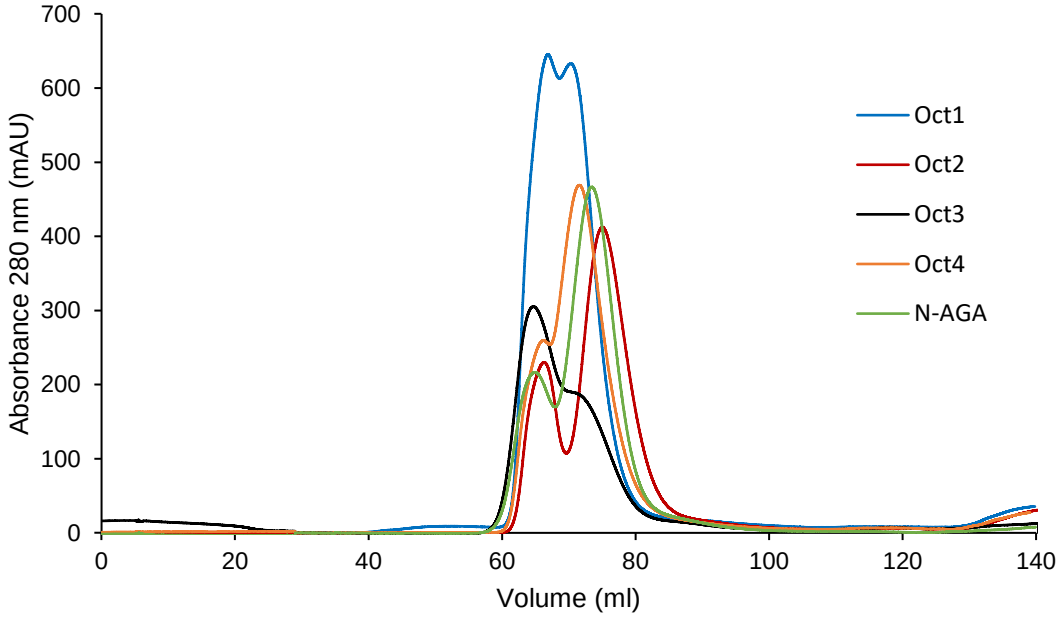


Figure 3.3.2: *A. octavius* and N-AGA protein purifications

A. SDS-PAGE gel (12 %) image of aminoacylase proteins after IMAC and SEC purification. From left to rights lanes contain: (P) Page Ruler prestained protein ladder, (O1) Oct1, (O2) Oct2, (O3) Oct3, (O4) Oct4, (N3) Nag3 (discussed later in the chapter), N-AGA.
B. SEC traces of Oct 1- 4 and N-AGA overlayed.

3.4 Verification of the identity of purified protein by mass spectrometry

An additional quality control check was performed on all purified proteins to confirm for each that the correct protein has been purified. Due to the availability of mass spectrometry (MS) equipment at the time of MS protein analysis, the proteins were tested using either intact protein electrospray ionisation mass spectrometry (ESI-MS), matrix assisted laser desorption/ionisation tandem mass spectrometry (MALDI-MS/MS) or liquid chromatography tandem mass spectrometry (LC-MS/MS). All mass spectrometry work was undertaken by the Centre of Excellence of Mass Spectrometry at York. Oct3 was confirmed with ESI-MS to contain the correct sequence as that given with two protein peaks being detected at both 44840 Da (molecular weight of Oct3 with hexa His tag) and 44872 Da (Oct3 plus 2 oxidations) (Fig 3.4.1). Several other peaks were also detected, most likely being truncated products of Oct3 indicating some protein degradation during storage of the sample. In the case of the highest intensity peak this can be matched to a truncated product of Oct3 with the mass of 25725 Da. This N-terminal truncation creates a protein excluding all but the last 14 residues of the N-terminal portion of the catalytic domain, leaving the entire oligomerisation domain and the smaller C- terminal section of the catalytic domain and suggests that the N-terminus is more vulnerable to proteolysis and degradation. Analysis of Oct3 sequence on Expasy peptide cutter shows pepsin does have a matching cleavage site to create this exact truncated product, although the action of other endopeptidases and aminopeptidases remains an alternate possibility. While this truncated protein does display much higher intensity than that of full length Oct3 this does not directly relate to the relative abundance of these species within the sample, instead correlating to each protein's response to the mass spectrometer. This is evident from SDS PAGE gels containing Oct3 displaying a much larger band for full length protein then any potential degradation product (Figs 3.3.1 C and 3.3.2 A).

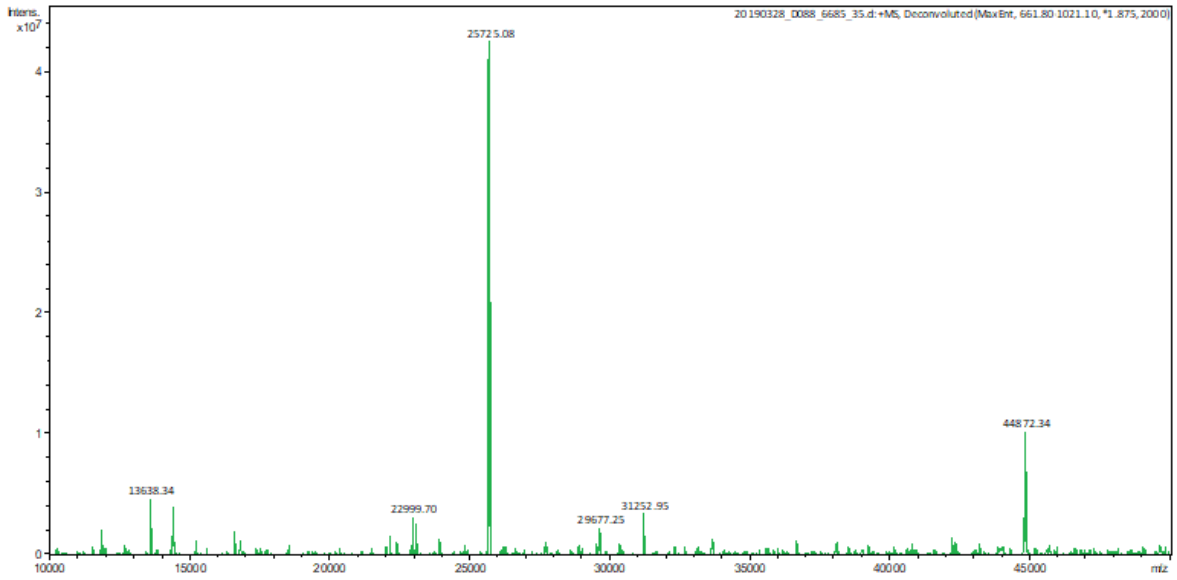


Figure 3.4.1: ESI-MS spectrum of Oct3

Deconvoluted positive ESI-Ms spectra of Oct3. Intensity peaks shown correspond to most intense protein peaks detected. Actual molecular weight of Oct3 is 44840 Da.

Oct1 and Oct2 were verified to have the correct protein sequence via LC-MS/MS. The protein was run on an SDS PAGE gel and the band at the correct molecular weight for an intact monomer was excised for use in a tryptic digest. The resulting protein fragments from the digestion were extracted over a 10 min LC-MS/MS acquisition with elution from a 50 cm microPillar column onto a Bruker maXis qTOF mass spectrometer for peptide mass fingerprinting. The detected fragments were compared to a database amended with the given sequences of Oct1 and Oct2. Both displayed strong matches to the given sequences with detected fragments matching to 48 % coverage of Oct1, with a score and 77 % of Oct2.

Oct4 and N-AGA were confirmed by MALDI-MS/MS from digestion of an excised SDS PAGE band of each protein. The digested fragments were ionised and ran down an ultrafleX III MALDI-TOF/TOF mass spectrometer. Tandem spectra obtained from each sample were then compared to a database containing the given sequences of each. Both proteins displayed good matches with the provided sequences with a coverage of 37 % for Oct4 and 17 % for N-AGA.

3.5 Identification of VFA generating *A. octavius* aminoacylase using a qualitative glutamine detection assay

With all *A. octavius* proteins and N-AGA successfully purified, attempts to determine if any of Oct1-4 could release either of the VFAs from their respective glutamine conjugate were conducted. Initial efforts to measure this activity utilised a spectrophotometric biochemical method with the use of a commercially available *in vitro* glutamine detection assay. This method would enable accurate monitoring of activity by the detection of L-glutamine which is produced stoichiometrically with that of the VFA when the precursor molecule is hydrolysed. The assay utilises a coupled biochemical endpoint reaction with the sample-derived glutamine first being deaminated by glutaminase to create L-glutamate and ammonia. The resulting ammonia is then used in a second reaction with 2-oxoglutarate to generate glutamate via glutamate dehydrogenase, causing oxidation of NADPH to $NADP^+$. As NADPH loss is equimolar to ammonia use in the reaction, the depletion of NADPH can be measured as a decrease in absorbance at 340 nm and allow quantification of glutamine levels in samples. As this assay detects the glutamine rather than 3M2H or HMHA it can be used as a standard test with all the candidate enzymes and doesn't require any specific adjustments for either VFA precursors.

To verify the suitability of this assay, a method validation experiment was performed using known standards of glutamine with the standard microplate protocol per the manufacturer's instructions. Experimental absorbance values were used to determine glutamine molarity using the manufacturers data analysis formulas and calculations. Attempts at method validation revealed a large discrepancy between experimentally calculated values and the actual known concentration of the standards (Fig 3.5.1). This pronounced error was observed across all glutamine concentrations despite the amount of glutamine present in the samples of ≤ 2 mM glutamine being within the stated range of detection for the assay. Experimental values for 0.1 mM also resulted in a supposed negative calculation which would imply this to be below the accurate limits of detection for this assay despite manufactures guidelines stating a lower detection limit of 68 µM.

These findings made this assay unsuitable for accurate quantification of enzyme activity with the candidate enzymes and VFA precursors. However, as this assay was the first method attempted for screening the candidate enzymes for VFA generating activity and did show sufficient signal across most glutamine ranges tested, it was used in a qualitative approach to identify any catalytically active enzymes.

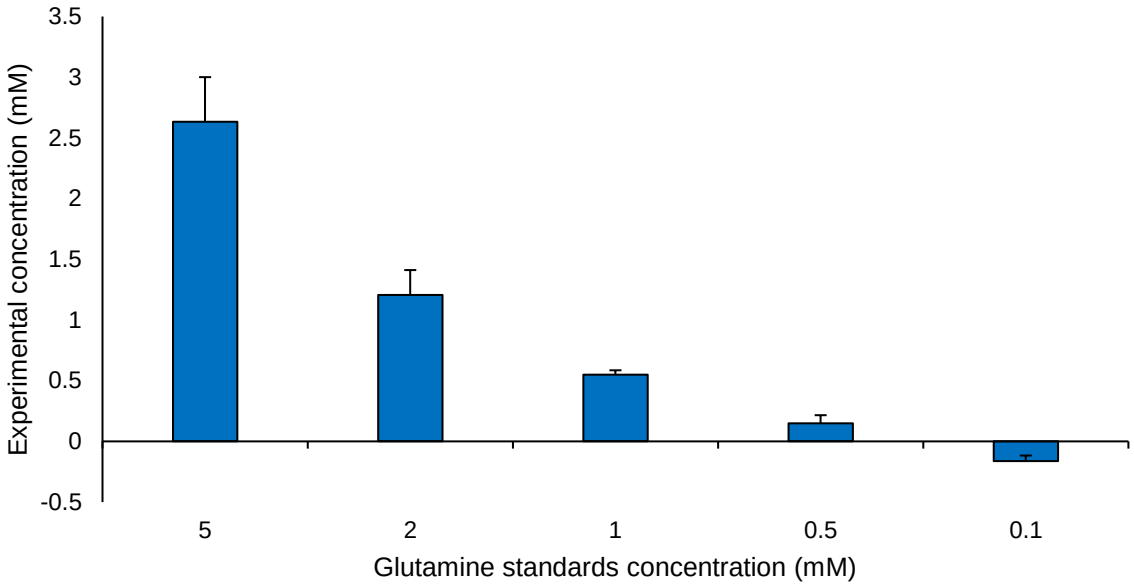


Figure 3.5.1: *In vitro* glutamine detection assay standards

Experimentally determined glutamine concentration of known glutamine concentration standards using glutamine detection assay. Error bars denote standard deviation (n=3).

All candidate enzymes were incubated with 2.5 mM of 3M2H-Gln or HMHA-Gln for 1 h at 37 °C in a pH 7.2 buffered medium before assessing the levels of glutamine released from VFA precursor hydrolysis. This screening did reveal activity for both substrates with some of the candidate enzymes (Fig 3.5.2). N-AGA displayed activity with both precursors as expected, matching previous finding by Natsch *et al*, and confirming that the assay can identify an active VFA generator. Oct3 displays a high activity with 3M2H-Gln but does not appear to have similar activity with HMHA-Gln, suggesting a difference in substrate specificity to N-AGA, and appears to be the major aminoacylase within *A. octavius* that is responsible for VFA generation. Oct1 and Oct2 appear to give very low levels of activity with both substrates, but the accuracy of this result is in doubt based on the large error observed from method validation (Fig 3.5.1). While Oct4 displayed higher activity than Oct1 and Oct2, this result is also not conclusive based on this assay alone.

Despite the limitations of the assay, Oct3 has been identified as a clear 3M2H producer and likely the main enzyme responsible for VFA generation in *A. octavius.*

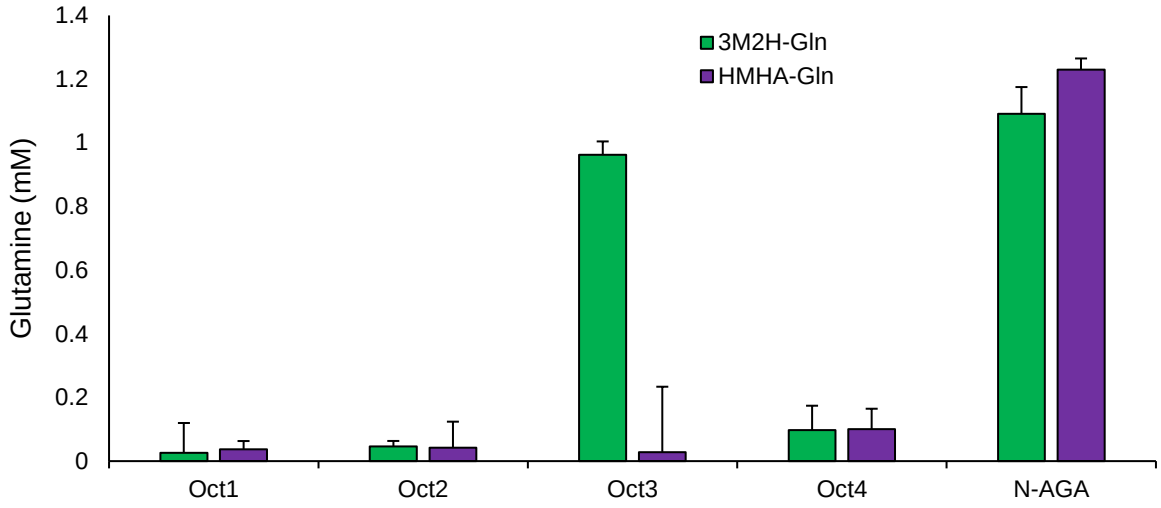


Figure 3.5.2: Glutamine detection assay screen of candidate *A. octavius* aminoacylases

Activity based on amount of glutamine formed after 1 h incubation of 2.5 µM enzyme with 2.5 mM 3M2H-Gln or HMHA-Gln. Error bars denote standard deviation (n=3).

3.6 Cloning of *A. nagyae* aminoacylases candidates

With the identification of Oct3 as a malodour generating aminoacylase, characterisation of candidate malodour contributing aminoacylases from the other known VFA producing *Anaerococcus* species, *A. nagyae*, was conducted in a same manner as for *A. octavius*. *Anaerococcus nagyae* has been reported to hydrolyse HMHA-Gln and release HMHA from an *in vivo* whole cell biotransformation assay while 3M2H release was not studied (Fujii et al., 2014). The ability of *A. nagyae* to generate HMHA differs to the activity displayed by Oct 1-4 in the glutamine detections screen and suggests that the two *Anaerococcus* species have different VFA generating capabilities. Study of the enzyme responsible for VFA generation in *A. nagyae* was conducted to identify the enzyme responsible for malodour generation in this species, determine whether this species has an enzyme able to hydrolyse 3M2H-Gln and study its similarity to that of Oct3.

Earlier bioinformatic analysis revealed the presence of five putative aminoacylase sequences within *A. nagyae* and noted that four of them displayed high similarity to a respective homologue within *A. octavius*, with these being labelled Nag1-4 matching the nomenclature of Oct1 -4 and the fifth protein being labelled Nag5 (Fig 3.2.1). Due to the high similarity of Nag1-4 to their counterparts in *A. octavius* it was hypothesised that Nag3, the closest homologue of Oct3, would be an orthologue and display activity with the VFA substrates. As there is no close representative to Nag5 within *A. octavius* that has been tested it is not possible to infer the likelihood of activity of this enzyme from the previous activity screening and bioinformatics analysis. Based on this Nag3 and Nag5 were chosen to be characterised *in vitro* in the same manner as Oct1-4.

The gene DNA sequence of *nag3* and *nag5* were each synthesised as doubled stranded gene-length DNA molecules (ThermoFisher) and cloned into a pET-21b(+) vector. This vector was chosen due to its similarities to the pGM vector used for the Oct and N-AGA proteins (Appendix A). The cloning was performed as detailed in methods but in brief, the synthesised genes each had their original stop codon removed (enabling the formation of a fusion protein with a C-terminal hexahistidine tag on the vector) and contained 20 bp at both 5' and 3' that were homologous to upstream and downstream respectively of the vector cloning site. The pET-21b(+) vector was linearised by PCR to contain homologous ends to that of the 5' and 3' ends of the insert at the ORF and enabled recombination of the vector and insert using ClonExpress II Ones step cloning kit. The resulting circularised insert containing vector was immediately transformed into *E. coli* DH5α chemically competent cells by the heat shock method and transformant colonies selected for with ampicillin resistance selection (100

µg/ml) on LB agar plates. Transformant colonies present after ~24 h incubation (37 °C) were picked for liquid culture growth (LB + 100 µg /ml ampicillin) overnight, before plasmid extraction and plasmid sequence verification by sanger sequencing (Eurofins) using sequencing primers homologous to the T7 promoter and terminator sequences upstream and downstream of the ORF respectively.

3.7 Purification and MS verification of *A. nagyae* aminoacylases

Both Nag3 and Nag5 were purified using the same protocol as that for the previous *A. octavius* proteins and N-AGA with a two-step purification of IMAC (Fig 3.7.1 A) and SEC (Fig 3.7.1 B). Immobilised metal affinity chromatography of Nag3 and Nag5 produced similar purification traces, in terms of intensity and elution time, to those expressed from the pGM vector system indicating that pET-21b vector can encode soluble protein with a functional histidine tag. Analysis of purification fractions by SDS PAGE for Nag3 showed a high yield of soluble protein is present and retained through from IMAC to SEC purification steps (Fig 3.7.1 C). The observed monomer band for Nag3 is present at roughly the correct position on SDS PAGE based on its actual molecular weight (45205 kDa). An SDS PAGE analysis of Nag5 purification fractions also showed a high yield of soluble intact monomeric protein present after purification and means all aminoacylase proteins purified to date have yielded soluble protein (Fig 3.7.2 A) (Table 3.7.1).

Purification of Nag3 and Nag5 via SEC also resulted in similar elution profiles to previous proteins with the presence of two unresolved species creating a widened peak (Fig 3.7.2 B). Nag5 displayed a slightly later elution time to all other purified aminoacylases which could be indicative of a difference in protein shape and/or oligomeric shape compared to all other protein given the similarity in monomeric molecular weight to the other proteins.

Both purified Nag proteins were confirmed to have the correct primary sequence by MALDI-MS/MS. A protein band for each protein was excised from a 12 % SDS gel after electrophoresis size separation. Each band underwent tryptic digest and the fragments analysed by MS in the same manner as Oct4 and N-AGA. This analysis confirmed both proteins matched the expected sequences with fragments detected to cover 50 % and 70 % of the full sequence of Nag3 and Nag5 respectively.

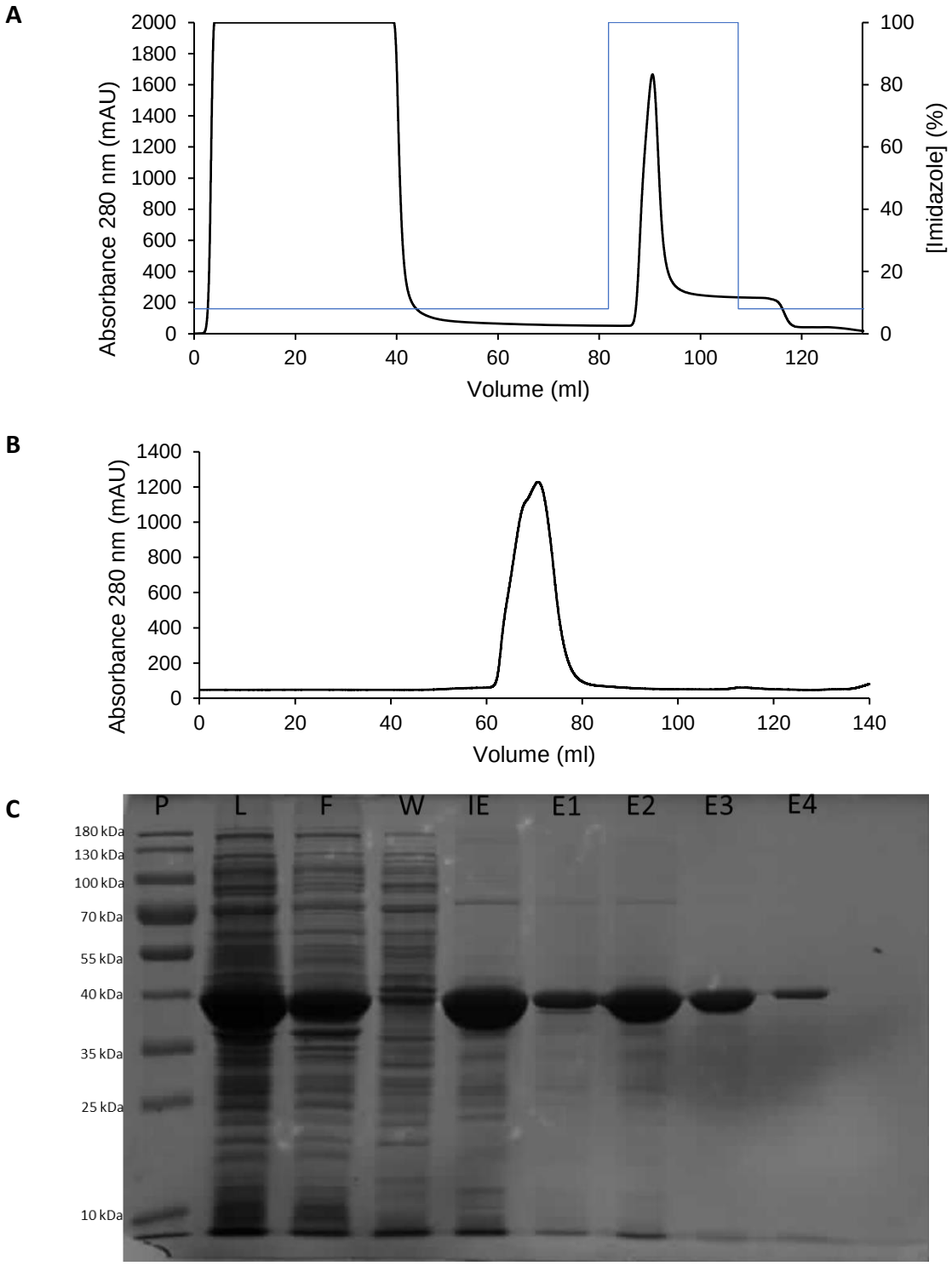
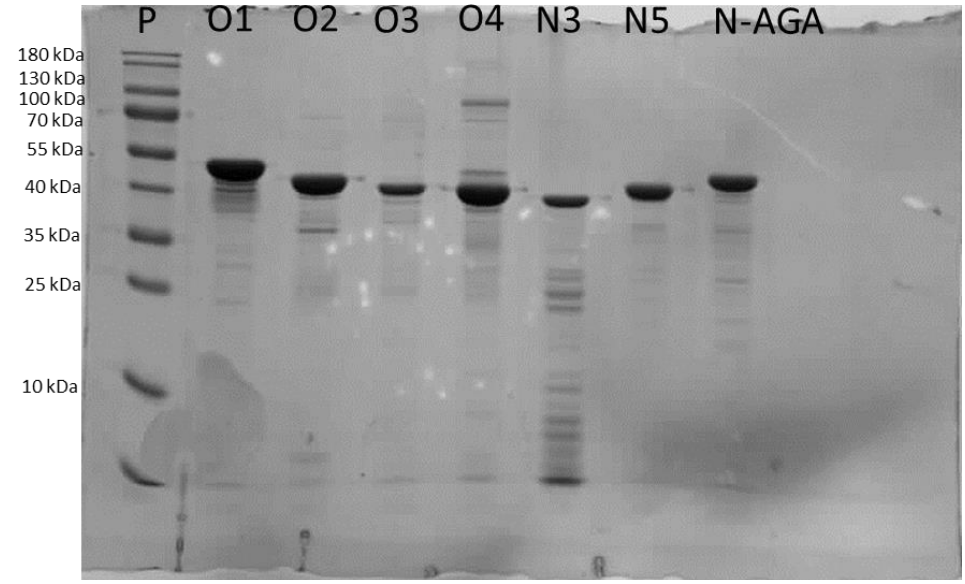


Figure 3.7.1: Purification of *Anaerococcus nagyae* aminoacylase, Nag3

A. Immobile metal affinity chromatography (IMAC) trace of the wash and elution of Nag3 from cell lysate of *E. coli* BL21 (DE3) recombinantly expressing Nag3 from a pET21-b vector . The black line details the A_{280} and the blue line represents the % concentration of imidazole present in the solution (percentage denotes 0 – 500 mM range of imidazole).
B. Size exclusion chromatography (SEC) of Nag3. IMAC purified elution fractions containing the highest purity of Nag3 were loaded onto a HiPrep 16/60 Sephacryl S-200 column.
C. 12 % SDS PAGE gel of Nag3 purification. Gel lanes contain (from left to right): (P) Page Ruler prestained protein ladder, (L) lysate of *E. coli* BL21 (DE3) expressing Nag3, (F) IMAC flow through, (W) IMAC wash through, (IE) IMAC pooled elution fractions of Nag3, (E1-4Oct3 SEC fractions 9 -13.

A



B

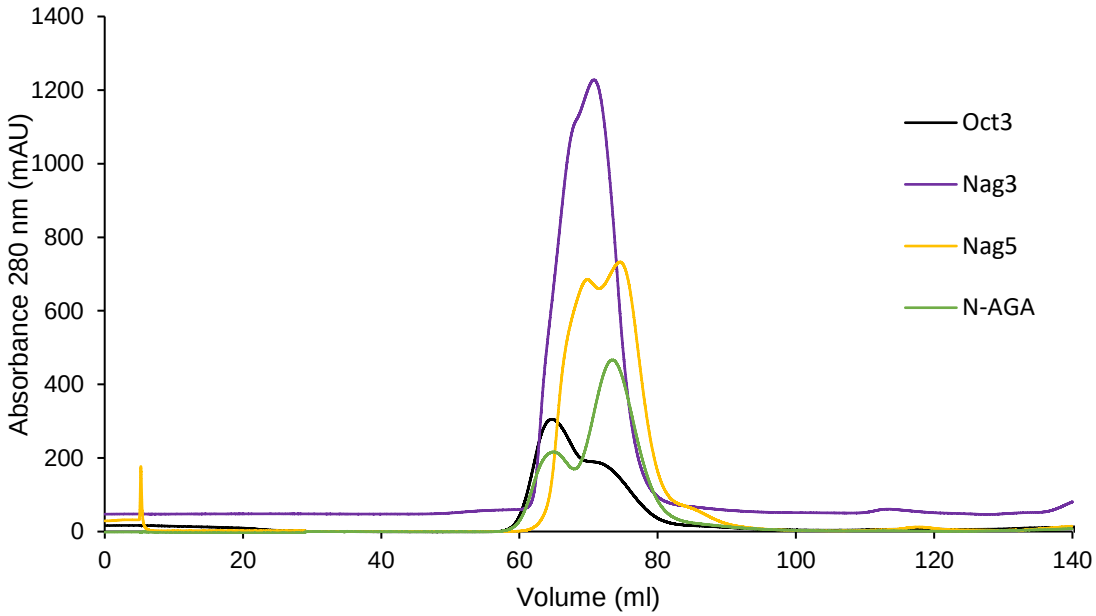


Figure 3.7.2: Purified aminoacylase SDS PAGE and Nag protein gel filtration comparison

A. SDS-PAGE gel (12 %) image of aminoacylase proteins after IMAC and SEC purification. From left to rights lanes contain: (P) Page Ruler prestained protein ladder, (O1) Oct1, (O2) Oct2, (O3) Oct3, (O4) Oct4, (N3) Nag3, (N5) Nag5 and N-AGA.
B. Size exclusion chromatography traces recording Abs_{280} against volume liquid for each purification run charted.

Table 3.7.1: List of purified aminoacylase proteins				
Protein	Host organism	Expression vector	MS verification method	Molecular weight, including His tag (kDa)
Oct1	*A. octavius*	pGM	LC-MS/MS	50.17
Oct2	*A. octavius*	pGM	LC-MS/MS	44.87
Oct3	*A. octavius*	pGM	ESI-MS	44.84
Oct4	*A. octavius*	pGM	MALDI-MS/MS	44.10
Nag3	*A. nagyae*	pET-21b(+)	MALDI-MS/MS	45.21
Nag5	*A. nagyae*	pET-21b(+)	MALDI-MS/MS	44.20
N-AGA	*C. sp. Ax20*	pGM	MALDI-MS/MS	44.81

3.8 Screening of Nag3 activity using a qualitative glutamine detection assay

The Nag3 protein was purified before Nag5 and was screened for VFA activity using the same glutamine detection assay as used for the Oct and N-AGA proteins. Nag3 was incubated with 2.5 mM 3M2H-Gln or HMHA-Gln for 1 h at 37 °C in a buffered solution with the amount of glutamine formed being measured using Megazyme L-glutamine detection assay.

As hypothesised by the bioinformatics analysis, Nag3 shares the ability of Oct3 to generate VFAs with both being able to generate 3M2H confirming these proteins as orthologous (Fig 3.8.1). However, an interesting difference does appear in hydrolysing HMHA-Gln with Nag3 displaying high activity more comparable to that exhibited by N-AGA rather than the almost negligible levels of Oct3**.** This result is surprising given the high sequence similarity of Nag3 to Oct3 (83.5 %) compared to N-AGA (30.5 %) which would have suggested Nag3 display an activity profile closer to its orthologue rather than it's distant *Corynebacterium* homologue.

While the glutamine detection screen did indicate that Nag3 is the likely main enzyme responsible for VFA generation in *A. nagyae*, it could be possible that Nag5 also contributes to VFA malodour generation in this species. The limitations in the accuracy of this assay also make it difficult to confidently infer anything between the differing activities displayed by the active VFA generating enzymes.

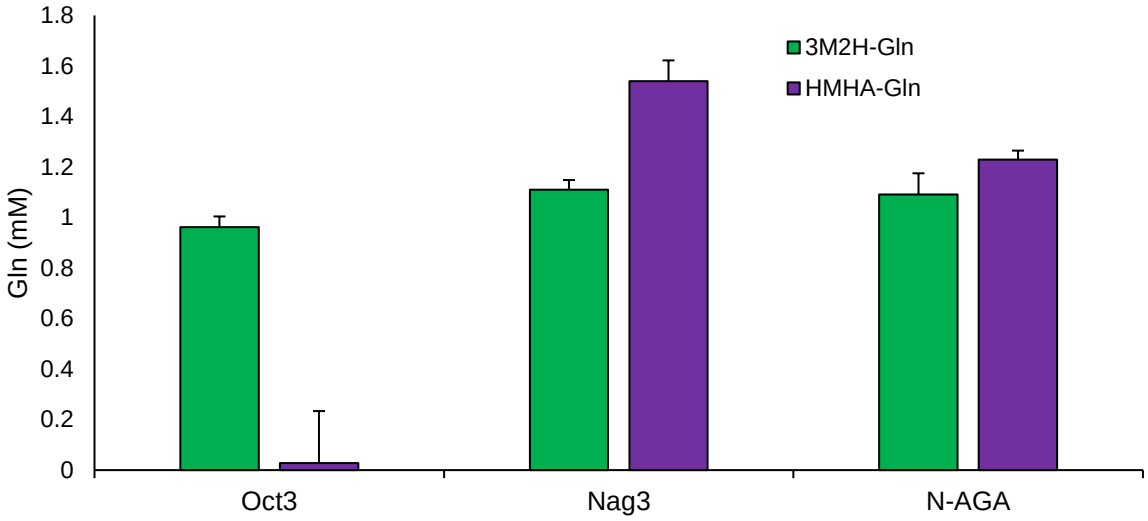


Figure 3.8.1: Nag3 glutamine detection assay screen

Activity based on amount of glutamine formed after 1 h incubation of 2.5 µM enzyme with 2.5 mM 3M2H-Gln or HMHA-Gln. Error bars denote standard deviation (n=3).

3.9 Ninhydrin assay development for monitoring VFA generating activity

The failure of the glutamine detection assay to provide reliable and accurate measurements of enzyme activity necessitated the development of another technique. This technique had to provide a more accurate quantification of enzyme activity and ideally be comparatively quick and straightforward like the glutamine detection assay to enable a reasonably high throughput. Potential options found within the literature included fluorescent detection of free amino groups with fluorescamine (Natsch et al., 2003), a spectrophotometric assay using carboxylic acid detection by ferric hydroxamate (Siedlecka et al., 2008) and amino acid detection using a coupled enzyme reaction with an L-amino acid oxidase and horse radish peroxidase to activate a dye (Andreo-Vidal et al., 2018). After assessing various options, a biochemical assay using ninhydrin was chosen as it fulfils the stipulated requirements and assay development was conducted.

Ninhydrin reacts with primary amines, including the α-amino group of amino acids (NH_2-CH(R)-COOH), to form Ruhemann's purple (RP) which can be measured spectrophotometrically at 570 nm (Moore and Stein, 1948) and has been used extensively to quantify levels of amino acids and peptides in a variety of biological fields (Friedman, 2004). Ninhydrin interacts with α-amino groups and does not react with amides (Warneford, 1916), therefore the reagent should only react with free glutamine produced alongside the VFA by malodour generating aminoacylases rather than the glutamine moiety present in both substrates. This results in the formation of RP, which is equivalent to the amount of glutamine present, being a proxy for VFA formation and enables quantification of enzyme activity. Ninhydrin has the capability to react with peptides and proteins due to the presence of an N-terminal residue amino group and the primary amine group of a lysine side chain and guanidine group of arginine however in the case of proteins this reaction has been shown be weak and unsensitive under most conditions, resulting in minimal effects upon signal generation in activity assays (Friedman, 2004; Nayuni et al., 2013).

Development and optimisation of the ninhydrin assay was performed using Nag3 because of its observed high activity with both VFA precursors and typically higher yield from purification compared to Oct3 and N-AGA. Due to the limited supply of both 3M2H-Gln and HMHA-Gln and its cost in procurement, initial work used a model substrate in place of the VFA precursors. One substrate used extensively in aminoacylase characterisation studies is N^{α}-acetyl-methionine (Ac-Met) (Jamdar et al., 2015; Lindner et al., 2003; Story et al., 2001; Tanimoto et al., 2008), likely due to its historic use in the characterisation of aminoacylase 1 and its biotechnological applications with enantiomer selection

from racemases (Gade and Brown, 1981; Tewari, 1990; Toogood et al., 2002). To facilitate a high throughput, the assay was attempted using a microplate format and microlitre volume scale which has been successfully performed in other studies (Iyamu et al., 2008) rather than the more common cuvette and millilitre approach (Moore and Stein, 1948; Sun et al., 2006; Zhang et al., 2013).

Before method development for a microplate was attempted, Nag3 was screened for activity with Ac-Met given the major differences of this model substrate to that of the VFA substrates. Incubations of 5 µM Nag3 with concentrations of Ac-Met ranging from 1 – 6.25 mM were prepared in a HEPES buffered medium (pH7.2) in a total volume of 200 µl in microcentrifuge tubes. The reaction was initiated with enzyme addition and incubated at 37 °C for 1 h before quantification. Substrate and enzyme only controls were performed with the omitted component being replaced with additional HEPES buffered medium and a positive control of 6.25 mM Met was included to assess ninhydrin reaction validity. A 50 µl volume of each reaction was added to 40 µl of a 2 % ninhydrin solution (3.5 g/L ninhydrin in 100 % ethanol) and heated to 80 °C for 30 min to catalyse the reaction before addition of 50 µl dH_20 to dilute and cool the sample. The Abs_{570} of the sample was then measured using a microcuvette in a spectrophotometer. The ninhydrin assay was able to detect free Met and generated no signal for the acetylated substrate (Fig 3.9.1 A). Nag3 was observed to be active with Ac-Met in a substrate concentration dependent manner with collected absorbance values exceeding enzyme only controls although minimally at 1 mM Ac-Met. The finding that Nag3 can catalyse the release of Met from its acetylated substrate suggests that this enzyme has promiscuous substrate range given the clear differences between Ac-Met and the VFA precursors in both the amino acid and acyl moieties.

A calibration curve of free methionine using the same ninhydrin assay conditions was attempted but generated a non-linear trendline across the concentration range used (Fig 3.9.1 B). A linear relationship could be established between 1.25 – 1.75 mM Met but this is far above the linear detection range noted in other studies (Friedman 2004; Sun et al. 2006). This lack of sensitivity was attributed to the ninhydrin solution used which was unsuitable for accurate detection for several reasons including: lack of a reducing agent which has been shown to aid RP formation (Moore 1968; Sun et al. 2006), absence of hydrindantin (a reduced complex of two ninhydrin molecules) which has been shown to prevent aminated ninhydrin being lost to a side reaction of hydrindantin formation rather than RP (Bottom et al., 1978; Friedman, 2004; Lamothe and McCormick, 1973) and the preparation and storage of the ninhydrin in air which oxidises the solution affecting both of the previous points.

Subsequent assay development utilised a commercially prepared solution of ninhydrin reagent (Sigma Aldrich) containing 2 % ninhydrin and 0.5 % hydrindantin dissolved in DMSO in a lithium acetate buffer

pH 5.2. This solution was aliquoted into smaller working volumes using an anaerobic cabinet and stored at 4 °C in sealed containers. This change, coupled with an increase in ninhydrin per reaction to 50 µl from 40 µl, led to a dramatic improvement in absorbance signal and assay sensitivity with a linear relationship observed from 0.05 mM to 2 mM methionine to create a calibration curve (Fig 3.9.1 C). Samples recording Abs_{570} values greater than 1 were further diluted with dH_2O as necessary.

Further refinements of the ninhydrin assay were undertaken using both experimental optimisation and the literature (Sun et al. 2006). These iterative improvements resulted in a quicker assay using less sample per reaction, improving throughput and reducing the need for excessive dilution, and a shortened heating step. To aid later biochemical characterisation of aminoacylases a method of enzyme quenching was developed to ensure accurate activity time courses could be collected. The M20 aminoacylase family are all metalloenzyme requiring divalent metal ions for activity and are known to be inhibited by metal collating agents (Rawlings and Barrett, 2013) with this effect having being previously observed for N-AGA (Natsch et al., 2003). Given this, ice cold EDTA was used as a quenching agent upon sample extraction for ninhydrin analysis.

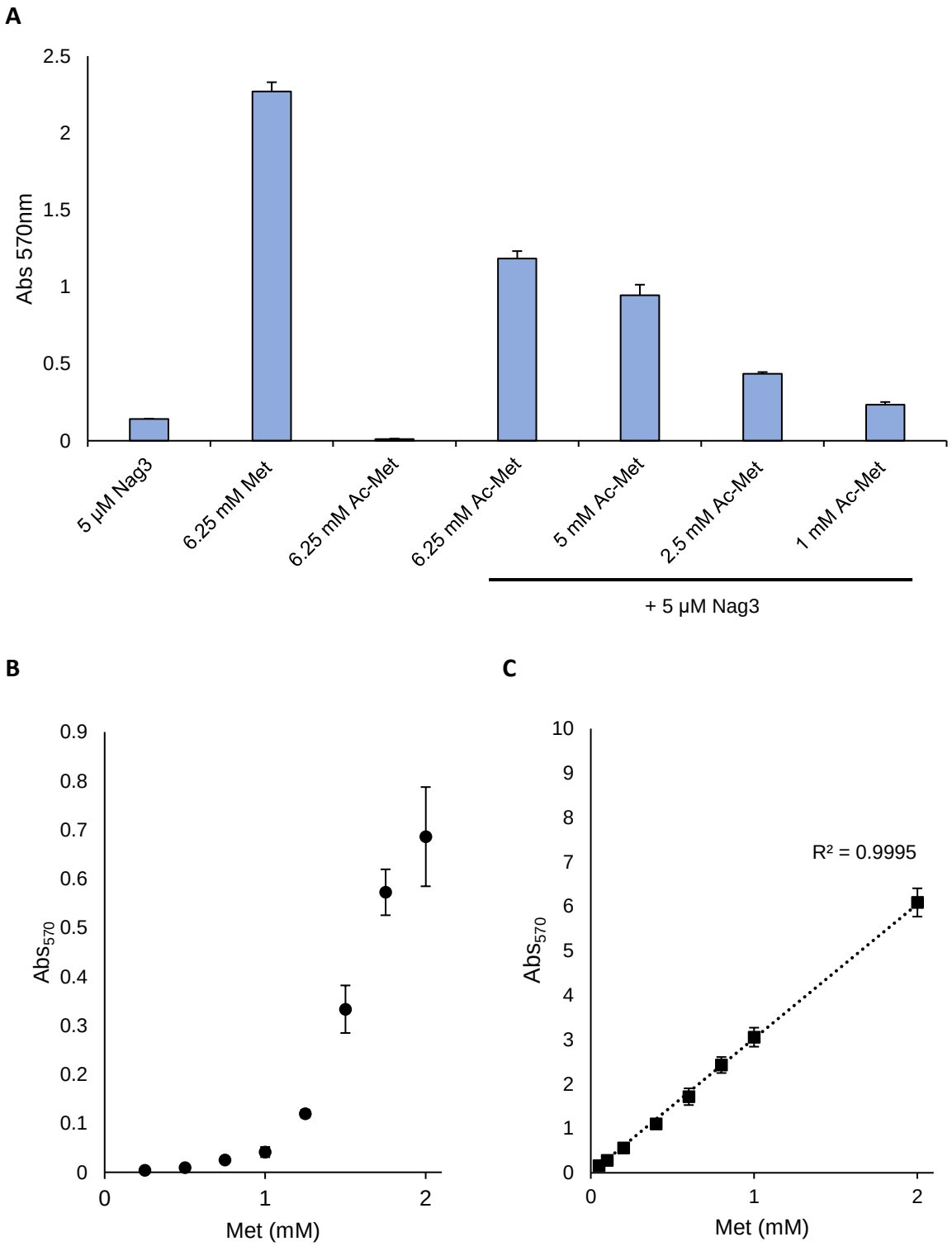


Figure 3.9.1: Ninhydrin assay development using a cuvette format

A. Preliminary assay validation of Nag3 activity with Ac-Met. Three controls of Nag3, Met and Ac-Met only were performed and displayed in that order from left to right. Error bars denote standard deviation (n=3).
B. L-methionine calibration curve from cuvette method and using 2 % ninhydrin solution in ethanol. Error bars denote standard deviation (n=3).
C. L-methionine calibration curve from cuvette method and using commercial ninhydrin reagent. Error bars denote standard deviation (n=3).

All improvements together coupled with full protocol adaptation to microplate format resulted in the following protocol which was hence used for all subsequent ninhydrin analysis assays unless otherwise stated. A 20 µl volume of experimental was added to 20 µl ice cold 25 mM EDTA before addition of 50 µl 2 % ninhydrin reagent. The solution was sealed and heated in a thermocycler for 10 min, 70 °C and cooled before mixing with 100 µl dH_2O. A 180 µl volume of the solution was transferred to a microplate well and the absorbance of the sample at 570 nm determined in a plate reader.

This optimised microplate method was used initially for time course activity assays of Nag3 with Ac-Met to ensure that accurate measurements could be taken. Reactions of 1.85 µM Nag3 with Ac-Met concentrations varying from 1 mM to 2.5 mM were made up to 200 µl with a HEPES buffered medium in a microplate. The reactions were initiated with enzyme addition and incubated 37 °C statically with 20 µl samples taken from each well at 0, 10, 20, 30, 40, 50 and 60 min time points. Collected samples were immediately quenched in ice cold EDTA and left on ice until the final 60 min time point was retrieved before commencement of the rest of the ninhydrin assay protocol detailed above. Absorbance values collected were converted to methionine concentration using a L-methionine calibration curve using the same ninhydrin protocol and microplate format as that used for the time course assay (Fig 3.9.2 A).

Nag3 displayed high activity with Ac-Met with significant levels of methionine being liberated at all substrate concentrations after 1 h (Fig 3.9.2 B). The activity of Nag3 with each concentration was accurately measured across all time points and increases in methionine concentration correlated with substrate concentration validating that this assay can make accurate quantitative measurements. The increasing levels of Met liberated over time display linearity across the shorter time points before beginning to plateau at later intervals, characteristic of an enzyme- substrate kinetic reaction. The observation of this effect using this assay further validates this assay's suitability for quantitative measurements as well as determination of Michealis-Menten kinetic parameters. All samples were only assayed after the last were collected indicating that the metal collation by EDTA is capable of quenching the enzyme reaction and making it ideal for subsequent experiments.

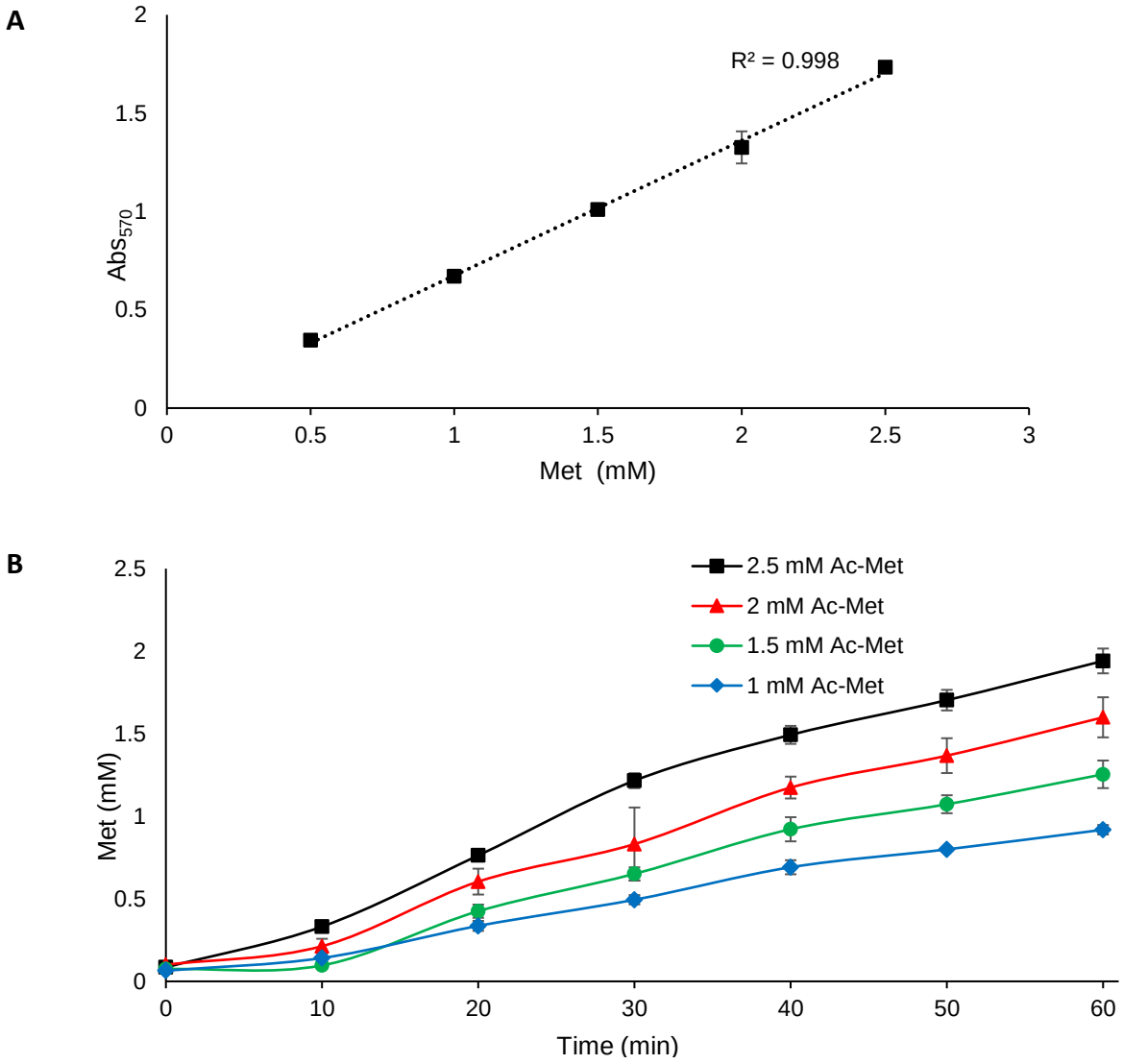


Figure 3.9.2: Nag3 time course activity assay with Ac-Met

A. L-methionine calibration curve from microplate method. Error bars denote standard deviation (n=3).
B. Nag3 activity time course with noted concentrations of acetyl-L-methionine. Error bars denote standard deviation (n=3).

3.10 Ninhydrin assay screen of purified aminoacylases for VFA activity

The developed microplate ninhydrin assay was used to screen all purified aminoacylases to validate the previous glutamine detection assay and enable quantitative measurements of activity. Each aminoacylase was assayed with 1.85 µM enzyme incubated with 2.5 mM 3M2H-Gln or HMHA-Gln in a 200 µl reaction using a HEPES buffer (pH7.2). The reactions were initiated by protein addition and incubated at 37 °C, with samples taken for ninhydrin assaying at 1 h, matching previous activity screening, as well as an additional time point of 24 h to identify very low activity that could be missed during the shorter incubation. The concentration of glutamine present in each sample was quantified using a calibration curve of free L-glutamine (Fig 3.10.1).

The ninhydrin screen generally matched that of the previous glutamine detection assay with Oct3, Nag3 and N-AGA displaying high levels of activity after 1 h and Oct3 exhibiting a clear substrate specificity for 3M2H compared to the more promiscuous Nag3 and N-AGA (Fig 3.10.2 A). Both N-AGA and Nag3 displayed activity with HMHA-Gln but this was to a lesser degree than with 3M2H-Gln, suggesting that 3M2H-Gln is the preferred substrate for all active enzymes. Low levels of 3M2H-Gln hydrolysis were also detected for Oct4 and Nag5 after 1 h indicating that VFA generation is not solely caused by a single enzyme in each of *A. octavius* and *A. nagyae.* However, the lower extent of this activity compared to Oct3 and Nag3, even after 24 h, would support the hypothesis that Oct3 and Nag3 are the primary producers of VFA malodour (Fig 3.10.2 B). Oct3, Nag3 and N-AGA continued to have the highest levels of activity after 24 h with Oct3 displaying the greatest increase in activity to attain a similar level of 3M2H-Gln hydrolysis as Nag3 and N-AGA as well as display clear activity with HMHA-Gln which is not significant after only 1 h. The finding that Oct3 can hydrolyse HMHA-Gln to a low amount indicates that Oct3 has a strong substrate selectivity against HMHA-Gln rather than being catalytically inactive with the substrate. No activity was observed for Oct1 and Oct2 with either substrate even after 24 h incubation indicating they are not responsible for VFA generation although the absence of any activity means it could be possible the purified enzymes are catalytically inactive.

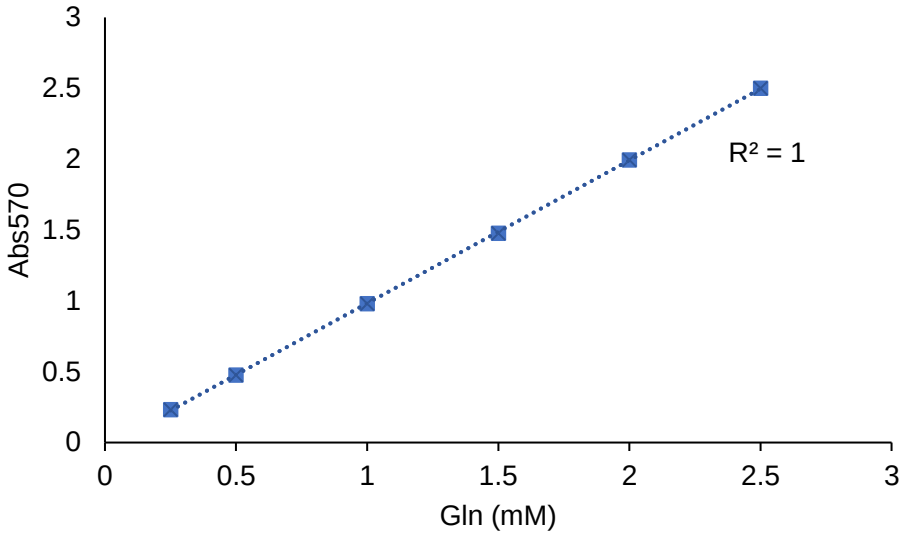
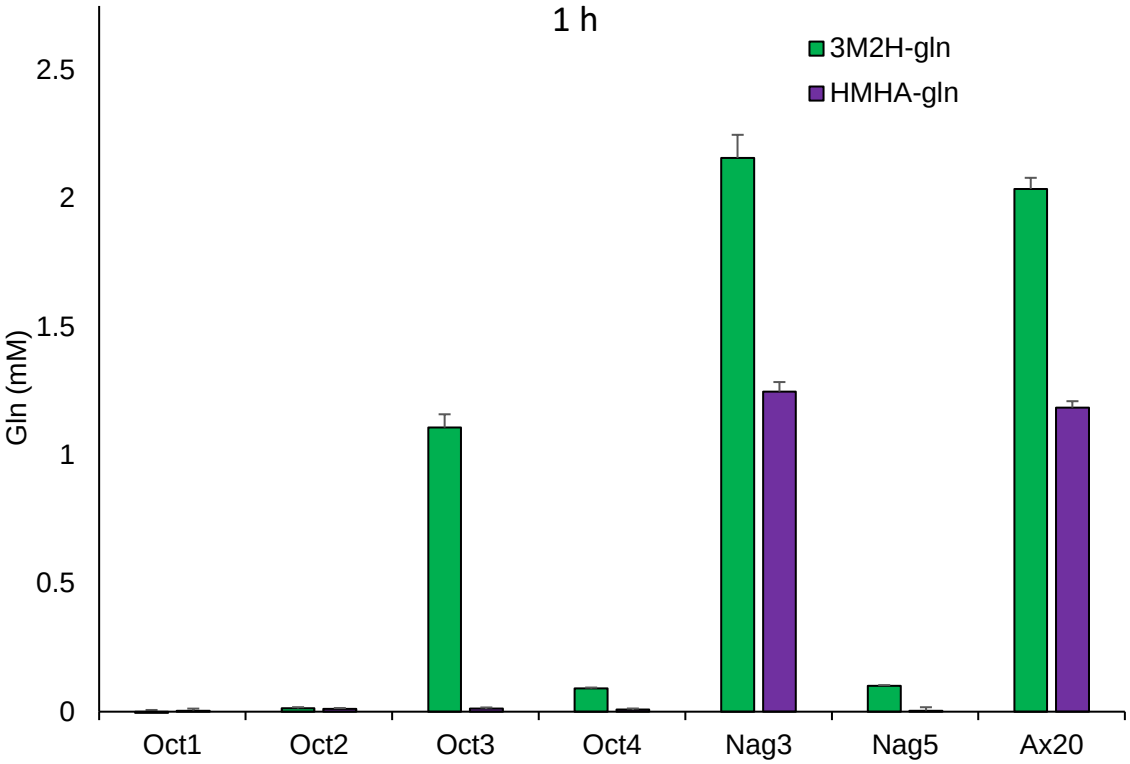


Figure 3.10.1: L-Glutamine calibration curve

Calibration curve generated using known concentrations of L-Glutamine. Quantification of glutamine in experimental samples uses the trend line displayed. Error bars denote standard deviation (n=3).

A



B

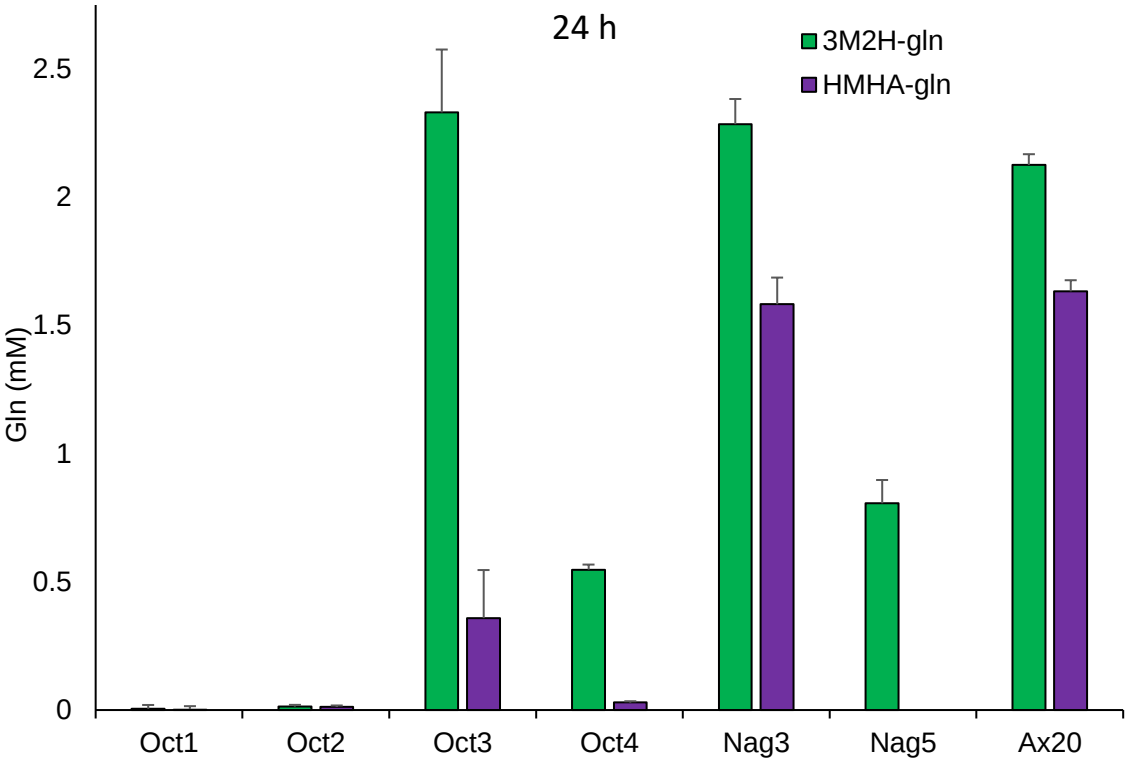


Figure 3.10.2: Aminoacylase VFA activity ninhydrin screen

A. Glutamine quantified via ninhydrin assay after 1 h incubation (37 °C) of 1.85 µM enzyme with 2.5 mM 3M2H-Gln or HMHA-Gln. Error bars denote standard deviation (n=3).
B. Glutamine quantified via ninhydrin assay after 24 incubation (37 °C) using the same reactions as used for the 1 h time point. Error bars denote standard deviation (n=3).

3.11 Discussion

This chapter has focused on the identification of the enzymes responsible for medium length VFA generation in *Anaerococcus*. A bioinformatics approach searching for proteins with sequence homology to the identified VFA producing aminoacylase N-AGA from *Corynebacterium* sp. Ax20 (Natsch et al., 2003) within *Anaerococcus* identified sequences across the genus, although none had a high identity to N-AGA. Preliminary internal work at Unilever had suggested that *A. octavius* has a role in VFA axilla malodour and this species contained four sequences identified by the bioinformatics work suggesting one of these candidates would be responsible for VFA generation. Five sequences with homology to N-AGA residing in *A. nagyae*, a related species to *A. octavius* (Veloo et al., 2016), were also discovered and of interest due to this species being previously identified as able to release HMHA from its glutamine precursor during an *in vivo* whole cell screen (Fujii et al., 2014). The two species are each other's closest relative within the genus (Veloo et al., 2016), and this similarity extends to the phylogenetic tree of putative aminoacylases where each *A. octavius* protein sequence has a related homologue within *A. nagyae* that resides either as the closest or a near sequence. Given the closeness of these two species and both being implicated in VFA malodour the logical hypothesis was that both can liberate VFAs via orthologous proteins. However, the fact that N-AGA can generate VFAs and has a sequence identity <40 % with most of the *Anaerococcus* aminoacylase sequences already implies that a high sequence homology alone is not a definitive indicator of activity. This necessitated the screening of all Oct proteins to identify which is active with the precursors.

Purification of all proteins gave high yields of all proteins with little need for expression optimisation. In fact, protein yields were also satisfactory using shorter 5 h induced expression cultures at 37 °C but the switch to express for 20 h at 20 °C was made due to its improvements in protein yield and ease. All proteins expressions would consistently yield > 60 mg protein from 1 L harvested cell culture with Nag3 even reaching > 120 mg indicating that not only are the proteins readily expressible from their vectors but also have a good cytoplasmic stability and exhibit little toxic effect upon the host cells. Size exclusion chromatography of all proteins exhibited either an asymmetrical single or double peak absorbance trace upon protein elution which is characteristic of multiple oligomeric states of the protein being present. The M20 family are known to typically form a homodimer to enable activity with key interactions occurring between residues from each monomer's oligomerisation domain and the substrates residing in their opposite numbers binding site (Jozic et al., 2002; Lindner et al., 2003; Natsch and Emter, 2020). This would suggest that the most likely oligomeric species present are that of monomer and dimer although there have been several observations of aminoacylases assembling

into a homotetramer, particularly in structural studies of these proteins (Botelho et al., 2011; Girish et al., 2013), including N-AGA (Natsch and Emter, 2020). However, this oligomeric state could well be a crystallographic effect formed in the asymmetric unit and could not be the oligomer state assumed in solution. The publishing of a monomeric structure from *Arabidopsis thaliana* (Bitto et al., 2009) would support this assumption as while some M20 members can form an active monomeric state this has only been noted in members that have a duplication in the oligomeric domain, thereby mimicking the interactions formed in their dimeric counterparts (Chang et al., 2010; Jozic et al., 2002), and is not present in *A. thaliana* ILL2 suggesting require dimerization would be required for activity. Interestingly though, N-AGA has been noted to assume a tetrameric organisation in solution upon addition of a competitive inhibitor (Natsch, 2015) although whether this would be a genuine occurrence upon addition of VFA precursors or is just an effect of inhibitor interaction in unclear.

The use of a commercially available assay to detect the amount of glutamine present from enzyme substrate incubation did enable identification of the major VFA liberating enzymes from those tested but also displayed a clear lack of accuracy when tested with known standards. This was surprising given that this kit has been used in several other studies from various fields to detect glutamine levels in solution, although more often to analyse media rather than to monitor *in vitro* enzyme activity (Daud et al., 2016; Kaushik et al., 2020; Sigmarsdottir et al., 2021). The glutamine levels being measured was well within the apparent ranges stated by the manufacturer for the microplate format used, with a supposed linear relationship between absorbance change and concentration for a glutamine range of 0.1-4 µg per assay sample. Given these conditions, a 2 mM glutamine standard would translate to a final amount of 2.92 µg in this assay which is well within the reported accurate range. This failure to accurately quantify even known standards could suggest a potential problem with one of the components of the kit although the shift to development of the ninhydrin assay means the exact cause of this inaccuracy is unknown.

Despite this inaccuracy the glutamine detection assay did identify the same active VFA liberating enzymes as the ninhydrin assay with it appearing that orthologous enzymes within *A. octavius* and *A. nagyae* are the principal producers of VFA malodour in each species. While some activity was displayed by other Anaerococcus enzymes, the low activity combined with it being most prominent only after 24 h incubation would suggest these enzymes have a promiscuous binding site that can accommodate the VFA precursors rather than them being one of their primary physiological substrates. It has been well documented that M20 members often display a diverse substrate range (Bitto et al., 2009; Cao et al., 2020; Jamdar et al., 2015; Lindner et al., 2003; Natsch et al., 2003; Tanimoto et al., 2008), and could be posited that this diversity is enabled by flexibility in the binding

site while retaining metal coordination and catalytic residue organisation to maintain the universal catalytic reaction of amide bond hydrolysis. In fact, examples of up to 10 Å movements in the catalytic and oligomerisation domains in response to substrate binding have been observed in the M20 metalloenzyme DapE (Nocek et al., 2018) and less extreme active site spatial adaptations to substrate binding have been observed in other members (Jozic et al., 2002).

It was surprising to note that the orthologous proteins display such a contrast in activity with HMHA-Gln despite the similarity in both sequence identity (84 %) and the two host species. It could be assumed that the preference of Oct3 for 3M2H-Gln would indicate adaptation specifically to this substrate at the expense of HMHA-Gln although this does not seem to translate to better activity with 3M2H-Gln when compared to the activity of the more promiscuous Nag3 and N-AGA. The ability of Nag3 to also hydrolyse Ac-Met would instead suggest that in this case a promiscuous binding site enables greater activity with VFA precursors than specialisation for a single substrate. End point activity assays do not provide definitive analysis of the factors behind these different activity profiles and are something which is explored in depth in chapter 4 with Michaelis-Menten kinetic determination for the VFA generating aminoacylases. Nevertheless, this clear activity difference raises an interesting question as to why Oct3 would display such a clear specificity for one precursor over the other. Both substrates are present in all adult's axilla sweat, with the notable exception of predominantly East Asian populations containing the SNP genotype NM_033151.3(*ABCC11*):c.538G>A p.(Gly180Arg) (Martin et al., 2010; Yoshiura et al., 2006), and evidence suggests that HMHA-Gln is the more abundant of the two (Natsch et al., 2003, 2006) with averaged quantities at mg/ L level detected in subjects (Troccaz et al., 2009). While there is no literature to date on the spatial structure and localisation of the microbiota of the axilla, both precursors are secreted from the same apocrine glands making it likely that all *Anaerococcus* species present would be exposed to both substrates. The most parsimonious explanation for these two species being able to hydrolyse the VFA precursors is for nutrient acquisition with use of the liberated L-glutamine, of importance in central nitrogen metabolism as well as glutamine biosynthesis (Reitzer, 2003). Furthermore, the VFAs are readily detectable by the human nose and form part of axilla BO suggesting no further metabolic breakdown of these odorous molecules, which would make glutamine utilisation the only, known, direct potential benefit to microbial growth. Oct3 selectivity therefore appears to be limiting the amount of nutrient available to its host compared to *A. nagyae* and *C. sp Ax20.* Although this does make the simplistic assumption that the transport mechanism for VFA precursor uptake has similar affinity for both molecules and this affinity is not radically different between homologous systems in each species. Precursor uptake will determine intracellular concentrations of substrates and could explain the *in vitro* activity profiles displayed here for the Oct3, Nag3 and N-AGA. Interestingly, activity with 3M2H-

Gln appeared the highest of the two precursors with all VFA liberating enzymes tested which would suggest this is the preferred substrate. This preference could indicate that 3M2H-Gln is a more generic substrate that is easily accommodated in M20 aminoacylases binding sites compared to HMHA-Gln which requires further specific adaptation.

The *in vitro* selectivity observed by Oct3 and not Nag3 does match with a previous *in vivo* study assessing HMHA release by *Anaerococcus* species (Fujii et al., 2014). Biotransformation assays measuring HMHA release from its glutamine conjugated substrate detected high levels of HMHA from *A. nagyae* after 24 h incubation (79 % HMHA release from 2 mM HMHA-Gln) while the second highest result was exhibited by *A. octavius* at only 1.3 %. Importantly these experiments were performed on cell lysates for each species, thereby measuring only the activity of the intracellular enzymes present rather than any effect of transport. The cell biomass used for the lysates was grown without the presence of HMHA-Gln indicating Nag3 is present in cells even in the absence of the VFA precursors and could suggest Nag3 also has additional functions within the cell. The promiscuous substrate range exhibited by other characterised M20 aminoacylases as well as Nag3 itself with activity with 3M2H-Gln, HMHA-Gln and Ac-Met would support this.

The identification of Oct3 and Nag3 has shown that their respective host species are able to generate VFA based malodour via homologous aminoacylases to that found in *C. sp Ax20* with N-AGA. This warranted further characterisation of these two orthologues in a biochemical and biophysical manner to enable a deeper understanding of this process.

Chapter 4 – Biochemical and biophysical characterisation of volatile fatty acid liberating aminoacylases

4 Biochemical and biophysical characterisation of volatile fatty acid liberating aminoacylases

4.1 Introduction

With the identification of Oct3 and Nag3 as the principle VFA releasing aminoacylases in their respective species, further characterisation of these proteins was conducted. The activity results observed for the orthologous aminoacylase in Chapter 3 clearly indicate a difference in substrate selectivity for the two precursor molecules so further biochemical characterisation was undertaken to investigate if this trend continues when tested with a broader range of acetylated substrates. Further biochemical experimentation to determine the Michaelis-Menten kinetics of Oct3, Nag3 and N-AGA with the VFA substrates was also performed to understand the variance in activity profiles exhibited.

Biophysical characterisation was also undertaken to explore the oligomeric state as well as determine the catalytic metal present in the active proteins. Commonly M20 aminoacylases assemble as an active dimer in solution although variation of this including monomer and homotetramer arrangements have been observed (Girish et al., 2013; Jozic et al., 2002) including N-AGA (Natsch, 2015; Natsch and Emter, 2020). Variation in catalytic metal present has also been observed with most utilising zinc but cobalt and manganese use has also been recorded (Girish and Gopal, 2010; Rawlings and Barrett, 2013).

All together this chapter aims to fully characterise Oct3 and Nag3 and enable a deeper understanding of VFA based malodour production in *Anaerococcus.*

4.2 Activity screen with acetylated amino acid

To explore the substrate specificity exhibited by the VFA liberating aminoacylases, an activity screen using N^{α}-acetylated amino acids was undertaken. This activity screen was also used with all purified aminoacylases in an effort to assign a putative role for these other aminoacylases.

The acetylated substrates were chosen to provide a broad representation of amino acid side groups as well as several structurally related to glutamine to explore the effect of the amino acid moiety upon enzyme activity. This gave a total of seven substrates to test with each of the seven purified proteins and were as follows: *N*-acetyl-L-methionine, *N*-acetyl-L-glutamate, N^{α}-acetyl-L-glutamine (Ac-Gln), *N*-acetyl-L-alanine, *N*-acetyl-L-phenylalanine, N^{α}-acetyl-L-arginine and N^{α}-acetyl-L-asparagine (Appendix B). To quantify levels of product formed in enzyme incubations, calibration curves for the free amino acids at known standard concentrations were created (Fig 4.2.1). All amino acids tested showed a linear relationship between amino acid concentration and absorbance with little variation observed between replicate measurements. Differences in colour development for each amino acid are observed despite each one containing a single α-amino group due to the chemical structure of each amino acid creating potential polar and/or steric parameters that affect formation of Ruhemann's purple (RP) (Friedman, 2004). Ninhydrin is known to not form RP with amides under standard conditions, preventing any potential additional colour formation with the R group of glutamine (Warneford, 1916). In the case of asparagine, a standard curve of absorbance at 340 nm was used (Fig 4.2.1 B) due to a specific ninhydrin-arginine derivative forming instead of RP which shifts the maximum absorption on a UV-visible spectrum to 340 nm from 570 nm (Friedman, 2004; Sheng et al., 1993), and has been observed in UV-visible spectra using this ninhydrin assay technique (data not shown).

Each enzyme (1.85 µM) was incubated (37 °C, 1 h) with the indicated 2.5 mM of substrate and the amount of free amino acid product released was calculated based on quantification using the respective standard curves (Fig 4.2.2). As previously noted, Nag3 displayed activity with Ac-Met but this trait was not observed in either Oct3 or N-AGA, suggesting that Nag3 has a much more promiscuous binding site and throwing doubt as to the use of Ac-Met as a model substrate (Jamdar et al., 2015; Lindner et al., 2003; Tanimoto et al., 2008). The highest activity for Ac-Met was observed with incubations of Oct4 which did also show low levels of activity with 3M2H-Gln. The extent of activity displayed here would suggest the physiological substrate of Oct4 is either Ac-Met and/or an acylated methionine molecule. Activity with Ac-Gln was restricted to Nag3 and N-AGA indicating tolerance in both proteins to variation in the acyl chain of a glutamine conjugated molecule whereas the lack of activity with Oct3 suggests a greater substrate selectivity in this protein that can

discriminate based on the acyl moiety. Despite the apparent promiscuity displayed by Nag3 this does not extend to hydrolysis of Ac-Glu or Ac-Asn which would indicate that both the amide group of the glutamine side chain and its length form an important interaction for catalysis to occur, presumably in substrate binding and positioning. The fact that N-AGA is only able to hydrolyse Ac-Gln would support previous biochemical characterisation of this enzyme where it appears to show a high selectivity for only glutamine containing molecules (Natsch and Emter, 2020; Natsch et al., 2003).

The only other acetylated substrate which could be hydrolysed was Ac-Ala in the presence of Nag5, which like Oct4 did display some activity with 3M2H-Gln, hinting that this protein's natural substrates may be acyl alanine derivatives. All other substrates screened showed negligible activity with any of the proteins suggesting that those amino acids are not found in any of the physiological substrates and functions of the aminoacylases. However, this screen only covers acetylated amino acids and doesn't investigate longer chain acyl groups which could have a profound impact upon substrate affinity and selectivity. Furthermore M20 members have been noted to display carboxypeptidase and aminopeptidase activity as well as aminoacylase activity (Jamdar, 2009; Jamdar et al., 2015) which again has not been explored in this particular screen. Both Oct1 and Oct2 displayed no activity with any of the acetylated substrates tested and no glutamine release was detected upon incubation with the VFA precursors which could suggest that two proteins have been purified as inactive or they function on different ligands to those tested here.

The results of this activity screen have given further evidence to support the hypothesis that Nag3 and, to a lesser extent, N-AGA are more promiscuous enzymes while Oct3 has a much more selective substrate range.

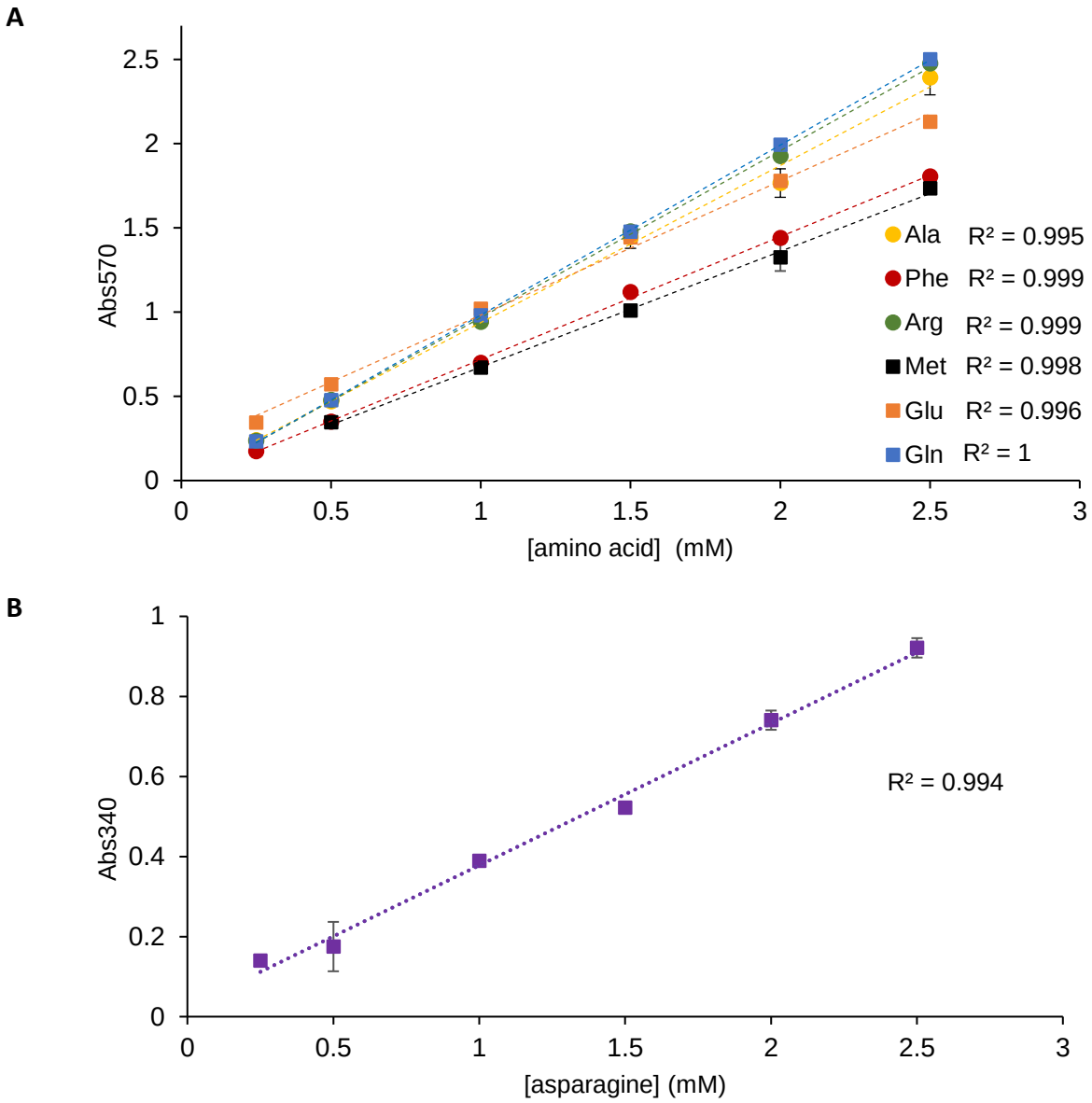


Figure 4.2.1: Amino acid standard calibration curves

A. Calibration curve of known concentrations of free L-amino acids generated using ninhydrin assay microplate format measuring absorbance at 570 nm. A line of best fit was calculated for each amino acid and used to quantify in subsequent experiments. Error bars denote standard deviation (n=3).
B. Calibration curve of known concentration of free L-asparagine generated using ninhydrin assay microplate format measuring absorbance at 570 nm. A line of best fit was calculated and used to quantify asparagine in subsequent experiments. Error bars denote standard deviation (n=3).

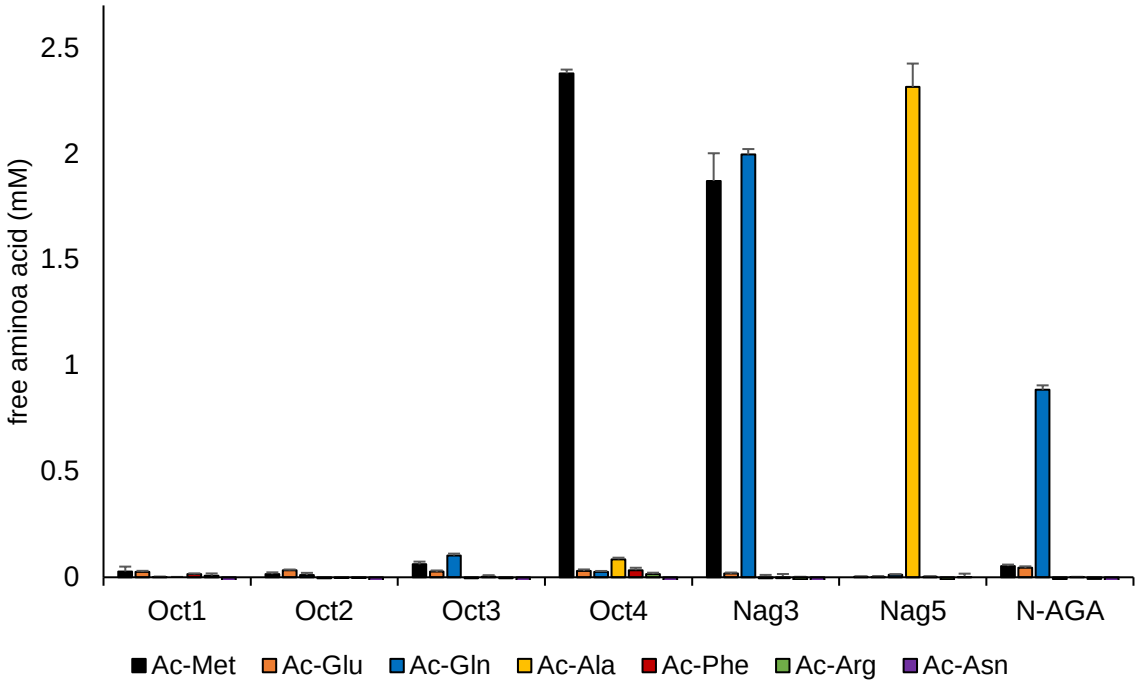


Figure 4.2.2: Acetylated amino acid activity screen

Respective amino acid product quantified via ninhydrin assay after 1 h incubation (37 °C) of 1.85 µM enzyme with 2.5 mM N^{α}-acetylated amino acid substrate. Error bars denote standard deviation (n=3).

4.3 Michaelis-Menten kinetics determination of active VFA liberating aminoacylases

The previous activity screen with the VFA substrates (**section 3.10**) clearly showed the selectivity of Oct3 for 3M2H-Gln while also displaying lower activity than Nag3 and N-AGA with this substrate. To investigate the different substrate selectivity displayed by Oct3, Nag3 and N-AGA for the precursors in greater detail the Michaelis-Menten kinetic parameters were determined for each protein.

Enzyme reactions at six different substrate concentrations were set up and a time course reaction initiated with enzyme addition. Samples of the reaction at time points (typically 0, 5, 10 and 15 min unless otherwise stated) were taken and quenched with EDTA before all glutamine was quantified by ninhydrin assay. The initial rates of reaction at each substrate concentration were calculated from the linear portion of the time course reaction and plotted against substrate concentration (Fig 4.3.1). A hyperbolic curve was fitted to the data points using nonlinear regression analysis in GraphPad Prism 9 and used to calculate the K_M and k_{cat} values for each enzyme with 3M2H-Gln and HMHA-Gln. Due to the very low rate of HMHA-Gln hydrolysis by Oct3 an accurate hyperbolic curve could not be fitted to the data points generated even at a higher substrate range (2 – 10 mM) and longer time points (0, 10, 20, 30 min). Further increases in substrate concentration were not attempted due to loss of pH buffering above 10 mM which would alter enzyme activity.

The kinetic experiments showed that all three enzymes have similar levels of affinity to 3M2H-Gln with K_M values ranging from 1.78 – 1.95 mM (Table 4.3.1) indicating that none of these enzymes display very high affinity for the VFA precursors. For Nag3 this affinity is even lower for HMHA-Gln with a rise of K_M to almost 3 mM, supporting the previous hypothesis that 3M2H-Gln is the preferred substrate. However, N-AGA displays a greater affinity for HMHA-Gln compared to 3M2H-Gln with K_M values of 1.65 mM and 1.95 mM respectively. The observation of better affinity with HMHA-Gln compared to 3M2H-Gln doesn't match previous *in vitro* kinetic work performed on N-AGA (Natsch et al., 2003) where K_M values of 0.2 mM for 3M2H-Gln and 0.74 mM for HMHA-Gln were calculated. The values from Natsch *et al* would indicate an even greater level of affinity than that observed here but differences in purification and assay conditions could be causative of this variation. The method used by Natsch *et al* also utilised a single time point after 15 min for kinetic determination which could have likely affected calculation of the initial rate of reaction.

Calculation of catalytic efficiency (k_{cat}/K_M) would suggest that N-AGA is the most efficient with both substrates with values a magnitude above those determined for Nag3 and Oct3. Of the Anaerococcus

enzymes, Nag3 displayed a better catalytic efficiency than Oct3 with 3M2H-Gln, again indicating that the more promiscuous enzyme also has better activity with both substrates.

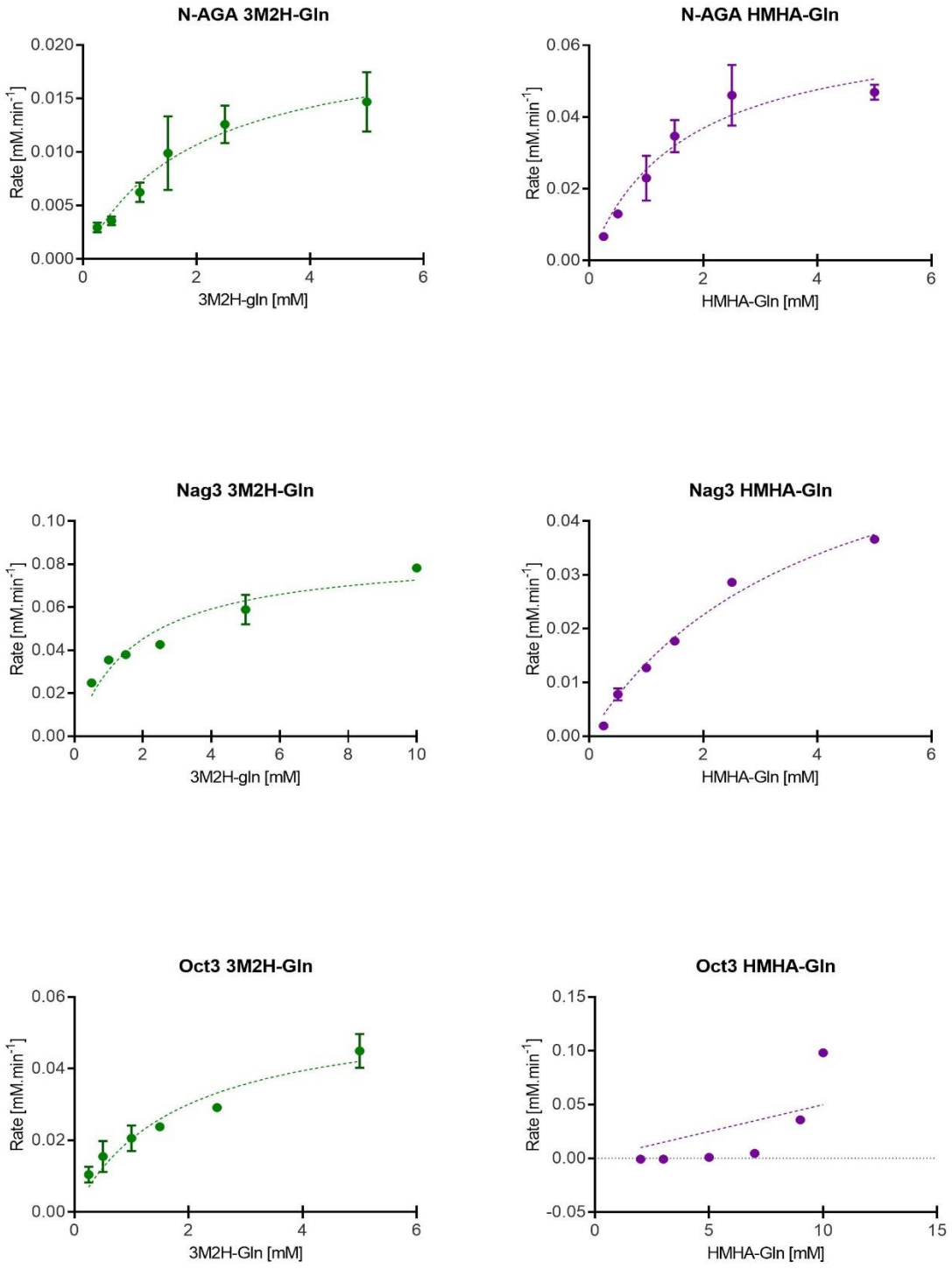


Figure 4.3.1: Michaelis-Menten kinetics with VFA precursors

Initial rates were calculated from the initial linear rate of formation of Gln from time-course reactions of enzyme and six substrate concentrations as indicated on plots. Dots represent experimental values from triplicate replicates. Lines are Michaelis-Menten fit using nonlinear regression analysis in GraphPad Prism 9.

Table 4.3.1: Michaelis-Menten kinetics with VFA precursors

	3M2H-Gln				HMHA-Gln			
Enzyme	**[E] (µM)**	**K_M (mM)**	**k_{cat} (min^{-1})**	**k_{cat}/K_M (min^{-1} mM^{-1})**	**[E] (µM)**	**K_M (mM)**	**k_{cat} (min^{-1})**	**k_{cat}/K_M (min^{-1} mM^{-1})**
Nag3	1.85	1.78 ± 0.15	46.2 ± 2.6	25.96	1.85	2.94 ± 0.22	36.4 ± 0.76	12.4
Oct3	1.85	1.82 ± 0.36	31.0 ± 0.81	17.03	-	-	-	-
N-AGA	0.05	1.95 ± 0.38	373 ± 23.3	191.3	0.1	1.65 ± 0.56	673 ± 103	407.9

4.4 Oligomeric state of Oct3

As previously noted during purification, all proteins displayed either two distinct or a single elongated protein peaks during SEC which was likely caused by the separation of two oligomeric states (Fig 3.3.3 and 3.7.2). M20 family members typically assume a homodimer oligomeric state although monomers and even homotetramers have also been observed (Rawlings and Barrett, 2013; Natsch and Emter, 2020). To assess the oligomeric state of Oct3 size exclusion chromatography – multi angle laser light scattering (SEC-MALLS) was performed.

Purified Oct3 was loaded onto a Superdex 200 HR10/30 column and a laser light passed through the eluent with the resulting light scattering in the Rayleigh scattering region being measured by detectors. The light scattering is measured at multiple angles and the intensities collected used in conjunction with refractive index measurements of the sample to determine the molecular mass of the protein eluate in solution. The Oct3 sample was run twice at different protein concentrations (3 mg/ml and 0.5 mg/ml) to check for any concentration dependent effect on oligomerisation. A control protein sample of 2 mg/ml BSA was also used to ensure correct instrument calibration.

SEC-MALLS of Oct3 displayed a similar UV trace to that of previous SEC runs with the failure to resolve the widened protein peak of Oct3 into two distinct entities based on oligomeric state as was observed for the BSA control (Fig 4.4.1). The Rayleigh ratio at both 3 mg/ml and 0.5 mg/ml displayed variation across the time of protein elution, indicative of a mixture of two oligomeric species. Calculation of the molecular weight of Oct3 in solution using both these ratios gave a value of 89 kDa suggesting the protein predominantly forms a homodimer conformation given the actual monomer weight of 44.8 kDa. An initial peak with 3 mg/ml Oct3 was also observed in all 3 recorded variables and created a very large intensity of light scattering (with a similar reduced pattern also being observed at 0.5 mg/ml) which coupled with early elution time of this signal would suggest the presence of some very large aggregates present in the sample likely caused by protein instability and degradation.

The finding of Oct3 to be a homodimer in solution matches the most common conformation within the M20 family and suggests Oct3 shares the similar mechanism of action in requiring a dimer for catalytic activity to occur (Jamdar et al., 2015; Lindner et al., 2003). Given the similarity in size and primary sequence between Oct3 and Nag3 it is assumed that Nag3 forms the same conformation in solution.

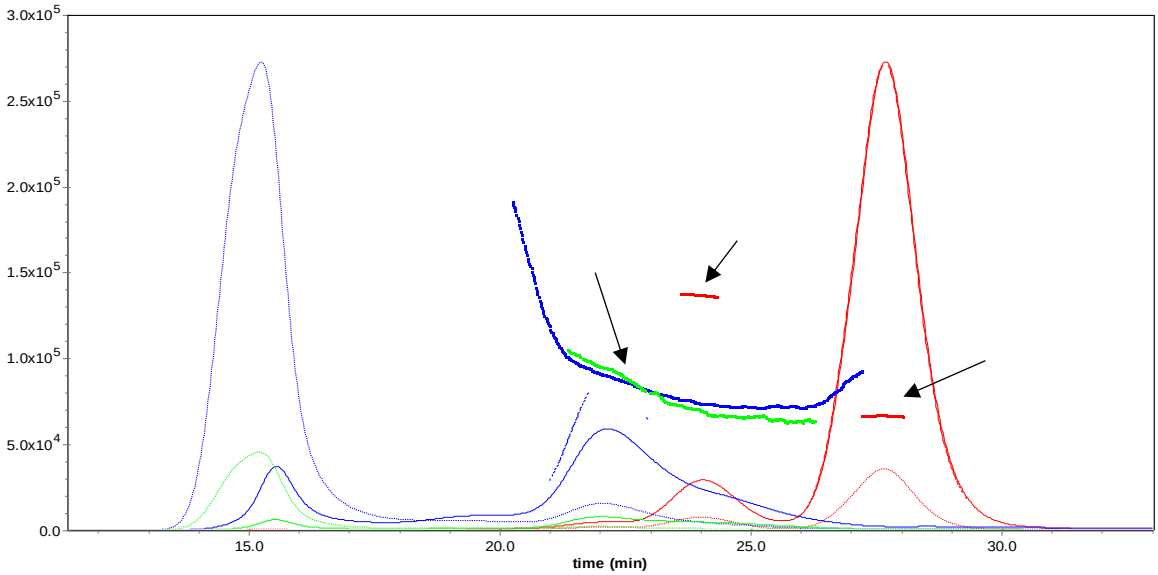


Figure 4.4.1: SEC-MALLS of Oct3

Oct3 was run at two different protein concentrations, 3 mg/ml and 0.5 mg/ml (denoted by the blue and green lines respectively). BSA (shown in red) was used as a calibration control. The solid lines represent the refractive index. The Rayleigh ratio is denoted as the dashed lines. The UV absorbance is shown as dotted lines. Estimated molecular mass of samples are indicated by black arrows. BSA dimer molecular mass = 130 kDa, BSA monomer = 66 kDa, Oct3 dimer and monomer mix = 89 kDa. SEC-MALLS experiment performed by Andrew Leech in the University of York Technology Facility.

4.5 Metal reconstitution assays

Members of the M20 aminoacylase family require divalent metal cations for activity, with this role most commonly being fulfilled by zinc ions, although examples of cobalt and manganese have also been recorded (Cao et al., 2020; Girish et al., 2013; Jamdar et al., 2015; Lin et al., 2007; Rawlings and Barrett, 2013). To determine the metal present in VFA liberating aminoacylases a metal reconstitution activity assay was attempted with Nag3.

Initial attempts involved pre-incubation of purified Nag3 with 10 mM EDTA to strip the metal cofactor before dialysis to remove the free and metal bound EDTA. This "metal free" Nag3 (henceforth referred to as mfNag3) was incubated with 1 mM of noted divalent metals and the amount of Gln liberated from a 1 h incubation with 2.5 mM 3M2H-Gln quantified by ninhydrin assay. This method failed to effectively remove the metal cofactor present in the protein as evidenced by a drop of roughly 20 % activity between the treated mfNag3 compared to untreated (Fig 4.5.1 A). Given that EDTA has been successfully used to quench previous activity reactions it would suggest that the inhibitory action of EDTA, in Nag3 at least, is not caused by complete stripping of the metal cofactor from the enzyme but instead the direct binding of the metals and EDTA within the active site. The formation of an inhibitory ternary complex has been well noted as a type of inhibition caused by metal chelating agents and various metalloenzymes (Auld, 1995; Ju et al., 2018). While this technique was unsuccessful in creating apo protein the resulting activity assay did show some effect of activity loss with some of the metals, most notably with copper, manganese and calcium which would suggest that at this concentration they likely have an inhibitory effect.

An on-column protein unfold refold procedure was used to try to fully strip the metal present. Protein samples of purified Nag3 were bound to an EDTA resistant HisTrap excel column and the protein unfolded with addition of 3 M guanidine HCl to a buffer also containing 25 mM EDTA. The unfolded and metal stripped protein was then refolded on column by buffer exchange to remove the denaturing agent and EDTA before the protein was eluted with imidazole as standard for IMAC purification. The refolded protein was then used in the same metal reconstitution assay performed on EDTA dialysed protein. Assaying of refolded Nag3 (rNag3) revealed the protein had been stripped of much of its metal cofactor as evidenced by the low activity of rNag3 containing no added metal (Fig 4.5.1 B). The low levels of activity still observed in this sample could be attributed to either a failure to remove all bound metal in the protein or perhaps more likely the presence of trace contaminants of divalent metals in buffers and labware. No activity was observed with the addition of zinc despite it being the presumed metal cofactor due to its prevalence in other M20 aminoacylases and being the stated cofactor for N-

AGA in both its biochemical characterisation (Natsch et al., 2003) and structural characterisation (Natsch and Emter, 2020). Instead, significant activity was only observed with the addition of cobalt which has been shown to enable activity in other M20D proteins (Cao et al., 2020; Jamdar et al., 2015; Lin et al., 2007). The activity level of rNag3 + 1 mM Co does not match that of native Nag3 which could be due to the concentration of metal used or that the refolding procedure resulted in only a fraction of active native protein correctly folding back into an active conformation.

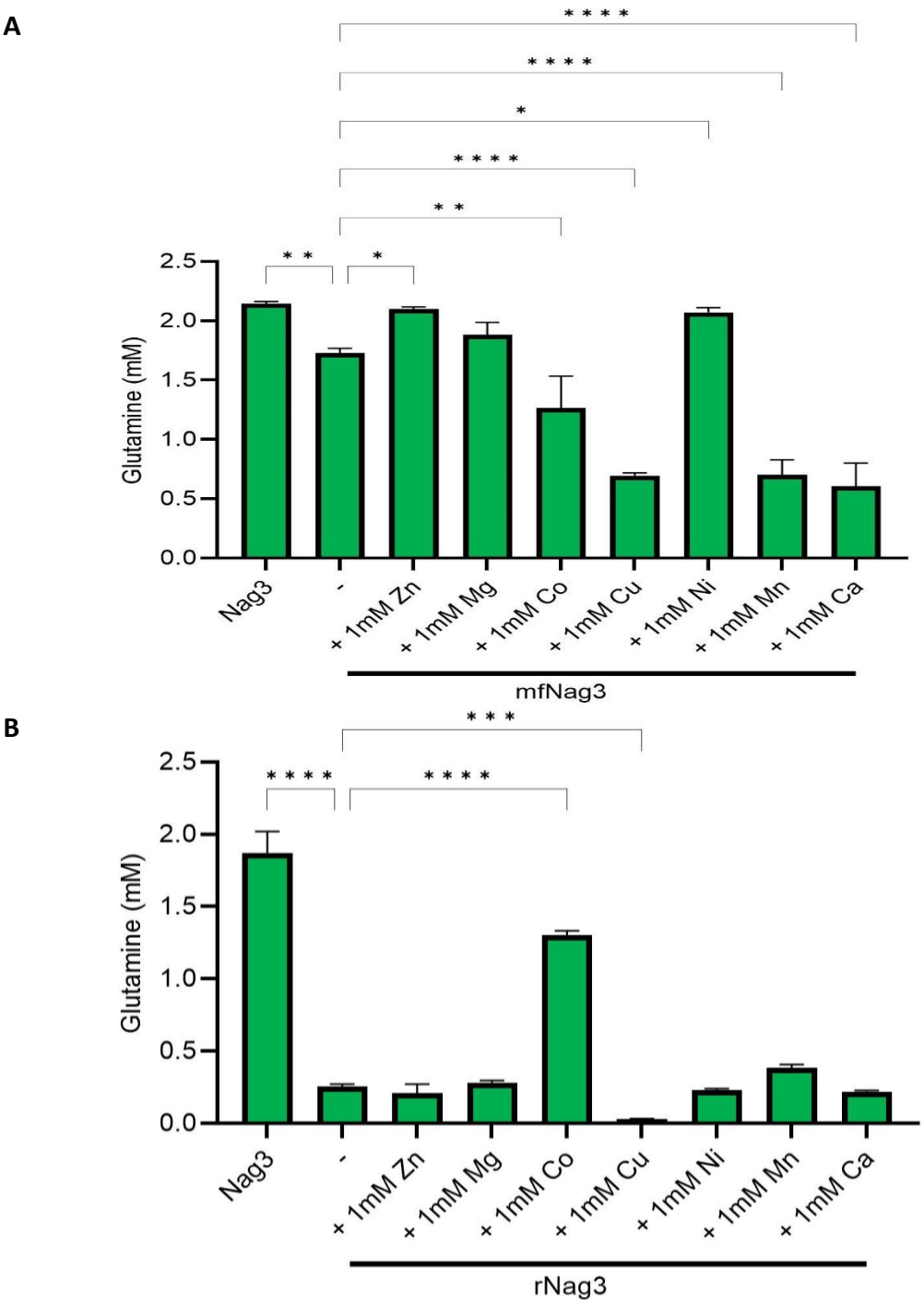


Figure 4.5.1: Nag3 metal reconstitution assays

A. Ninhydrin activity assay performed on incubations of 1.85 µM native or metal free Nag3 (mfNag3) with 2.5 mM 3M2H-Gln for 1 h, 37 °C. Samples were amended with stated concentration of metal before reaction initiation upon enzyme addition. – denotes reaction of mfNag3 and 2.5 mM 3M2H-Gln without any metal added. Error bars denote standard deviation (n=3). Asterisks show statistical significance from one way ANOVA compared to mfNAg3 without metal.

B. Ninhydrin activity assay performed on incubations of 1.85 µM native or refolded Nag3 (rNag3) with 2.5 mM 3M2H-Gln for 1 h, 37 °C. Samples were amended with noted concentrations of metal before reaction initiation upon enzyme addition. – denotes reactions of rNag3 and 3M2H-Gln without any added metal. Error bars denote standard deviation (n=3). Asterisks show statistical significance from one way ANOVA compared to rNAg3 without metal.

The surprising result of zinc not restoring activity in apo Nag3 led to a repeat of this experiment for N-AGA as a control due to its previous metal characterisation of zinc in the literature (Natsch and Emter, 2020; Natsch et al., 2003). Purified N-AGA underwent the same unfold refold procedure as Nag3 but the subsequent metal reconstitution assay also included an experiment at 50 µM of each metal as well as 1 mM to test for metal concentration dependent effects (Fig 4.5.2). At both metal concentrations the refolded N-AGA (rN-AGA) without any added metal displays low but not complete abolishment of activity, similar to Nag3, suggesting that not all the native metal has been removed. The two metal concentration assays were performed on separate protein stocks that had undergone the unfold refold procedure, indicating that this is the extent of metal stripping attainable by the protocol used. In the presence of 1 mM metals, a complete abolishment of activity is observed in the addition of zinc seemingly contradicting previous literature where zinc is the believed native metal bound in N-AGA (Natsch and Emter, 2020; Natsch et al., 2003). This inhibition however is not observed at 50 µM Zn which shows increased activity to that of rN-AGA (although not as high as other samples or native N-AGA), indicating that high zinc concentrations appear to have an inhibitory effect. At 1 mM concentrations of metal Co, Mn and Mg facilitate high activity with Co and Mn being higher even than native N-AGA indicating that multiple divalent metals can facilitate and even enhance activity. Even at 50 µM metal both CO and Mn show higher activity than that displayed by Zn making it difficult to conclude which is the native metal and even raising the possibility that the protein may contain multiple different metal ions making them cambialistic.

With the observation of a zinc concentration dependent effect for N-AGA, an activity assay with 3M2H-Gln and various concentrations of zinc was performed with both native Nag3 and N-AGA to investigate if this effect can be repeated in Nag3 (Fig 4.5.3). Zinc inhibition was recorded for both aminoacylases when Zn ≥ 1 mM suggesting that Nag3 is affected in a similar manner to N-AGA by high zinc concentrations.

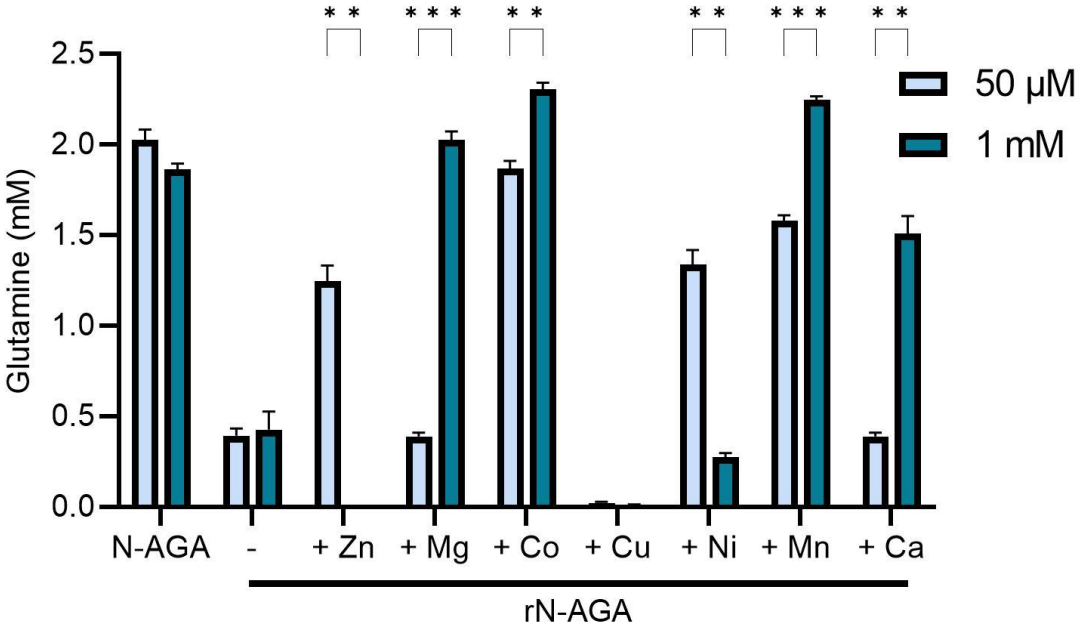


Figure 4.5.2: N-AGA metal reconstitution assay

Ninhydrin activity assay performed on incubations of 1.85 µM native or refolded N-AGA (rN-AGA) with 2.5 mM 3M2H-Gln for 1 h, 37 °C. Samples were amended with noted concentration of metal where applicable before reaction initiation with enzyme addition. Error bars denote standard deviation (n=3). Asterisks show statistical significance from multiple unpaired t-test between metal concentration.

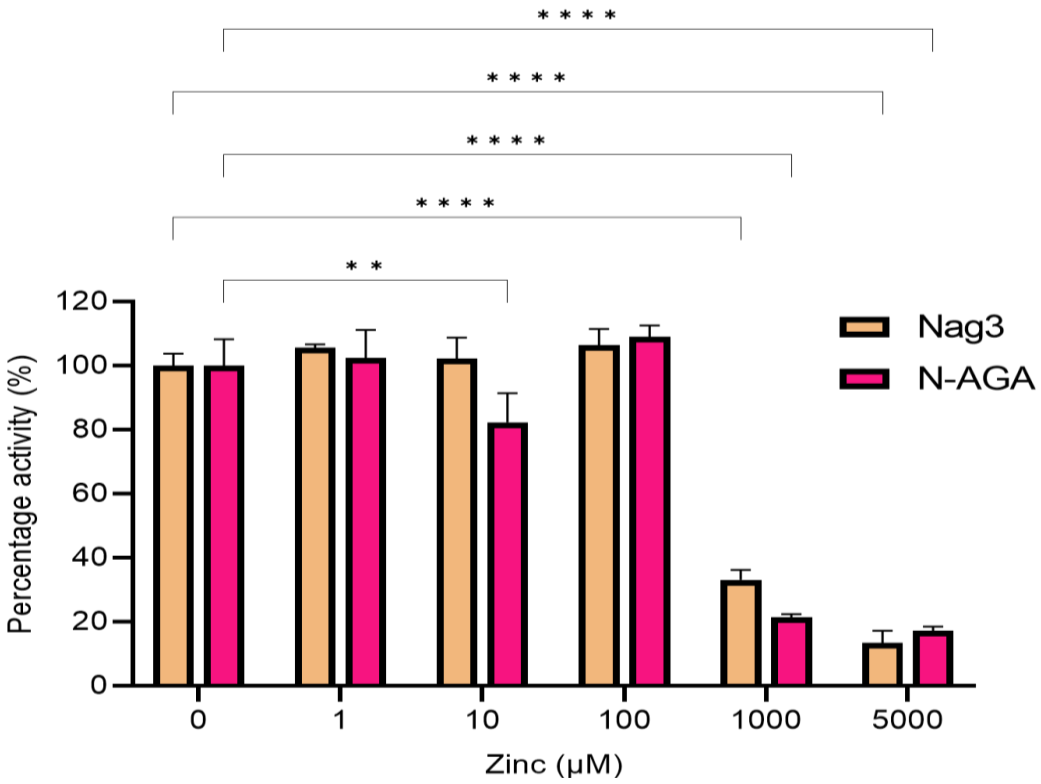


Figure 4.5.3: Nag3 and N-AGA activity assay varying zinc concentration

Ninhydrin activity assay performed on incubations of 1.85 µM Nag3 and N-AGA with 2.5 mM 3M2H-Gln for 1 h, 37 °C. Samples contained stated concentrations of zinc throughout reaction. Data shown as percentage activity using 0 µM zinc for each protein as 100% activity. Error bars denote standard deviation (n=3). Asterisks denote statistical significance from twoway ANOVA compared to 0 µM Zn for each protein.

4.6 Circular dichroism of refolded VFA liberating aminoacylase proteins

Whilst it can be inferred from the activity assays that protein used in the unfold refold procedure was able to refold back into its active conformation, additional evidence of protein refolding was acquired using circular dichroism (CD). The use of CD to look at the secondary structure of refolded protein compared to active native protein and denatured samples should provide evidence as to correct refolding if refolded and native protein CD spectra match.

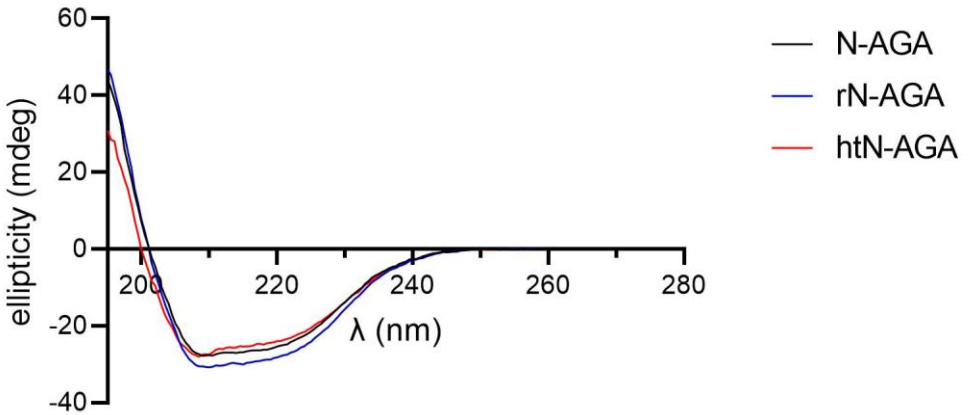
Data was collected for both N-AGA and Nag3 native and refolded proteins as well as an additional heat denatured sample for each. Purified protein samples were diluted to 0.2 mg/ml into 10 mM NaH_2PO_4 pH 7.4 buffer before CD spectra obtained in a 1 mm path length quartz cuvette in a J-1500 CD spectrophotometer. Measurements of ellipticity were obtained across a 185 to 260 nm wavelength range alongside measures of HT voltage. Analysis of collected data was limited to wavelength generating HT values less than 600 V due to inaccuracy in detection above this value. Estimation of protein secondary structure was performed using BeStSel (Beta Structure Selection) (Micsonai et al., 2018).

CD spectra for N-AGA samples showed agreement between rN-AGA and native N-AGA throughout the range of 260 to 195 nm (Fig 4.6.1). Analysis of CD spectra data using BeStSel revealed a similar estimation of secondary structure content for native N-AGA (BeStSel CD: 35% alpha helix, 27.6% beta strand, 36% other) to that observed in the PDBsum entry for the crystal structure of N-AGA (PDB: 36.2% alpha helix, 22.9% beta strand, 37.6 % other). When analysing rN-AGA data the estimated secondary structure does retain similarity (39.3% alpha helix, 16.7% beta strand, 39% other) giving a good indication that rN-AGA is correctly refolding into an active state. The heat treated N-AGA (htN-AGA) produces a relatively similar spectra to that of N-AGA and rN-AGA and secondary structure content analysis would suggest similarity in composition (32.6 % alpha helix, 24.3 % beta strand, 42.8 % other). As this denaturation was a limited time exposure to high temperature (95 °C, 10 min), it could be likely that N-AGA has assumed an inactive but folded state comprising of alpha helices after cooling.

CD spectra were obtained for Nag3 samples across the same wavelength range of 195 – 260 nm although greater variation between native, refolded and heat treated Nag3 was observed compared to N-AGA. Secondary structure analysis using BeStSel for native Nag3 estimated a composition of 28.4 % alpha helix, 34.2 % beta strands and 37.3 % other which did differ to that obtained for rNag3 with estimations of 34.9 %, 25.1 % and 41 % for each respective category. Estimations for htNag3 were 11.1 % alpha helix, 44.1 % beta strand and 44.7 % other which does differ to a greater extent to native and

refolded Nag3 than they do to each other, implying rNag3 is more likely to have assumed a similar active fold to that of native purified protein. The observation of rNag3 activity with 1 mM Co does also indicate active protein must be present in the sample (Fig 4.6.1 B).

A



B

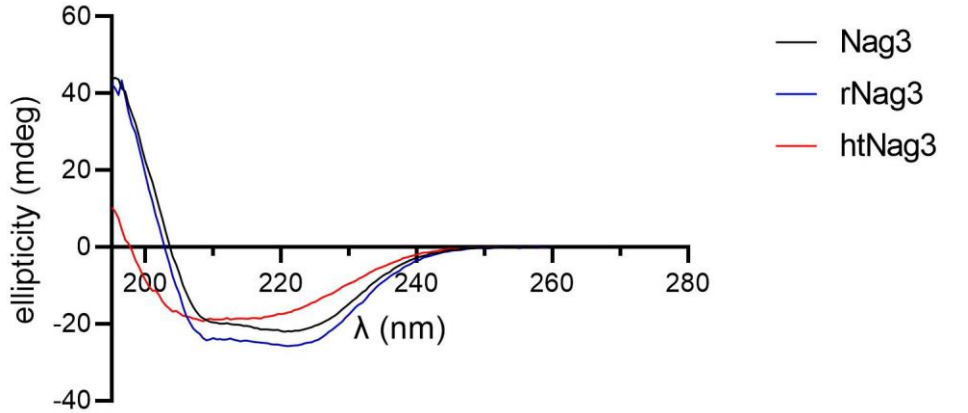


Figure 4.6.1: Circular dichroism of refolded aminoacylases

Measures of ellipticity (mdeg) were obtained from the wavelengths 185 to 260nm. Data displayed has been cut to appropriate points for respective proteins due to HT voltage values > 600 V for lower wavelength measurements impacting result accuracy. Both proteins were analysed in native active state (black), refolded protein (blue) and heat-treated protein (red). All data here has been zeroed against a 10 mM sodium phosphate buffer blank. CD spectra were collected for: **A.** N-AGA protein, refolded N-AGA (rN-AGA) and heat-treated N-AGA (htN-AGA). **B.** Nag3 protein, refolded Nag3 (rNag3) and heat-treated Nag3 (htNag3).

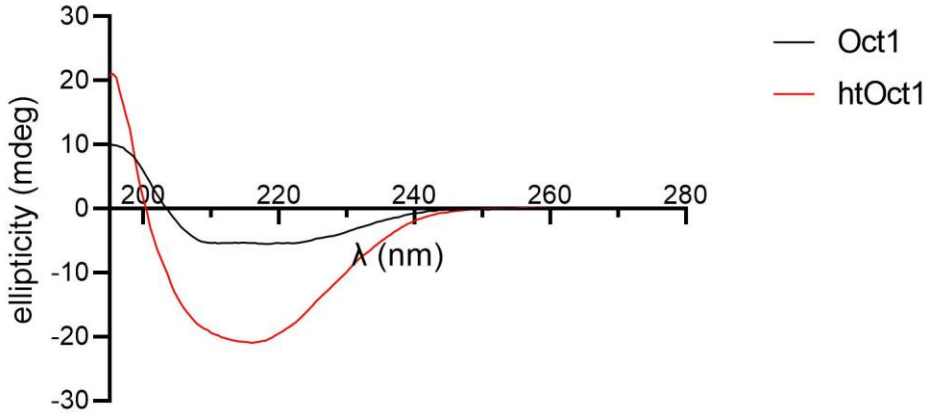
4.7 Assessing folded state of Oct1 and Oct2

Circular dichroism was also performed for Oct1 and Oct2 to verify if the purified proteins were in a folded state as no activity was observed in any activity assay performed. A comparison of CD spectra of native protein to that of heat denatured should show clear differences between the two indicating a transition from a correctly folded state to a misfolded one. Evidence of this fold change would suggest that both proteins were purified in their active, folded states and their inactivity in all previous biochemical activity screens is not due to catalytic inactivity.

CD spectra were obtained for both proteins in the same manner as detailed for Nag3 and N-AGA, for native and heat denatured protein samples (95 °C, 10 min) (Fig 4.7.1). The spectra collected for each protein showed clear differences upon heat treatment to that collected for purified protein, suggesting that after heat denaturation the protein has misfolded into a different folded state than originally. However, the data collection for Oct1 appeared noticeably different to all other aminoacylases, suggesting the protein is not folded similarly and could be in an unactive fold. Estimation of secondary structure content with BeStSel for native Oct1 highlighted this difference with an estimation of only 3. 9 % alpha helix structures, 52.1 % beta strand and 44.1 % other. The stark difference in alpha helixes compared to Nag3 and N-AGA (28.4 % and 35 % respectively) would suggest the protein tested is not folded properly (assuming aminoacylase homologues share structural similarity). Oct2 produced a more comparable spectra to Nag3 and N-AGA, extending to secondary structure estimation (30.7 % alpha helix, 30.6 % beta strand and 38.7 % other). Heat denaturation showed a clear change in spectra and secondary structure estimation with BeStSel (12.3 % alpha helix, 39.5 % beat strand and 48.4 % other), indicating that purified Oct2 is in an active folded state.

The CD spectra analysis of these two proteins would suggest that Oct2 is purified in an active folded state and its observed inactivity is due to failure to screen with a tolerated substrate or absence of cofactor. Oct1 appears to not be in a similar folded state as all other tested aminoacylases suggesting that this protein has likely appeared inactive in all screening due to catalytic inactivity.

A



B

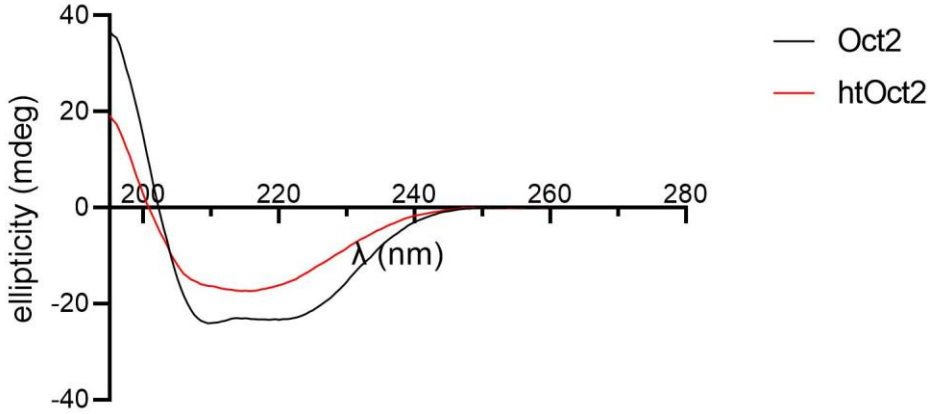


Figure 4.7.1: Circular dichroism of Oct1 and Oct2

Measures of ellipticity (mdeg) were obtained from the wavelengths 185 to 260nm. Data displayed has been cut to appropriate points for respective proteins due to HT voltage values > 600 V for lower wavelength measurements impacting result accuracy. Both proteins were analysed in native state (black) and heat-treated protein (red). All data here has been zeroed against a 10 mM sodium phosphate buffer blank. CD spectra were collected for: **A.** Oct1 protein and heat-treated Oct1 (htOct1). **B.** Oct2 protein and heat-treated Oct2 (htOct2).

4.8 Analysis of metal content of N-AGA structure

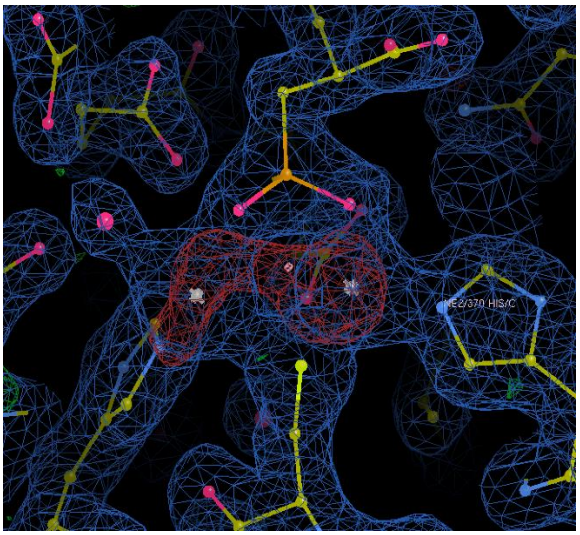
The results of the metal reconstitution activity assay for N-AGA show some difference to that of previous biochemical evidence seen in the literature (Natsch et al., 2003), making it difficult to ascertain the native metal present. To try resolve this difference the published structure of N-AGA (PDB code: 6SLF) was studied to check if the published structural data could give a more definitive answer on the native metal ion.

The structure contains two modelled Zn atoms within the binding site and at the correct positions to be coordinated by the five metal interacting residues that are conserved in all M20D family aminoacylases. However, upon observing the electron density maps a clear difference in the modelled structure density to the experimental density can be seen when looking at the Fo-Fc map, with negative contours around the modelled Zn atoms (Fig 4.8.1). This difference in density suggests that the modelled Zn atoms, which are present in each of the protomers within the homotetramer asymmetric unit, have been incorrectly placed within the structure. The crystallisation conditions published for Ax20 do not include any additional metal added suggesting the metal has not been exchanged during crystallisation.

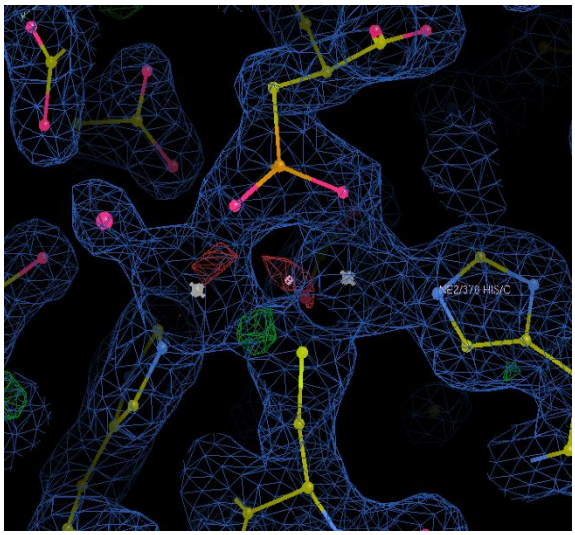
As the density maps show Zn as being incorrectly placed within the structure, the zinc atoms were replaced with different divalent metals that were used in the metal reconstitution assays to see if any would better fit the experimental density (Fig 4.8.1). The Zn atoms were substituted with other metal ions using Coot before running through REFMAC5 and visualising the structure and density maps within Coot again. Both Co and Mn provide a much better fit to the density than previously observed with Zn which could suggest that one of these metals is present in the crystal structure. These metals also showed high activity with N-AGA in metal reconstitution assays further supporting this structural analysis. Calcium was also modelled but appears too small to fit the density as evidenced by the Fo-Fc(+ve) σ green contours suggesting that it is not the present metal and ruling out any smaller metal ions such as magnesium. An evaluation of B factors of the modelled metal ions and nearby atoms was also performed to corroborate these finding (Table 4.8.1). A large difference between the B-factors of nearby atoms can be indicative of misplaced or incorrect metal atom. For each metal modelled, four nearby atoms were also analysed to allow comparison both to the specific metal in that structure and as a control check between each model to ensure that refinement had not significantly altered other atoms within the structure. The collected B-factors supports the idea that Zn is not the metal present in the crystal structure of N-AGA, with Co and Mn being more consistent to nearby residues B-factors and Ca showing a clear difference to its respective nearby residues. This data is suggestive that Zn is

not the metal cofactor present in the published Ax20 structure. However, analysis of interaction distances between the present zincs and coordinating residues shows an average distance of 2 Å between both His N atoms and glutamate O atoms with Zn^{2+} ions which is within the average interaction distances seen in zinc metalloenzymes (Tamames et al., 2007). Metal binding residues appear positioned appropriately to facilitate zinc coordination making it difficult to accurately conclude the exact metal present in the structure.

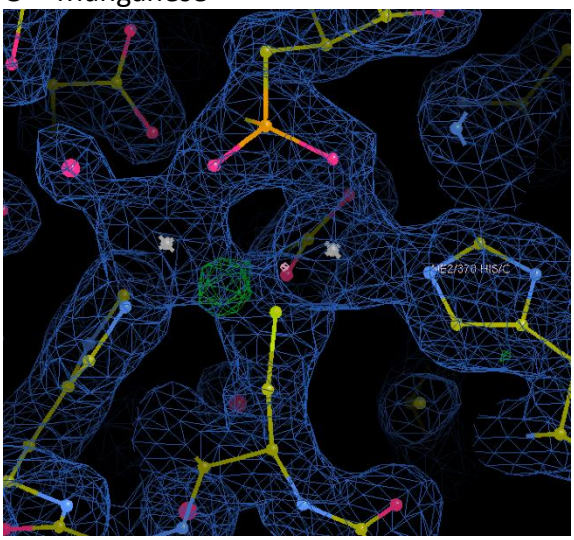
A Zinc



B Cobalt



C Manganese



D Calcium

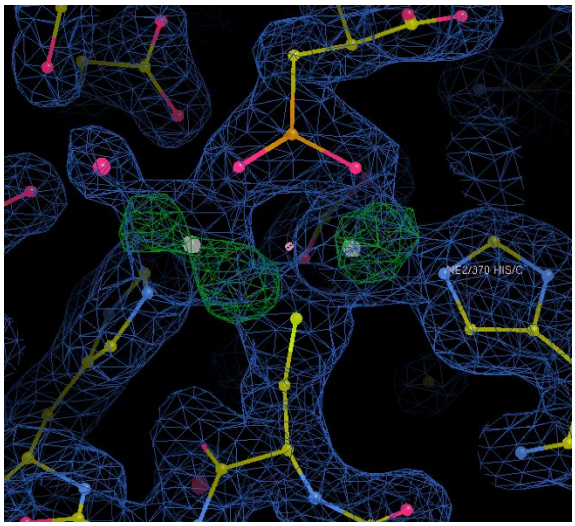


Figure 4.8.1: Electron density map of N-AGA with varied metal cofactors

A view into a single monomer metal binding site within the Ax20 structure. The two modelled metal atoms present in the structure are denoted as grey spheres. The Fo-Fc(-ve) σ map is shown by a red wireframe. Fo-Fc(+ve) σ is shown as a green wireframe (Fo-Fc maps contoured to 3 Å RMSD). 2Fo-Fc σ map is shown as a blue wireframe (contoured to 1.5 Å RMSD). The N-AGA structure (PDB 6SLF) was visualised in Coot.

Table 4.8.1: B-factors of modelled metal atoms in N-AGA

Zn		**Co**		**Mn**		**Ca**	
O Glu141	17.54	O Glu141	17.62	O Glu141	17.71	O Glu141	17.98
O LJ8	17.41	O LJ8	17.6	O LJ8	17.71	O LJ8	18
N His370	16.26	N His370	16.34	N His370	16.48	N His370	16.91
S Cys105	16.78	S Cys105	16.75	S Cys105	16.86	S Cys105	17.13
Zn	23.16	Co	20.83	Mn	18.79	Ca	13.5

4.9 Discussion

With the identification of Oct3 and Nag3 as VFA liberating aminoacylases, efforts detailed in this chapter focused on the further characterisation of these two enzymes biochemically and biophysically. The use of a ninhydrin assay for biochemical characterisation enabled the detection of a variety of amino acids with the same assay, something which would not have been possible with a more specific assay such as the commercial glutamine detection assay first employed. The high throughput provided by the ninhydrin assay also enabled screening of all purified proteins for aminoacylase activity with a range of acetylated amino acid substrates. While aminoacylases have been often noted to have a wide substrate range (as discussed in section 3.11), this was not observed during this screening indicating substrate affinity is highly influenced by the amino acid moiety. Despite Ac-Met being commonly used as a model substrate and showing activity with Nag3, only Oct4 was also able to catalyse methionine release. The stark differences between Ac-Met and the VFA precursors raises interesting questions as to how this promiscuous activity in Nag3 is achieved, especially coupled to the fact that the only other hydrolysed substrate was Ac-Gln indicating a level of substrate selectivity is maintained. Examples of activity with both of these substrates has been observed in other enzymes but they also displayed activity with a variety of other substrates (Jamdar et al., 2015; Koreishi et al., 2005, 2009). In the case of three M20D aminoacylases BcepM20D, DradM20D and HmrA (from *Burkholderia cepacia, Deinococcus radiodurans* and *Staphylococcus aureus* respectively), activity was low for both substrates with relative activities ranging from 15-30 % for Ac-Met and 10-40 % for Ac-Gln. This level of activity was also exhibited for numerous different acetylated amino acids tested with each enzyme displaying noticeable activity with at least seven different substrates. In comparison Nag3 has a much more restrictive substrate range that even lacks activity with Ac-Ala, which due to its small size and hydrophobic side chain could have been hypothesised to be a likely universal substrate for most aminoacylases. In fact, Ac-Ala was observed to have either the highest or second-best activity of the acetylated amino acid substrates tested for BcepM20D, DradM20D and HmrA (Jamdar et al., 2015), evidencing its potential as a model substrate for promiscuous aminoacylases. Given this it would suggest that while Nag3 is more promiscuous than Oct3 and N-AGA it does not appear to be as promiscuous as many other aminoacylases and displays clear substrate specificity. Admittedly this statement must be made with the awareness that the biotechnological interest in aminoacylases could lead to a bias in published literature to focus into promiscuous enzymes that could be used with various acylated substrates in commercial processes. The activity recorded with Ac-Met for Nag3 could therefore be seen as a surprising, tolerated substrate rather than an indication of a broad substrate range. The lack of activity of Oct3 with Ac-Gln or any other acetylated amino acids does match the

hypothesis postulated after VFA precursor screening that this enzyme is more selective in substrate activity than Nag3 or N-AGA with it appearing that specificity is conferred by both the amino acid and acyl moiety in this enzyme. The observation of activity for N-AGA only with Ac-Gln does match previous literature that noted activity for that enzyme only for acylated glutamine molecules and not any other amino acid conjugates (Natsch and Emter, 2020; Natsch et al., 2003). That characterisation of N-AGA did also note that increased substrate affinity did correlate to increasing acyl chain length which again matches screening here with Ac-Gln activity appearing lower than that achieved with 3M2H-Gln in chapter 3.

While the other proteins are not the major focus of this work the high activity Oct4 and Nag5 displayed for Ac-Met and Ac-Ala, respectively, could hint at their physiological substrates and role. Oct4 was able to hydrolyse almost 100 % of the Ac-Met substrate after 1 h, suggesting its native substrate is likely methionine containing. Methionine plays a key role in cell growth as a source of sulphur and generation of the vital methyl and amine donor molecule S-adenosylmethionine (SAM), a conserved process from bacteria to eukaryotes (Berger et al., 2003; Chiang et al., 1996; Parkhitko et al., 2019). Hence methionine scavenging and recycling is an important process for microorganisms to enable cell function and would place an emphasis on maximising methionine utilisation. With *A. octavius* being present in the axilla microbiome an obvious methionine source can be proposed in human derived peptide and amino acid uptake. Eukaryotic proteins are typically N-terminally acetylated with this residue commonly being the initial methionine residue unless cleaved (Jörnvall, 1975). As such it could be postulated that Oct4 could act upon Ac-Met scavenged from sweat containing degraded protein and peptides. Evidence of bacterial protein N-terminal acetylation has also been observed (after the necessary deformylation), raising the possibility that host protein turnover could also be a source for *N*-acetylated methionine (Brown et al., 2017; Ouidir et al., 2015). The high activity of Nag5 with Ac-Ala would propose that natural substrates are acylated alanine molecules. In a similar speculation to that for Ac-Met, scavenging of human derived N-terminal acetylated peptides and amino acids could be a likely source of Ac-Ala due to its propensity (alongside other small hydrophobic residues) to become the N-terminal residue in the case of methionine cleavage, as dictated by the protein sequence recognition of met-aminopeptidases (Frottin et al., 2006).

The acetylated amino acid screen does allow some exploration of substrate range and putative roles for the purified proteins, but the chosen substrate list is not extensive. Likewise no work was conducted with longer chain acylated amino acids (VFA precursors aside) or di/tri-peptides of which many aminoacylases, including N-AGA, have been noted to also have activity with (Ishikawa et al., 2001; Jamdar et al., 2015; Natsch et al., 2003). Hence it is likely that the full extent of substrate ranges each protein exhibits has not been explored. This is the most likely explanation for the absence of

activity observed for Oct2 throughout all activity screens given its observed folded state from CD spectra analysis. Although, it is interesting to note that both Oct4 and Nag5 did display activity with 3M2H-Gln, especially after 24 h, as well as with Ac-Met and Ac-Ala whereas both Oct1 and Oct2 have appeared inactive in all biochemical activity screening which could be suggestive that both are catalytically inactive.

Determination of the Michealis-Menten kinetics for Oct3, Nag3 and N-AGA with the VFA precursors revealed each had a similar level of affinity for 3M2H-Gln ranging from 1.78 – 1.95 mM. Kinetic analysis for HMHA-Gln was unsuccessful for Oct3 but showed a clear difference in affinity between N-AGA and Nag3 (1.65 and 2.94 mM respectively). These values combined with comparison of catalytic efficiency indicate N-AGA as the most efficient enzyme out of the three. It is therefore interesting to note that other *Corynebacteria,* including those containing close homologues, do not share this high activity (James et al., 2013b), with it instead appearing to be just restricted to the *C.* sp Ax20 species. Assuming that all axilla microbiome studies to date have not simply failed to isolate high other highly active VFA producing *Corynebacterium* it seems remarkable that this single isolate has adapted to so efficiently liberate VFAs. This species has also been shown to contribute to thioalcohol based malodour with the ability to generate 3M3SH (Bawdon et al., 2015; Emter and Natsch, 2008), making this strain appear to be a potent malodour generating strain. All affinity values determined for Oct3, Nag3 and N-AGA are at millimolar concentration which would imply the requirement of an efficient transport mechanism to enable sufficient intracellular substrate concentrating for significant enzymatic activity. Similar levels of affinity for acylated amino acid substrates have been observed in other aminoacylases (Jamdar et al., 2015; Lindner et al., 2003; Parker et al., 2011) suggesting that the VFA liberating aminoacylases display relatively normal levels of affinity for their substrates compared to other aminoacylases. The affinity values obtained for N-AGA are larger (3M2H-Gln K_M = 1.95 mM and HMHA-Gln K_M = 1.65 mM) than those noted for previous characterisation of this enzyme (3M2H-Gln K_M = 0.2 mM and HMHA-Gln K_M = 0.74 mM) (Natsch et al., 2003), but differences in protein purification as well as kinetic assay methodology could be the cause of these effects. Natsch *et al* used a fluorescamine assay to detect L-glutamine formation which reacts with the primary amine made available after substrate hydrolysis to form a fluorescent product that can be measured and has been used extensively for amino acid and peptide quantification (Derayea and Samir, 2020).

Biophysical characterisation of Oct3 showed it assumes a homodimer in solution matching the commonly observed oligomeric state in aminoacylases (Lindner et al., 2003; Rawlings and Barrett, 2013), with dimerization being shown to be key for catalytic activity (Lindner et al., 2003, 2005). While this was the presumed oligomeric state for all the *Anaerococcus* aminoacylases, the prevalence of homotetramer aminoacylases deposited on the PDB would suggest this is also a likely conformation

in solution as well. In fact evidence of N-AGA assuming a tetrameric state has been observed both in its published structure and also in solution upon addition of a high affinity competitive inhibitor (Natsch, 2015; Natsch and Emter, 2020). As the inhibitor mimics the transition state assumed by a VFA precursor molecule it was proposed that such an effect could also occur upon substrate binding. Large conformational changes upon substrate binding have been observed in the M20 family with DapE, a *N*-succinyl, L-diaminopimelic acid desuccinylase, from *Haemophilus influenzae* showing a ~10 Å shift in binding site position between product bound and apo structures (Nocek et al., 2018). This major shift was further shown to be key for catalysis with movement of a histidine residue (found on one monomer's dimerization domain) to interact with an oxygen atom of the succinic acid within the other monomer's binding site enabled activation of the scissile carbonyl bond. The His residue is conserved within N-AGA and is shown in the structure to be coordinating an oxygen atom of the phosphinic acid group of the inhibitor, a position that would be fulfilled by the carbonyl of 3M2H-Gln and HMHA-Gln. It is likely that N-AGA can undergo a similar level of conformational shift upon substrate binding and this movement enables a homotetramer arrangement to form with a dimerisation of dimers. Such an effect still raises the question as to the biological function of such a change in oligomeric state or whether it is a coincidental effect. It would also be interesting to note if a similar effect also occurs within the *Anaerococcus* aminoacylases.

Attempts to determine the metal cofactor of Nag3 with a metal reconstitution assay using EDTA dialysed protein failed to effectively remove the bound metal and still enabled protein activity even without the supplementation of metal ions. Due to the failure in the removal of metal the results of this assay were not studied in detail due to the unknown effect that the EDTA dialysis may have had upon the protein.

An on-column unfold refold procedure proved much more effective at removing the bound metal cofactor and revealed that the presumed metal cofactor, zinc, did not restore activity for rNag3, with only cobalt showing a significant increase. While zinc is the most commonly found metal present in M20 aminoacylases, there have been examples of other divalent ions bound such as cobalt (Cao et al., 2020), manganese (Girish and Gopal, 2010; Girish et al., 2013) and even nickel (Minasov et al., 2005). However, the examples of manganese observed in two M20 aminoacylases from *Staphylococcus aureus* COL0085 and Sapep were found in X-ray diffraction crystal structures of each protein that had been crystallised in the presence of manganese which has likely replaced any native metal present. In a similar trend, the nickel found in YxeP from *Bacillus subtilis* was observed from a crystal structure which did not have any added nickel in the crystallisation condition but the presence of a C-terminal His tag could imply that nickel may have been carried over from IMAC purification of the protein. Further experimentation with both Nag3 and N-AGA revealed that zinc appears to have a

concentration dependent inhibitory effect upon activity for both enzymes between 100 – 1000 µM zinc. High zinc concentrations have been observed to inhibit activity in other M20 aminoacylases such as SmAcy from *Stenotrophomonas maltophilia* (Cao et al., 2020) and BcepM20D from *B. cepacia* (Jamdar et al., 2015) and in the case of BcepM20D this was shown to be concentration dependent with zinc being determined as the native metal by inductively coupled plasma mass spectrometry (ICP-MS). Zinc concentration inhibition has been recorded for zinc metalloenzymes although typically for mononuclear zinc metallopeptidases, with zinc binding to an inhibitory site proximal to the catalytic zinc in the active site (Larsen and Auld, 1989, 1991). The mechanism of inhibition for dinuclear zinc metalloenzymes is not clear although evidence of additional zinc binding sites outside of the active site have been reported to inhibit activity in molecular dynamic experiments (Tomić et al., 2021).

Screening of the effect of metal ions upon rN-AGA revealed activity with various divalent metal ions. Several M20 aminoacylase have also been shown to exhibit activity in the presence of non-native metal ions such as BcepM20D, DradM20D, SmAcy, PhoAcy and carboxypeptidase Ss1 (Cao et al., 2017; Colombo et al., 1992; Jamdar et al., 2015; Occhipinti et al., 2003; Tanimoto et al., 2008) indicating that this is common feature of aminoacylase enzymes as well as being a common trait of metalloenzymes generally (Valdez et al., 2014). In the case of PhoAcy a switch in metal cofactor from the native zinc to manganese also led to a change in substrate specificity with acetylated amino acids raising the possibility that regulation of metal concentrations intracellularly could be used to regulate enzymatic activity and metabolism. This area was not further explored with the VFA liberating aminoacylases but could be an interesting area of further research to assess whether different metal cofactors affect substrate selectivity between the VFA precursors and other acetylated amino acids.

It was surprisingly that zinc had an inhibitory effect upon rN-AGA at a millimolar concentration which would appear to contradict previous N-AGA characterisation (Natsch et al., 2003). However, zinc inhibition detailed in this thesis was shown to be concentration dependent which was not explored by Natsch *et al.* (2003) who proposed the metal cofactor as zinc from treatment with EDTA oro-phenanthroline and subsequent restoration of activity with the addition of an "excess of Zn^{2+}". The lack of quantitative measurement of zinc makes it difficult to ascertain whether the inhibitory effect may have been missed with the amount of free zinc ions available. Furthermore, those experiment were performed on cell extracts of *C.* sp. Ax20 which would expose N-AGA to the entire cellular metal content, suggesting the possibility that other metal ions may be affecting activity. Natsch *et al* did also note partial activity restoration with the addition of Mn^{2+} which does match the results detailed here. This was the only other divalent ion mentioned as providing activity (although no list of tested metals was given) which could be attributed to the difference in *ex vivo* and *in vitro* environments these experiments were conducted in.

Use of a biochemical assay to study the effect of different metals has clearly illustrated the difficulty in being able to accurately ascertain the bound metal present in the protein using this method. Metal reconstitution assays with N-AGA suggests a flexible metal binding site that can accommodate, and utilise for catalytic activity, various divalent metals. Given this flexibility it must also be considered that the bound metal cofactor might not actually be that found *in vivo* within the host species. A clear example of this is peptide deformylase from *E. coli* which was initially observed to be bound to zinc *in vitro* before being shown to bind Fe^{2+} *in vivo* with evidence that during hydrogen peroxide stress this switches to Mn^{2+} to retain activity under oxidative stress (Anjem and Imlay, 2012; Jain et al., 2005). It therefore must be considered that not only might the metal bound be non-native but also that the metal found could be flexible depending upon host cellular conditions.

Analysis of the published structure for N-AGA, revealed that while two Zn^{2+} atoms have been placed in the structure this does not seem to match the electron density map of the structure. While structural analysis with potential cofactors could not accurately conclude which metal is present in the structure both manganese and cobalt created better fits to the data, and both also enabled activity with rN-AGA, which could suggest that zinc is not the currently bound metal present in the structure. The decision to place zinc within the structure despite these difference could have been due to knowledge of M20 aminoacylases utilising zinc as well as the previously mentioned biochemical characterisation of N-AGA metal cofactor (Natsch et al., 2003). The bond distances measured between the zinc atoms and interacting residues are also within the norm for zinc binding metalloenzymes (Tamames et al., 2007). Macromolecular crystallography can allow accurate determination of bound metals either through distinct anomalous signal during X-ray diffraction or through analysis of X-ray fluorescence (XRF) spectra. Unfortunately neither have been stated for N-AGA structural characterisation (Natsch and Emter, 2020), but this was explored during structural determination of Oct3 and Nag3E138A as detailed in chapter 5. ICP-MS, and the related ICP-optical emission spectroscopy (ICP-OES), is another technique that has also been used extensively in metalloenzymes to determine the bound metal, including for several M20 aminoacylases (Botelho et al., 2011; Jamdar et al., 2015; Lin et al., 2007; Tanimoto et al., 2008). Attempts to determine the bound metal of N-AGA, Oct3 and Nag3 using this technique were attempted using an external service but failed to yield any usable data due to high variation in metal concentration between replicate samples including the surprising detection of silver in some replicates despite each triplicate samples being technical repeats from the same protein stock (**data not shown)**. It was concluded that an experimental error was the likely cause of the variation and silver detection, and this work was not continued.

In summary biochemical characterisation of purified proteins revealed that Nag3 is the most promiscuous of the VFA liberating aminoacylases although this promiscuity does not extend as far as

other notable M20 aminoacylases that often display much broader substrate ranges. This would suggest that Oct3, Nag3 and N-AGA are displaying noticeable specificity to their VFA precursor substrates with kinetic experiments showing that N-AGA has the highest catalytic efficiency for both substrates. Failure to accurately calculate affinity values for Oct3 with HMHA-Gln, with the values likely to be > 7 mM, indicates differences must be present in the binding site compared to Nag3 and N-AGA to give such a clear absence in activity. SEC-MALLS concluded that Oct3 forms a homodimer in solution and is the likely same state assumed by Nag3 due to sequence similarity. Metal reconstitution assays indicated the difficulty in their use to accurately assign the native metal in the aminoacylases while also highlighting the flexibility in metal binding inherent to N-AGA and other M20 aminoacylases. Structural data can provide a more definitive means of metal determination, but study of the published N-AGA underlines the importance of analysing published structures rather than direct acceptance of the proposed structural model. While the N-AGA structure does appear to be not be bound to zinc this does not mean that this enzyme, and indeed Nag3 and Oct3, are not zinc metalloenzymes *in vivo*.

Chapter 5 – Mechanistic characterisation of Oct3 and Nag3

5 Mechanistic characterisation of Oct3 and Nag3

5.1 Introduction

To further the characterisation of the Oct3 and Nag3 detailed in chapter 4, work to investigate the specific role of residues within the VFA liberating aminoacylases was conducted. Site-directed mutagenesis of residues within Oct3 and Nag3 were performed to allow the elucidation of specific residue roles. These mutants were purified and assayed *in vitro* in both end point and, where applicable, kinetic ninhydrin assays with both 3M2H-Gln and HMHA-Gln to assess their impact upon enzyme activity and deduce residue function.

Mutagenesis and subsequent activity assaying has been performed on other M20 aminoacylases, perhaps most notably in human Aminoacylase 1 (hAcy1) where extensive structural, biochemical and mutagenesis characterisation was performed (Lindner et al., 2003, 2005, 2008). Mutagenesis has also been performed on an M20D aminoacylase previously although primarily focused on the effects metal interacting residues had upon the metal content of the aminoacylase from *Pyrococcus horikoshii* (*Pho*ACY) (Ishikawa et al., 2001). The mutagenesis planned for Nag3 and Oct3 focused mainly on potential substrate interacting as well as some conserved catalytic and metal binding residues to validate these enzymes are bona fide M20D aminoacylases and gain insight into the reaction with VFA precursors. The identification of key residues involved in substrate binding and catalysis would inform on the catalytic mechanism as well as potentially highlight causes of the differing substrate selectivity of the orthologous proteins. Further understanding of the mechanism of activity could also have clear biotechnological value in the development of novel inhibitors for the deodorant sector.

Finally structural characterisation of Oct3 and Nag3 was performed using X-ray crystallography to explore the structural arrangement of these proteins and allow a structural analysis between the orthologues and the published N-AGA structure (Natsch and Emter, 2020).

5.2 Selection and design of mutagenesis variants

The choice of mutagenesis candidates occurred prior to the solving of structures of both Nag3 and Oct3 and was instead devised using the published structure of N-AGA alongside protein homology models of Nag3 and Oct3 and a multiple sequence alignment (MSA) of the three. The structure of N-AGA has the benefit of being high resolution and containing a bound competitive inhibitor, termed LJ8, which closely mimics the VFA precursors thereby likely forming the same interactions that would be observed between the enzyme and their malodour substrates (Fig 5.2.1). As evidenced from the MSA of aminoacylases in chapter 3 (Fig 3.2.2), the low sequence similarity between the proteins still identifies key areas of conservation. This trend is more evident when analysing a MSA of Oct3, Nag3 and N-AGA where significant portions of the protein sequence display homology (Fig 5.2.2). This also extends to inhibitor interacting residues in N-AGA, which show high levels of similarity for most positions. Based on these factors, it was deemed acceptable to devise mutagenesis candidates for Oct3 and Nag3 using this structure.

A total of 14 mutant candidates were selected, 3 within Oct3 and the rest for Nag3, to attempt to explore three different areas of research. First, validation that both proteins are M20D aminoacylases that require the hypothesised key catalytic residues for action. Secondly, the elucidation of residues involved in substrate binding and finally an attempt to understand the differences in substrate selectivity exhibited by the two enzymes. To prevent the duplication of work, Nag3 was chosen as the primary target of mutagenesis with Oct3 only being mutated for enzyme validation under the first theme and exploration of the substrate differences in the last research area.

All but one mutant was generated by the same site directed mutagenesis method. An inverse PCR was performed on pET21b-Nag3 or pGM3 using a mutagenic forward primer and a reverse primer to amplify the entire vector-gene sequence. The mutagenic primers contained the mutated bases within the middle of the oligonucleotide sequence that was otherwise homologous to the native gene sequence. One mutant failed to generate a PCR product using this method and was instead mutated with a mutagenic reverse primer which included the desired mutagenic nucleotides at the 5’ end of the oligonucleotide. All amplified, linear mutated vector-gene DNA was then blunt end ligated and transformed into *E. coli* DH5α cells and selected for with ampicillin resistance on LBA plates after 24 h growth (37 °C). Transformants were picked for plasmid extraction and subsequent sanger sequence verification of successful mutagenesis before transformation into *E. coli* BL21 (DE3) for protein expression (Table 5.2.1).

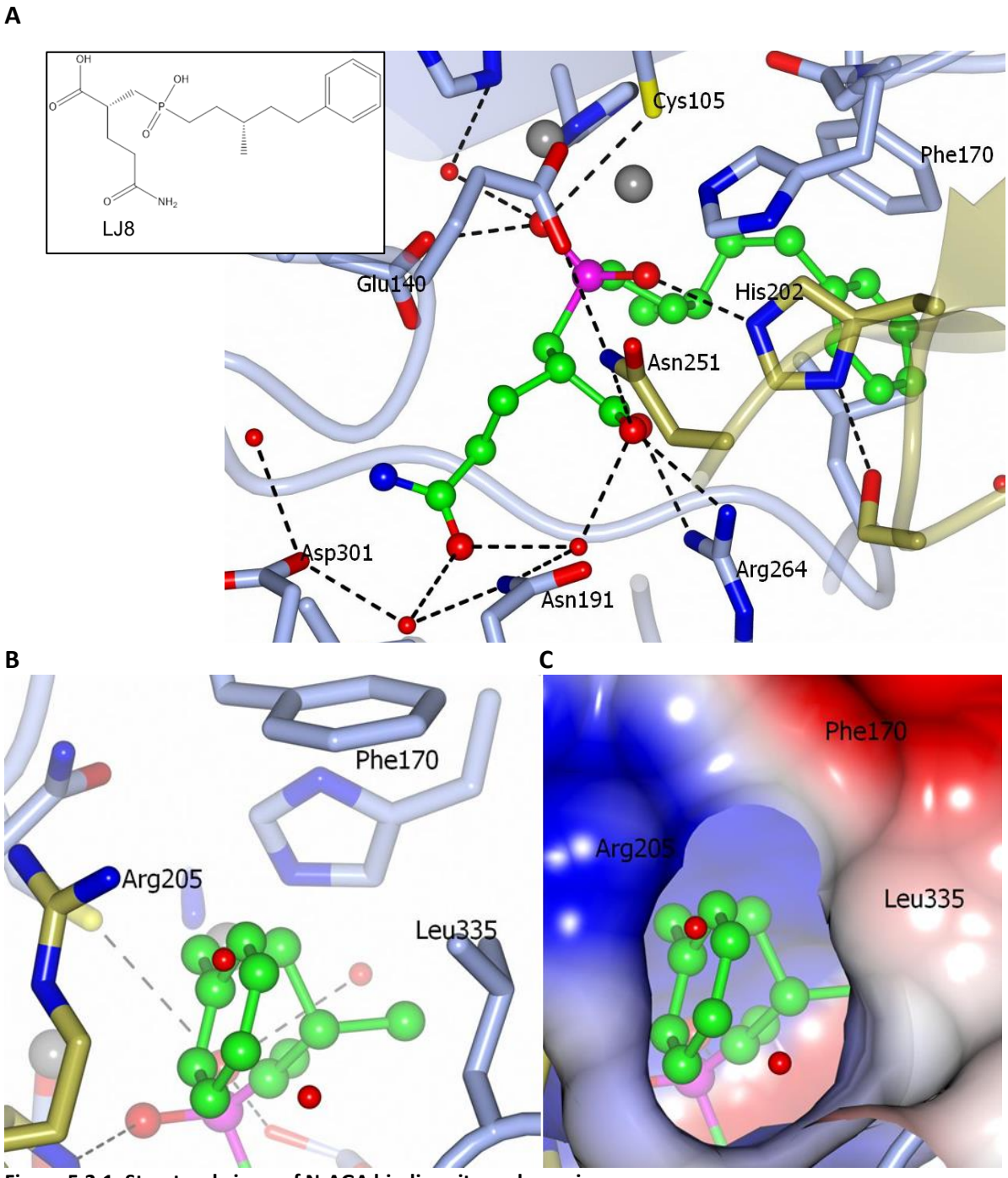


Figure 5.2.1: Structural views of N-AGA binding site and opening

Monomers A and B are distinguished by grey and gold colouring respectively. The competitive inhibitor, LJ8**,** is shown in green. Residue numbering if based on N-AGA primary sequence. **A.** View into a single monomer's binding site, labels denote residues that have been selected for mutagenesis. Chemical structure of the competitive inhibitor LJ8 is shown. **B.** View of the binding site opening which is formed principally by three residues. Arg205 from monomer B, Leu335 and Phe170 from monomer A. **C.** The same view of the binding site opening is depicted with an electrostatic interaction map.

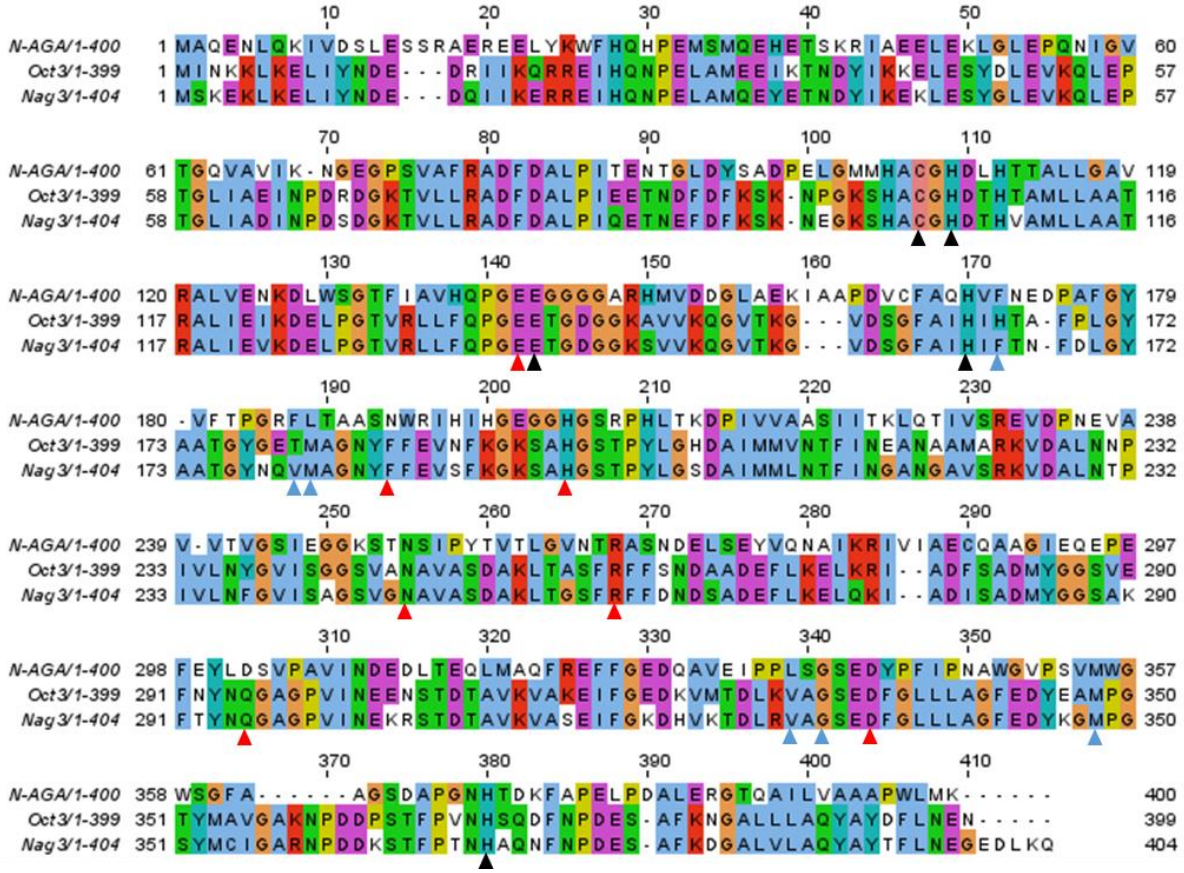


Figure 5.2.2: Multiple sequence alignment of N-AGA with Oct3 and Nag3

Translated sequences of N-AGA, Oct3 and Nag3 were aligned using Clustal Omega. Using Clustal colouring to denote conservation. Triangles highlight key binding site residues according to the structure of N-AGA (PDB 6SLF) with the colours denoting different groups: black indicates Zn coordinating residues, red to inhibitor interacting residues coordinating with the Gln and phosphinic acid moiety and blue refers to the hydrophobic pocket forming residues that accommodate the large acyl group of the inhibitor. N-AGA shares 32% and 33% identity with Nag3 and Oct3 respectively. Oct3 and Nag3 have 84% sequence identity.

Table 5.2.1: Mutagenesis targets and primer sequences		
Mutant	**Forward primer sequence**	**Reverse primer sequence**
Conserved metal binding and catalytic residues		
Nag3C103A	AAGCCATGCAgcgGGTCATGATAC	TTGCCTTCGTTTTTGCTTTTG
Nag3C103D	AAGCCATGCAgaTGGTCATGATAC	TTGCCTTCGTTTTTGCTTTTG
Nag3E138A	CAGCCTGGTGcgGAAACCGGT	AAACAGCAGACGAACGGTGC
Oct3C103A	AAGCCATGCAgcgGGTCATGATAC	TTACCCGGATTCTTGCTTTTG
Oct3E138A	TCAGCCTGGCGcgGAAACCGG	AACAGCAGACGAACGGTG
Putative Gln moiety interacting residues		
Nag3N247A	TAGCGTTGGTgcgGCAGTTGC	CCTGCGCTAATAACACCA
Nag3R260A*	TTTTTTGATAACGATAGCGCAG	cgcAAAGCTACCGGTCAGTTTGGC
Nag3Q295A	TACCTATAATgcgGGTGCAGGTC	AATTTGGCGCTACCACCA
Nag3Q295D	TACCTATAATgAtGGTGCAGGTC	AATTTGGCGCTACCACCA
Nag3H197A	TAAAAGTGCAgcgGGTAGCAC	CCTTTGAAGCTCACCTCG
Nag3F186N	AGGTAACTATaacTTCGAGGTG	GCCATAACCTGATTATAACC
Nag3F186A	AGGTAACTATgcgTTCGAGGTG	GCCATAACCTGATTATAACC
Putative residues for HMHA-Gln/3M2H-Gln discrimination		
Nag3F165H	CATTCATATCcaTACCAATTTCGATC	GCAAAACCGCTATCAACAC
Oct3H165F	CATTCATATTttcACCGCATTTCC	GCAAAACCGCTATCAACAC

Underlining denotes the codon of the mutated residue. Lower case indicates a substituted nucleotide to that of the original gene sequence. * denotes mutant generated using mutagenic 5' reverse primer.

5.3 Nag3 and Oct3 are M20D aminoacylases

Mutagenesis of metal binding and catalytic activity residues was performed for both Oct3 and Nag3 to focus on verifying that both are M20D aminoacylases that utilise the typical metal-dependent mechanism of action of aminoacylases. The metal interacting cysteine (Nag3Cys103 and Oct3Cys103) is conserved in all M20D aminoacylases and acts as a bidentate ligand to both zinc atoms and therefore is key for metal coordination and subsequent activity. The catalytic glutamate residue (Nag3Glu138 and Oct3Glu138) is conserved throughout the M20 family and believed essential for activity with its role as a catalytic base to activate the attacking hydroxyl molecule (Rawlings and Barrett, 2013; Rawlings et al., 2018). Both residues are found in to be coordinating zinc within the binding site of the published N-AGA structure and observed to be interacting with the metal cofactor and bound competitive inhibitor (Fig 5.3.1). Substitution of Cys103 and Glu138 to Ala in both proteins would confirm these conserved roles with the mutations expected to prevent catalysis and abolish any activity with their preferred substrates. An additional mutant for Nag3 at the Cys position was also created with the introduction of an aspartate, Nag3C103D. The characteristic marker of an M20D protein is the presence of the bidentate cysteine residue whereas other subfamilies typically use an Asp for this role (Rawlings et al., 2018). This mutant was produced to assess whether Nag3 could retain some level of activity, despite the expected disruption that a larger residue would be expected to have upon metal coordination.

All mutants expressed and purified to similar yield and purity (or better) to their native proteins, with nearly all producing similar UV traces upon SEC, suggesting no effect, in most of the mutants, upon protein folding or oligomeric state has been introduced the mutagenesis (Appendix C). Only Oct3E138A and Nag3N247A displayed difference to their native protein SEC UV traces with both eluting slightly later than WT proteins suggesting potential minor changes to protein shape and/ or oligomeric state. Mutants were tested in an end point ninhydrin activity assay with both VFA precursors at both 1 h and 24 h timepoints to assess whether any decreased activity due to total abolishment or severe reduction in rate. Both Nag3C103A and Nag3E138A clearly displayed no activity with both substrates even after 24 h indicating catalytic inactivity as to be expected (Fig 5.3.2). Activity was recorded for Nag3C103D with 3M2H-Gln, at reduced levels to native Nag3, but not HMHA-Gln with the protein now exhibiting a substrate selectivity similar to that of Oct3. The Oct3C103A mutant also showed complete abolishment of activity, but a small level of activity was observed for Oct3E138A after 1 h which raised to almost native Oct3 levels after 24 h (Fig 5.3.2).

Michaelis-Menten kinetic parameter determination was attempted for Nag3C103D with 3M2H-Gln but failed to yield a typical hyperbolic curve across the substrate range tested (Fig 5.3.3). Higher

substrate concentrations were not attempted due to its effects upon pH. The data points on the substrate velocity graph could be argued to almost be assuming a sigmoidal pattern, raising an interesting possibility of allosteric control, although the lack of such observations for Nag3 would suggest a more parsimonious explanation of the assay being unsensitive to the low activity rates exhibited by the mutant at lower substrate concentrations. Likewise, Oct3E138A kinetic experiments with 3M2H-Gln were unable to accurately detect sufficient product formation to enable parameter determination.

The retention of some activity by Oct3E138A with 3M2H-Gln was surprising given the suspected role of that conserved glutamine in catalytic activity. The fact that the next residue is also a glutamine, E139, which is spatially close in the N-AGA structure and involved in zinc coordination in M20D aminoacylases, which could suggest that this residue is fulfilling the role of E138, albeit much less effectively. Although this does make it surprising that a similar effect was not also observed for Nag3E138A which also has E139.

Nag3C103D did retain some activity indicating that aspartate can fulfil the role of a bidentate metal coordinating ligand within M20D aminoacylases despite differences in other metal interacting residues. The general arrangement of metal ligand residues in M20 proteins (His/Asp, Asp, Glu, Glu/Asp, His) is different to that of M20D proteins and the VFA liberating aminoacylases (Cys, His, Glu, His, His) which would suggest that alterations would have an adverse effect upon activity with impacts upon metal binding residue orientation. Notably after 24 h incubation only 3M2H-Gln hydrolysis has occurred indicating a change in substrate selectivity in this mutant to mirror that observed for Oct3. Such an effect suggests the exact binding site conformation and metal cofactor coordination in Nag3 are critical for activity, with HMHA-Gln hydrolysis especially sensitive to any changes. Given the difference between the two VFAs is the hydroxyl group present in HMHA, it suggests either a polar interaction or steric hinderance prevents correct substrate positioning for activity.

These mutants have verified that the proteins are M20D metalloenzymes that utilise a bidentate Cys to coordinate their catalytic metal ions and use a catalytic Gln residue, that is conserved across the larger enzyme family, to act as a catalytic base to prime a hydroxyl for reaction with the substrate.

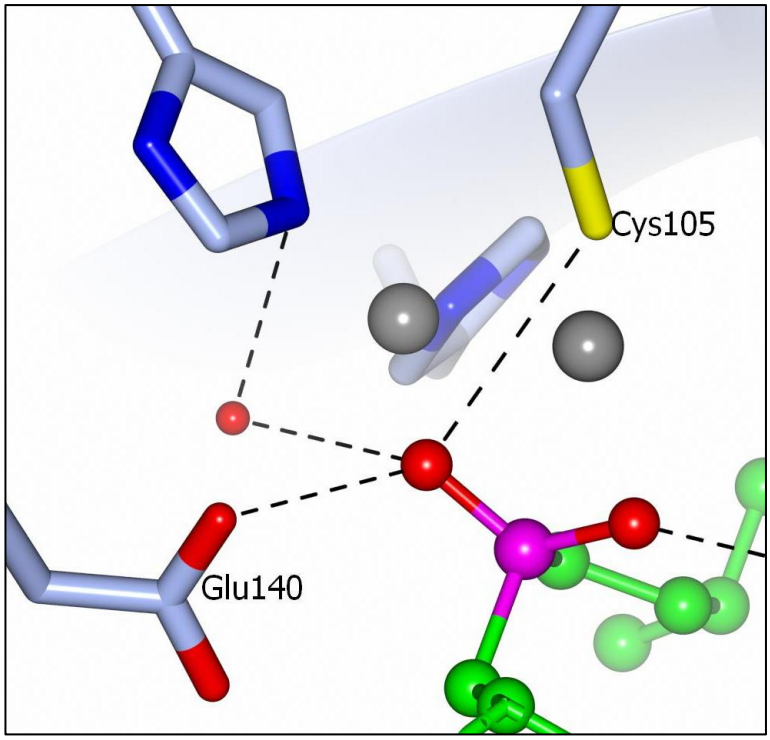


Figure 5.3.1: Close view of N-AGA monomer binding site

Close up view of a single monomer's bidentate metal coordinating Cys105 and the catalytic Glu140 residues. Grey speres denote zinc atoms and green structure highlights the bound competitive inhibitor. N-AGA structure (PDB: 6SLF) visualised with numbering referring to N-AGA primary sequence.

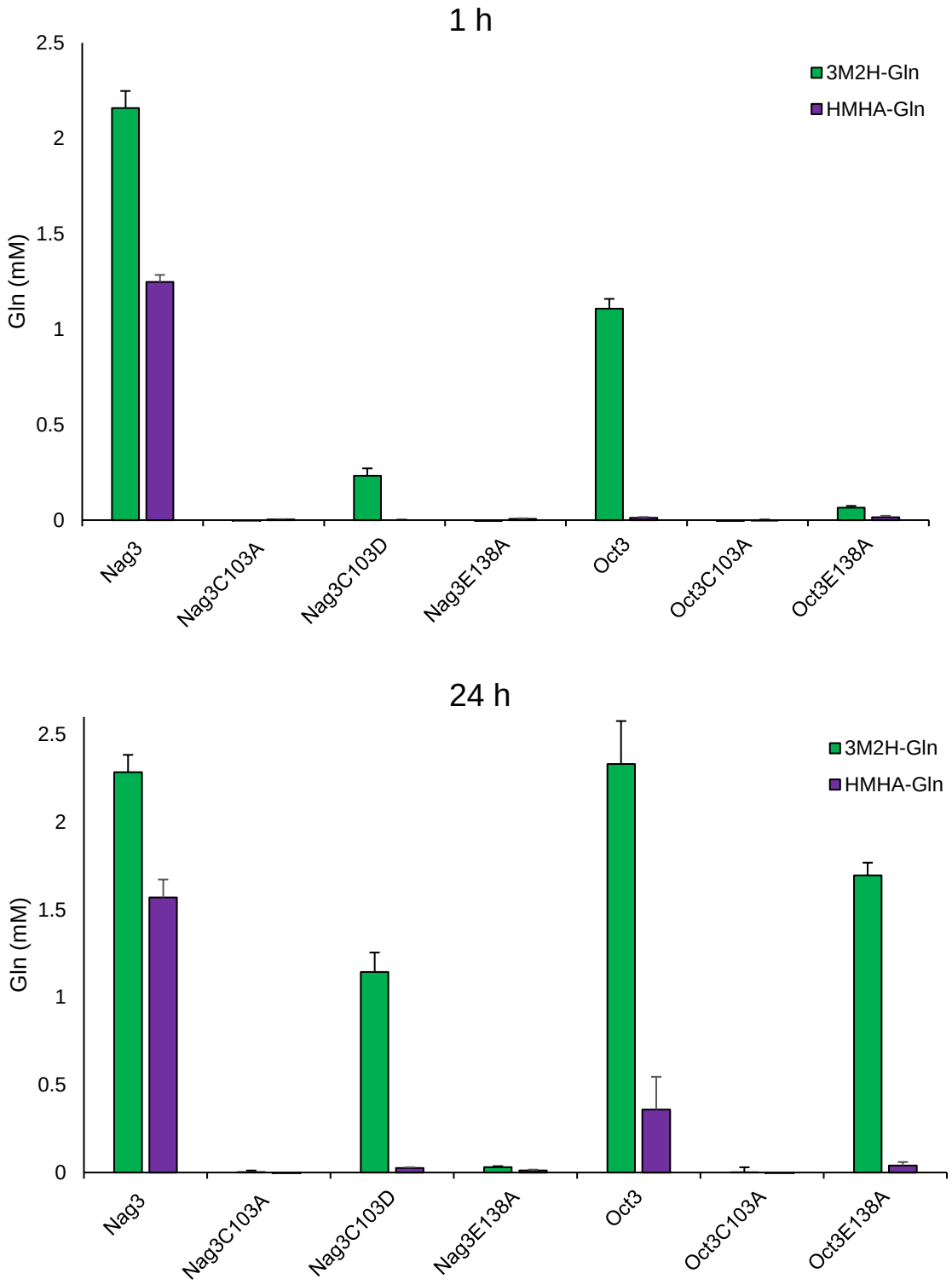


Figure 5.3.2: Endpoint activity of *Anaerococcus* metal interacting and catalytic residue mutants with VFA precursors

End point ninhydrin activity assay was performed for mutants and native enzymes after 1 h and 24 h incubation. Reactions initiated upon addition of 1.85 µM enzyme to 2.5 mM 3M2H-Gln or HMHA-Gln and incubated (37 °C). Error bars denote standard deviation (n=3).

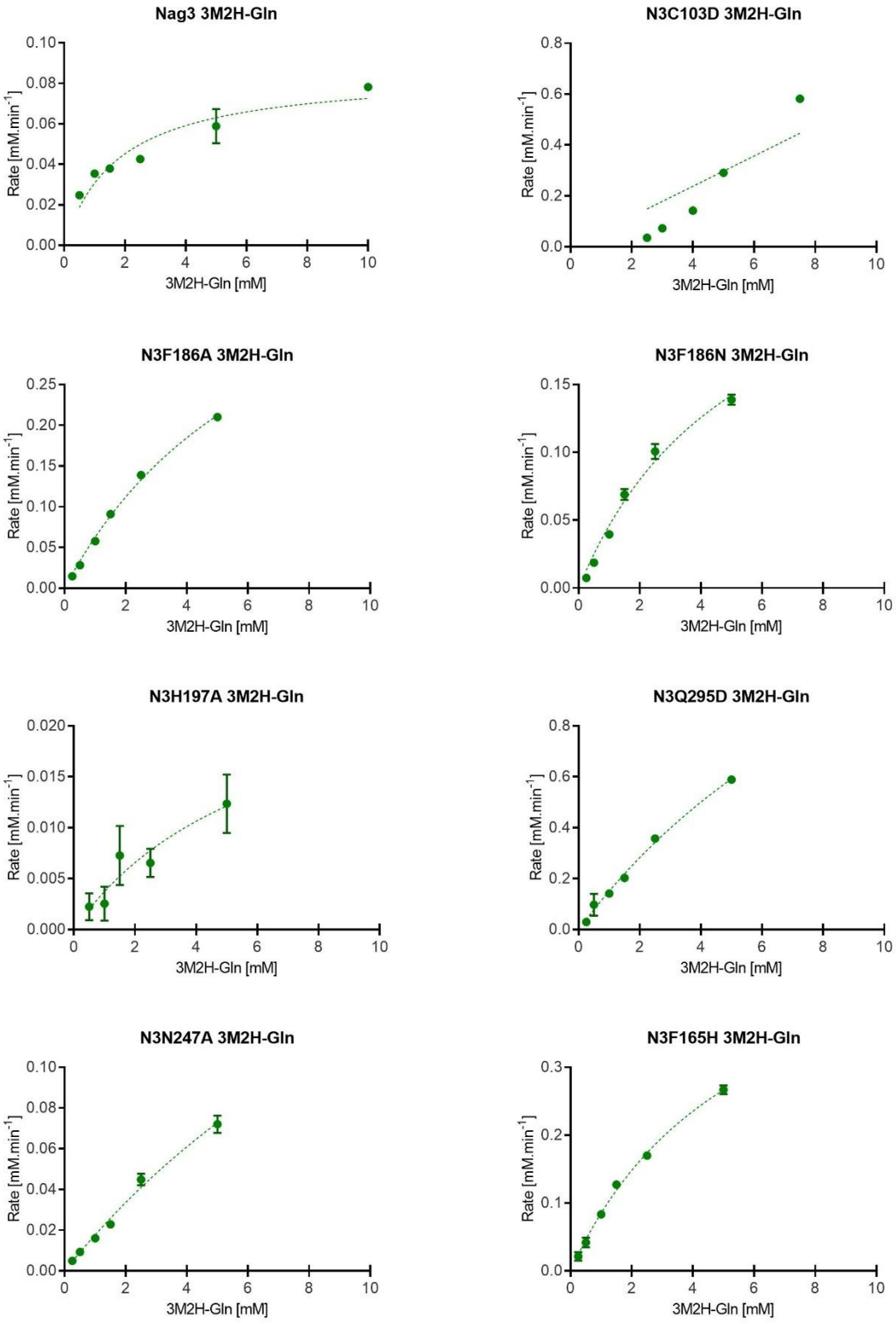


Figure 5.3.3: Substrate velocity plots for biotransformation of 3M2H-Gln by mutants

Initial rates were calculated from the initial linear rate of formation of Gln from time-course reactions of enzyme and six substrate concentrations as indicated on plots. Dots represent experimental values from triplicate replicates. Lines are Michaelis-Menten fit using nonlinear regression analysis in GraphPad Prism 9.

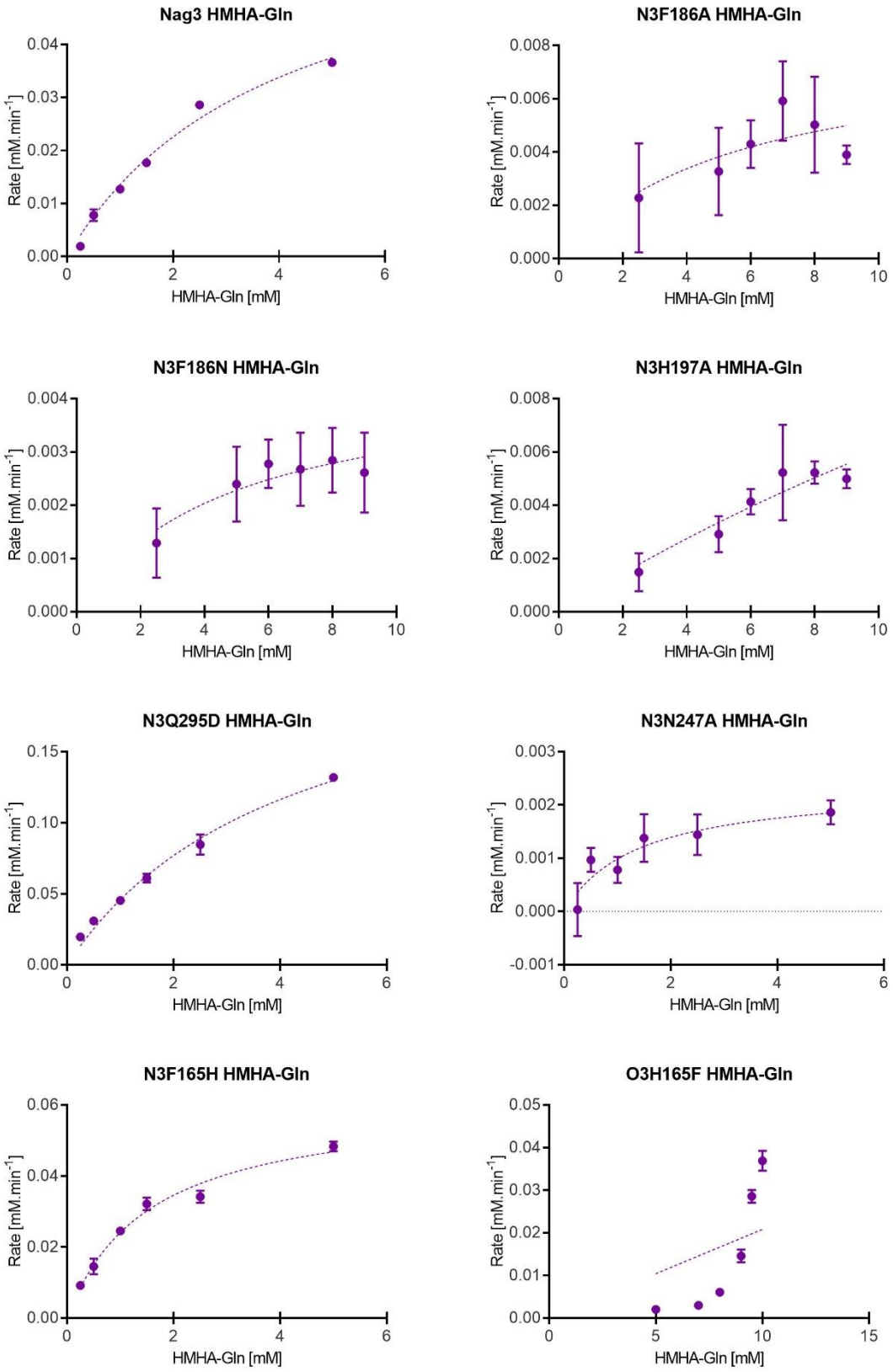


Figure 5.3.4: Substrate velocity plots for biotransformation of HMHA-Gln by mutants

Initial rates were calculated from the initial linear rate of formation of Gln from time-course reactions of enzyme and six substrate concentrations as indicated on plots. Dots represent experimental values from triplicate replicates. Lines are Michaelis-Menten fit using nonlinear regression analysis in GraphPad Prism 9.

Table 5.3.1: Michaelis-Menten kinetics for VFA precursors by mutants

	3M2H-Gln				HMHA-Gln			
Enzyme	**[E] (µM)**	**K_M (mM)**	**k_{cat} (min^{-1})**	**kcat/ K_M (min^{-1} mM^{-1})**	**[E] (µM)**	**K_M (mM)**	**k_{cat} (min^{-1})**	**kcat/ K_M (min^{-1} mM^{-1})**
Nag3	1.85	1.78 ± 0.15	46.2 ± 2.6	25.96	1.85	2.94 ± 0.22	36.4 ± 0.76	12.4
N3C103D	1.85	-	-	-	-	-	-	-
N3F186A	1.85	7.39 ± 0.54	284 ± 13.2	38.43	1.85	5.61 ± 1.40	4.38 ± 0.99	0.78
N3F186N	1.85	5.41 ± 0.41	160 ± 9.98	29.57	1.85	4.71 ± 0.85	2.39 ± 1.21	0.51
N3H197A	1.85	6.47 ± 0.75	15.0 ± 1.03	2.32	3.70	39.8 ± 8.89	8.11 ± 2.17	0.20
N3Q295D	1.85	13.53 ± 0.83	1186 ± 72	87.66	1.85	4.21 ± 0.29	129.2 ± 6.11	30.69
N3N247A	1.85	12.35 ± 0.32	132.9 ± 0.65	10.13	1.85	1.40 ± 0.39	1.28 ± 0.11	0.91
N3F165H	1.85	5.717 ± 0.24	308.8 ± 9.9	54.01	1.85	1.55 ± 0.09	33.1 ± 0.26	21.35
O3H165F					1.85	-	-	-

5.4 Characterising glutamine moiety substrate interactions in Nag3

To investigate substrate binding and affinity in Nag3 mutants, I focused upon residues predicted to interact with the glutamine moiety and amide bond in the VFA precursor molecules. The acetylated amino acid screen described in chapter 4 demonstrated that the amino acid present in different substrates has a large impact on activity. Previous biochemical characterisation of N-AGA also noted clear activity for only glutamine containing conjugates while often displaying much more tolerance in variation in acyl chain (Natsch and Emter, 2020; Natsch et al., 2003, 2004). The hydrophobic nature of the acyl moieties present in the VFA substrates and inhibitor, and their presence in an overall hydrophobic pocket of the protein binding site also made it difficult to identify key interacting residues to study.

All residues chosen for mutation in Nag3 to explore this area display some level of conservation to those in N-AGA and were also observed to form specific interactions with the bound competitive inhibitor in the structure. A noticeable exception to this is Asn191 in N-AGA which aligns to Phe186 in Nag3 (and Oct3). The mutants generated and their aligned counterpart residue in N-AGA are as follows: Nag3H197 + N-AGAH202, Nag3F186 + N-AGAN191, Nag3Q295 + N-AGAD301, Nag3N247 + N-AGAN251 and Nag3R260 + N-AGAR264 (Fig 5.4.1). Alanine substitution mutants were generated for all listed Nag3 residues as well as F186N and Q295D to match the amino acid found in N-AGA.

The side chain of N-AGA Asn191 was shown to coordinate a water network involved in interaction with the glutamine amide functional group and the carboxylic acid group in the N-AGA structure and so appears to have a key effect upon substrate recognition (Fig 5.4.1). As noted earlier this residue is not conserved in Nag3 or Oct3, leading to mutation of this residue to Ala to abolish any potential effect of phenylalanine and Asn to try introducing the effect that residue has within N-AGA. Endpoint activity assaying of both Nag3F186A and Nag3F186N showed no cleavage of HMHA-Gln after 1 h but did show some activity after 24 h (Fig 5.4.2). While 3M2H-Gln activity showed some reduction after 1 h but was comparable to Nag3 after 24 h. Both mutants therefore cause an activity profile similar to that observed for Oct3 with substrate selectivity against HMHA-Gln. As both mutants have a similar effect it is likely that the native Phe fulfils a secondary role in effecting binding site residue orientation given its lack of functional groups itself. Both mutants led to a lower affinity with both substrates (3M2H-Gln: F186A K_M 7.39 mM and F186N K_M 5.41 mM, HMHA-Gln: F186A K_M 5.61 mM and F186N K_M 4.71 mM) than native Nag3 indicating adverse effects upon substrate binding (Table 5.3.1). Higher turnover rates were observed for both with 3M2H-Gln, as a result of higher V_{max} values for each protein which does lead to better catalytic efficiency than Nag3. However, such high affinity values would make this

supposed increase in efficiency worthless under the standard physiological conditions experienced by *A. nagyae* and *A. octavius* where such levels of substrate would appear unrealistic.

N-AGA H202 was observed to be interacting with an oxygen of the phosphinic acid group of the inhibitor (Fig 5.4.1) which would mimic the positioning of a tetrahedral intermediate formed during substrate hydrolysis (Lindner et al., 2003). This residue resides on a second monomer's oligomerisation domain and points directly into the binding site which has been shown before to be key to activity in human aminoacylase 1 (Lindner et al., 2003) and is highly conserved throughout the M20 protein family, including all Oct and Nag proteins (Fig 3.2.2 and Fig 5.2.2). Nag3H197A displayed noticeable reduction in activity with both substrates at both 1 h and 24 h timepoints suggesting this residue likely fulfils a similar role in Nag3. A marked increase in affinity was observed for both substrates compared to native Nag3 with K_M 6.47 mM for 3M2H-Gln and 39.8 mM for HMHA-Gln as well as a decrease in turnover rate supporting the previous conclusions that this residue has an important role in substrate binding and activity.

N-AGA D301 was observed to be involved in interactions of the glutamine moiety R group with coordination of water molecules (Fig 5.4.1). Interactions with the amide functional group of glutamine was hypothesised to be key in enabling discrimination between different amino acid acyl conjugates. This role was believed to be conserved in Nag3 and Oct3 with the conservation of residue properties with the presence of a Gln in this position. Mutation to alanine does show a stark loss of activity with both substrate after 1 h with activity for 3M2H-Gln showing a greater increase to native Nag3 levels than HMHA-Gln activity after 24 h (Fig 5.4.2). This would further support the assumptions that 3M2H-Gln is the preferred, and likely better tolerated, substrate of the two. Mutation to aspartate to mimic N-AGA showed only minor reductions in activity in endpoint assays suggesting that conservation of chemical properties is crucial for activity. Nag3Q295D did show a higher affinity K_M 13.53 mM with 3M2H-Gln but also a much higher velocity in kinetic experiments that lead to a highly increased turnover rate and consequently much higher catalytic efficiency than Nag3. However, with a calculated K_M far greater than the plotted data point the accuracy of these values could be doubtful. Affinity with HMHA-Gln was also slightly higher for this mutant at 4.21 mM compared to Nag3 but again an increase in turnover rate was observed, leading to an improvement in catalytic efficiency.

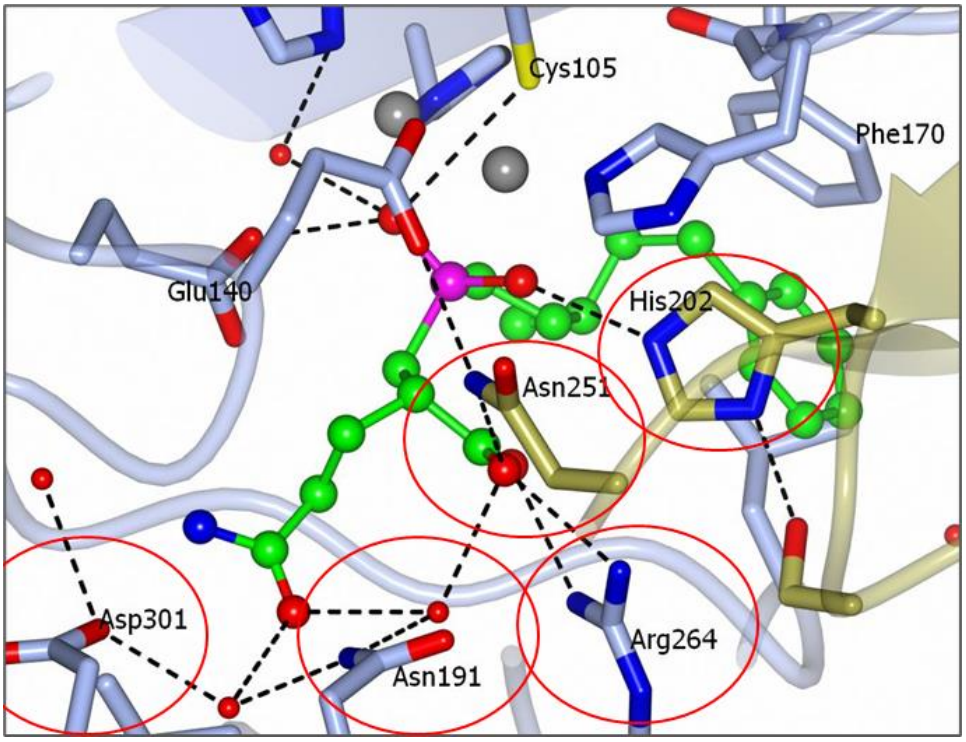


Figure 5.4.1: Binding site of N-AGA

View of N-AGA binding site containing the competitive inhibitor. N-AGA residues whose reciprocal residues in Nag3 were mutated to explore substrate binding are highlighted with red rings.

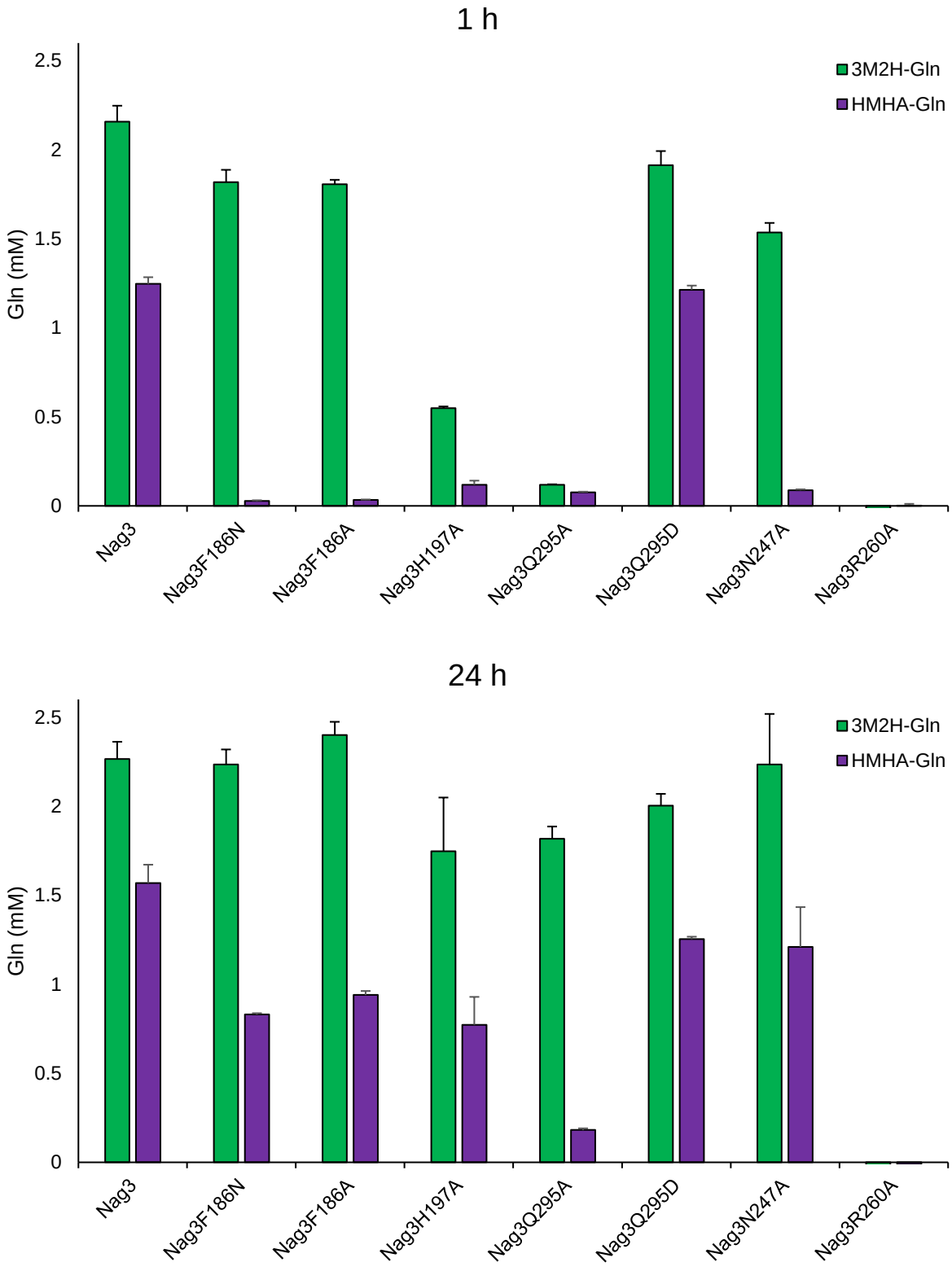


Figure 5.4.2: Endpoint activity of Nag3 glutamine interacting residue mutants

End point ninhydrin activity assay was performed for mutants and native enzymes after 1 h and 24 h incubation. Reactions initiated upon addition of 1.85 µM enzyme to 2.5 mM 3M2H-Gln or HMHA-Gln and incubated (37 °C). Error bars denote standard deviation (n=3).

N-AGA N251 from one monomer's oligomerisation domain was observed to point into the binding site of another monomer and interact with the carboxylic group of the inhibitor's glutamine moiety as well as coordinate Glu141 which is one of the key metal interacting residues (Fig 5.4.1). This residue is fully conserved in Nag3 and widely conserved in the M20 family, with evidence suggesting it plays a role in activity (Jozic et al., 2002; Lindner et al., 2005). Mutation to alanine did cause a drop in activity, most notably with HMHA-Gln after 1 h, suggesting a role in activity (Fig 5.4.2). Catalytic efficiency displayed a clear reduction with both substrates compared to Nag3 further implicating this residue as fulfilling an import role in enzyme catalysis.

N-AGAR264 was observed to form key interactions with the carboxylic group of glutamine and is another highly conserved residue in M20 proteins (Fig 5.4.1). Mutation to alanine in Nag3, Nag3R260A, led to complete abolishment of activity with both substrates even after 24 h incubation (Fig 5.4.2), confirming it plays key role in enabling hydrolysis of substrates. The absence of any activity would suggest that carboxylic interaction forms a major part of substrate interaction and positioning.

5.5 Exploring Oct3 and Nag3 substrate selectivity

To investigate the cause of the difference in activity Oct3 and Nag3 display with HMHA-Gln, binding site residues were examined to find differences between the two, particularly for any which showed conservation in Nag3 and N-AGA, which both readily hydrolyse HMHA-Gln, and not Oct3. One area that did display such differences were the residues which form part of a binding site opening in the inhibitor bound N-AGA structure. This opening is formed principally by three residues in N-AGA of F170, L335 from one monomer catalytic domain and R205 from another monomer oligomerisation domain (Fig 5.2.1 B). Sequence alignment of N-AGA with Nag3 and Oct3 showed no conservation of Arg205 and conservation of chemical properties for L335 with the presence of a valine in both *Anaerococcus* enzymes (Fig 5.2.2). Meanwhile, F170 is fully conserved in Nag3 but not in Oct3 where it is instead a His residue. The difference in properties of this residue between the two enzymes made it of particular interest to see if it could play a role in substrate specificity against HMHA-Gln in Oct3.

To test the effect of this residue, reciprocal mutations were made in Nag3 and Oct3 to swap the His and Phe present with each other, leading to the generation of Nag3F165H and Oct3H165F. If these residues did have a large impact upon substrate selectivity it was hypothesised that Nag3H165F would display a reduction in HMHA release while Oct3H165F would have an increase. Given the similar size of histidine to phenylalanine it would infer that a charge interaction between the positive charge of His165 and the hydroxyl group of HMHA-Gln and HMHA would be the likely cause. Although interaction with other residues in the binding site that significantly affect binding site shape and size could be an equally possible scenario. Initial hypothesis focused on the binding site opening seen in N-AGA having an effect on the exchange of substrates and products although several studies have noted major domain movements upon substrate or inhibitor binding in M20 members from open apo conformations to closed bound ones (Agarwal et al., 2007; Lundgren et al., 2007; Nocek et al., 2018). Therefore, it could be possible that the binding site opening has little effect upon activity if the VFA liberating enzymes undergo similar levels of movement and can assume a more accessible active site before substrate binding. Instead, a role in acyl binding could be more plausible given the positioning of the aligned residue Phe170 in the acyl binding pocket of N-AGA.

Endpoint activity assays did display an increase in HMHA release after 24 h incubation for Oct3H165F compared to Oct3 while Nag3F165H appeared to have little effect upon activity with either substrate (Fig 5.5.1). Kinetic experiments for Nag3F165H showed a higher affinity for 3M2H-Gln with K_M 5.72 mM compared to Nag3 but also an increase k_{cat} and catalytic efficiency, due to an increase in V_{max}. While this would suggest that introduction of a histidine has reduced affinity for the VFA substrates, this did not occur for HMHA-Gln where surprisingly a lower affinity value was determined at K_M 1.55

mM compared to 3.96 mM for native Nag3. As a result, it would appear that this histidine has actually enhanced HMHA-Gln activity rather than reduced it in Nag3. Kinetic parameter determination with Oct3h165F with HMHA-Gln was attempted but failed to allow accurate fitting of a hyperbolic curve to the data (Fig 5.3.4) similar to that observed for native Oct3 with HMHA-Gln (Fig 4.3.1). This would suggest both have very high affinity values which could that are far beyond the substrate range usable within the kinetic assay conditions. Given this inability to define affinity values it is therefore difficult to determine whether introduction of a phenylalanine at this position has improved catalysis of HMHA-Gln, although it does clearly show that this residue is not solely responsible for substrate selectivity in Oct3.

It is evident that substrate selectivity in Oct3 and Nag3 is achieved by a combination of several residues within the binding site likely effecting the overall conformation rather than due to a single amino acid. This can be evidenced by Nag3C103D and Nag3F186A/N showing a severe reduction in HMHA release while not directly interacting with the acyl moiety.

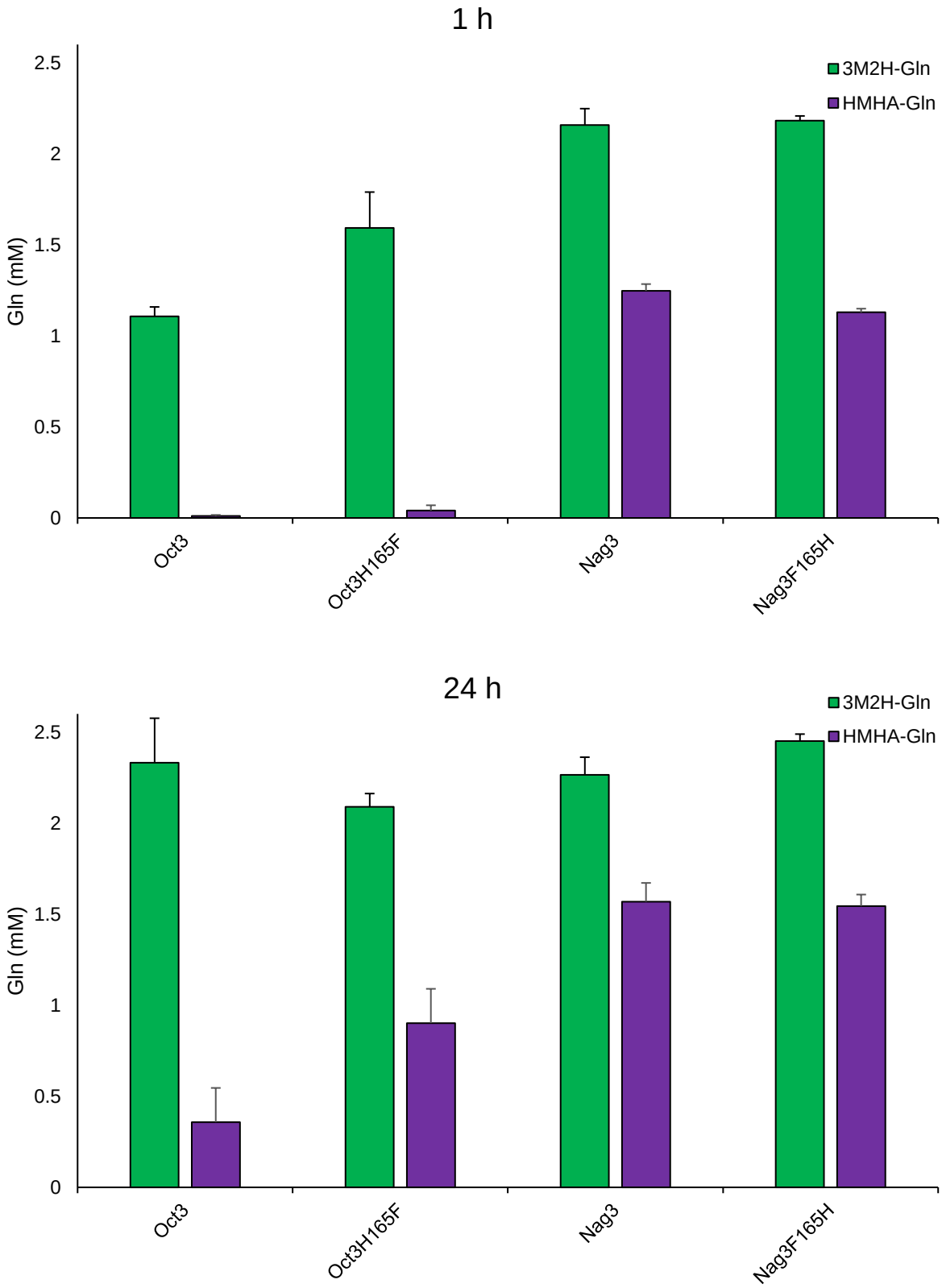


Figure 5.5.1: Endpoint activity of Oct3 and Nag3 substrate specificity mutants

End point ninhydrin activity assay was performed for mutants and native enzymes after 1 h and 24 h incubation. Reactions initiated upon addition of 1.85 µM enzyme to 2.5 mM 3M2H-Gln or HMHA-Gln and incubated (37 °C). Error bars denote standard deviation (n=3).

5.6 X-ray crystallisation and structural determination of Nag3

Attempts at structural characterisation of both Oct3 and Nag3 were conducted since their identification as active VFA liberating aminoacylases to inform upon substrate binding and the catalytic mechanism of the proteins.

Purified Nag3 protein was used in crystallisation trials of commercially available crystal screens in a 96 well-plate sitting drop format. Several different screens were tested, namely JCSG+, Crystal screen HT and PEG/ION with different concentrations of Nag3. Crystals of Nag3 were obtained for a single condition where the reservoir solution was composed of 0.2 M $MgCl_2.6H_20$, 0.1 M sodium cacodylate pH 6.5, 50 % (v/v) PEG 200 and the drop solution contained 15 mg/ml Nag3. The crystals were picked with in solution containing reservoir solution and 20 % glycerol. An in-house diffraction experiment confirmed the presence of protein in these crystals. These crystals were sent to the Diamond Light Source (DLS) for data collection on beamline i04.

Data sets were collected from two crystals at DLS, however problems were encountered during processing and ultimately we concluded these crystals were not suitable for structure determination. Repeated attempts to create better diffracting crystals of Nag3 at this condition were performed but again yielded either non-diffracting or poorly diffracting crystals.

To generate better protein crystals, Nag3 was cloned into a pETYSBLIC-3C vector which encodes a cleavable N-terminal hexa-His tag (Appendix A). This protein was purified to similar yield and purity as previous Nag3 batches and used in crystallisation experiments with both N-terminal His tagged Nag3 as well as untagged Nag3 after treatment of purified protein with 3C protease to cleave the linker region and remove the purification tag. While crystals were obtained for both N-terminally His tagged and untagged Nag3 these did not yield improved diffraction data sets upon testing at Diamond beamlines.

An improved data set was eventually obtained for the Nag3 mutant, Nag3E138A. This was selected for crystallisation trials as it is catalytically inactive allowing crystallisation experiments on the substrate bound protein as well as the unliganded form. Crystals were obtained for Nag3E138A with 3M2H-Gln, HMHA-Gln and without ligand under the same condition used for crystallisation of Nag3 (0.2 M $MgCl_2.6H_20$, 0.1 M sodium cacodylate pH 6.5, 50 % PEG 200) although at a lower protein concentration of 20 mg/ml. All crystals grown in the presence of the VFA precursors had similar difficulties with diffraction to previous Nag3 attempts upon testing at Diamond, but usable data sets were generated for apo Nag3E138A crystals on the i04 beamline. Data processing was carried out via Xia2 and the space group was determined to be P1 (Table 5.6.1)

The structure of Nag3E138A was solved by molecular replacement using Molrep, with an AlphaFold model of a Nag3 monomer as the search model (Fig 5.6.1). This revealed four protomers within the asymmetric unit, as observed for N-AGA. Iterative rounds of REFMAC and model building in COOT were carried out, with assistance from Reyme Herman and Professor Eleanor Dodson. The diffraction data were overall relatively poor, but they did enable a structural model to be created that was judged to be solved with a R value of 0.221 and a R_{free} of 0.282 (Table 5.6.1).

Table 5.6.1: Nag3 data collections and refinement statistics		
Data statistics	Nag3E138A (denotes highest resolution shell)	Oct3 (denotes highest resolution shell)
X-ray source	Beamline i04	Beamline i04
Wavelength (Å)	0.91589	0.91589
Space group	P1	$I4_122$
Cell constraints a, b, c (Å)	75.313, 79.052, 86.026	141.118, 141.118, 131.036
Cell constraints α, β, γ (°)	79.981, 74.480, 68.106	90.0, 90.0, 90.0
Overall resolution range (Å)	82.60 - 2.58	99.79 – 2.98
Inner shell resolution range (Å)	82.60 – 10.95	99.79 – 8.94
Outer shell resolution range (Å)	2.65 – 2.58	3.16 – 2.98
Overall completeness (%)	98.7 (97.8)	100 (100)
R_{merge}[a] for all I+ and I-	0.148 (1.759)	0.203 (2.184)
<I/σ(I)>	7.3 (0.7)	10 (2.1)
Overall multiplicity	3.5 (3.1)	15.7 (15.4)
CC1/2[b]	0.992 (0.438)	0.996 (0.514)
Refinement statistics		
Program	REFMAC5	REFMAC5
Number of observations	190728	216582
Number of reflections, Free R set	54901, 2771	13819, 626
R[c], Rfree[d]	0.221, 0.282	0.177, 0.228
Total number of atoms	12192	3060
Average B factor	61.51	83.92
RMS deviations – Bonds	0.0109	0.0100
RMS deviations – Angles	1.896	2.02
Ramachandran plot outliers (%)	0.95	2.53

a Rmerge = ∑HKL∑I|II - <I> |/∑HKL∑I <I> where II is the intensity of the *i*th measurement of a reflection with indexes hkl and <I> is the statistically weighted average reflection intensity.
b CC 1/2 is the correlation coefficient between two randomly selected half data sets as described in Karplus & Diederichs (2012).
c R-factor = ∑||Fo| - |Fo||/∑|Fo| where Fo and Fc are the observed and calculated structure factor amplitudes, respectively.
d R-free is the R-factor calculated with 5 % of the reflections chosen at random and omitted from refinement.

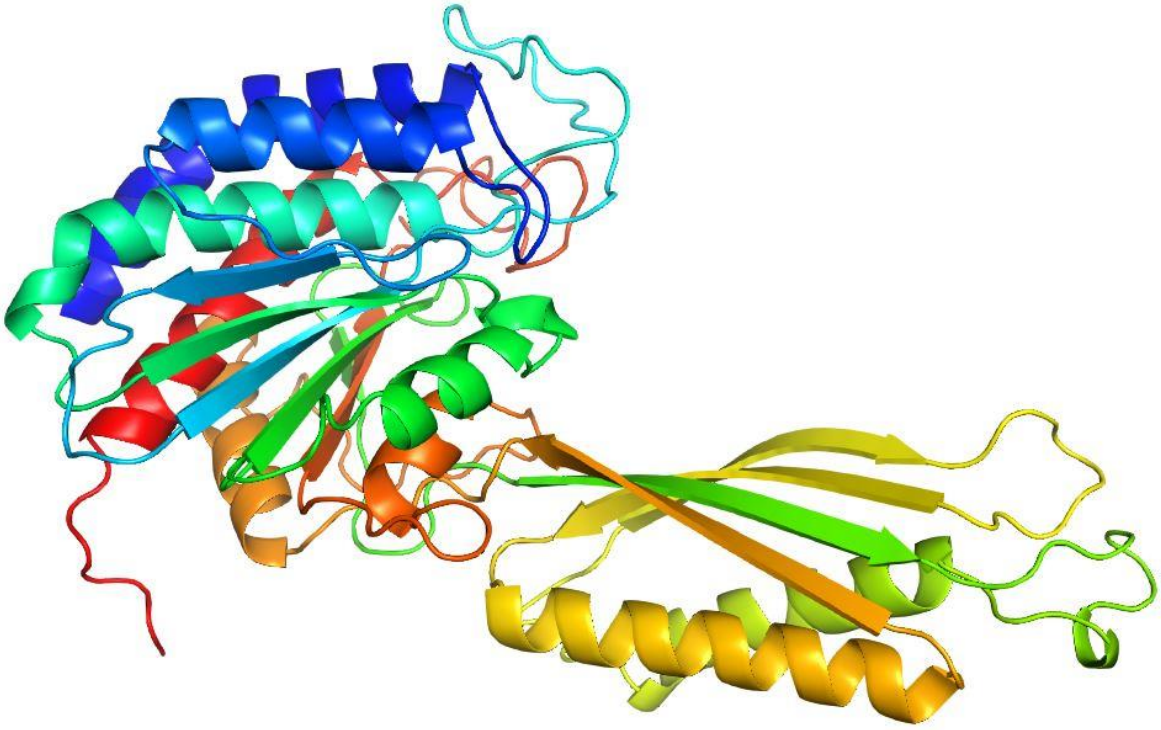


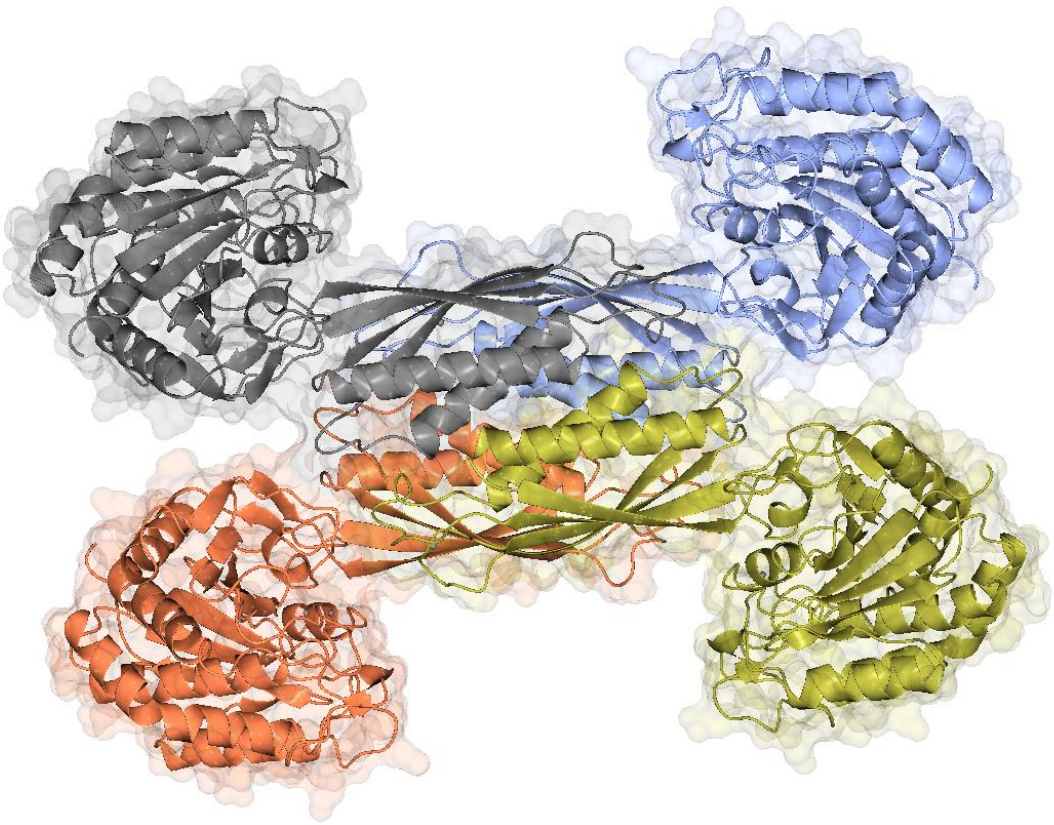
Figure 5.6.1: AlphaFold model of Nag3

Nag3 monomer model generated using AlphaFold (Jumper et al., 2021). The model has high (>90) predicted local distance difference test values for all residues except the N-terminal Met and last four residues of the C-terminal.

Nag3E138A has the typical fold and domain organisation of other M20 aminoacylases with a larger catalytic domain containing both N and C termini and forming an αβα sandwich fold as well as containing all five of the metal interacting residues (Figure 5.6.2A). The oligomerisation domain sequence is inserted within the middle of the protein primary sequence and forms an αβ fold as expected. These two domains are linked by flexible linker regions that would allow significant interdomain movement which has been proposed for this family (Nocek et al., 2018). Nag3E138A assumes a similar butterfly-like homotetramer to that of N-AGA and other aminoacylases (Botelho et al., 2011; Girish et al., 2013; Natsch and Emter, 2020), with interaction between the oligomerisation domains of each monomer. The residues coordinating the two metal ions were also observed in an expected arrangement despite the potential disruption introduced by the nearby Glu138Ala mutation (Figure 5.6.2 B). X-ray fluorescence was performed on the crystal used for structural determination alongside two other N3E138A crystals. Fluorescence spectra were collected on the i24 beamline at DLS, and auto data processing identified arsenic and zinc in all three samples due to the characteristic energy peaks detected (Fig 5.6.3). These peaks are identified by comparison to fluorescence spectra of each element. The high levels of arsenic are caused by the cacodylate buffer present in the crystallisation condition while the zinc peaks suggest it is the metal cofactor. No exogenous source of zinc was added to the crystallisation condition making it likely that this zinc is the native cofactor in recombinantly expressed Nag3. This enables a definitive answer to the identification of the native metal cofactor of Nag3 explored in chapter 4.

Superimposition of the homotetramer of Nag3E138A with that of N-AGA showed that while visually similar there are clear difference in the catalytic domain positioning between the two proteins, with an RMSD of 3.39 Å across 789 residues (Fig 5.6.4 A)**.** Superimposing a single chain from each tetramer displayed a closer similarity with an RMSD of 2.11 Å across 279 residues but highlights the differences between the orientation of the tetramer arrangement between the two (Fig 5.6.4 B). A closer view of the metal binding site of the two structures does show close similarity although N-AGA displays an overall tighter zinc coordination sphere than that observed with Nag3E138A (Fig 5.6.5 A). Measurement of distances between metal interacting residues and zinc atoms in N-AGA reveals lengths of approximately 2 Å between histidine N atoms and glutamate O atoms with Zn^{2+}, which is within the range of Zn-O and Zn-N bond distances in metalloproteins (Tamames et al., 2007). In Nag3E138A these interaction distances are greater with averaged measurements of 2.7 Å and 3 Å for histidine and glutamate residues respectively, which are beyond ideal bond distances and indicates disruption of metal coordination in this structure. The most likely source of this disruption would be the Glu138Ala substitution effecting nearby residue coordination. The difference electron density map (Fo-Fc) surrounding the zinc also suggests disruption to this area (Fig 5.7.1 B).

A



B

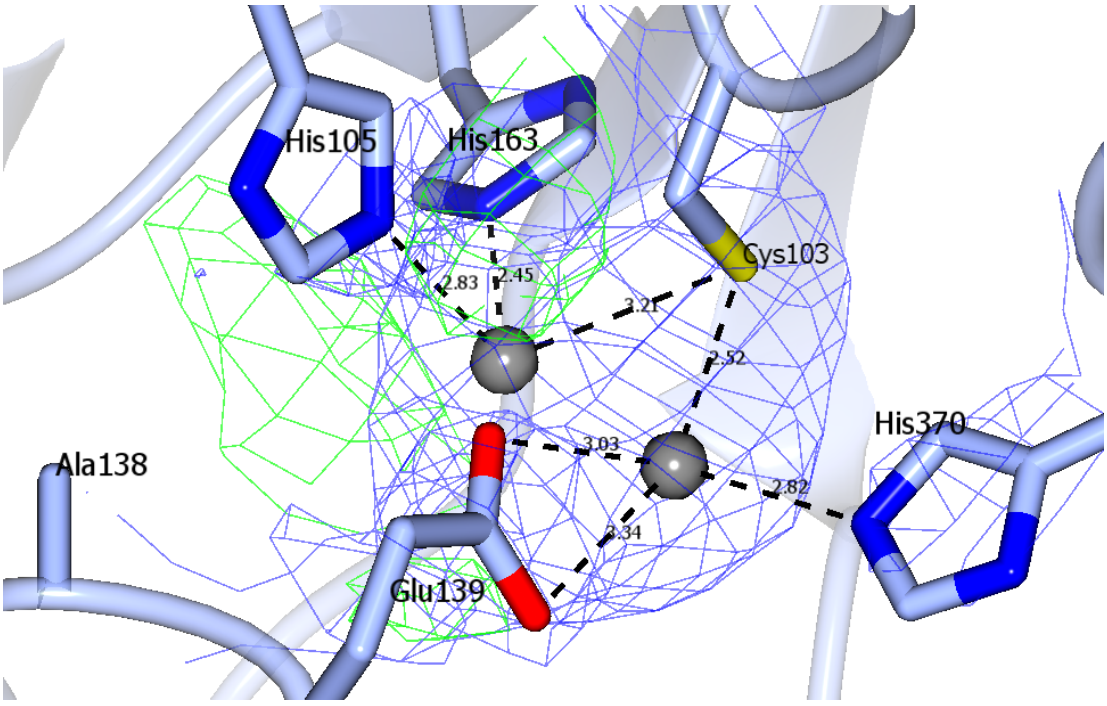


Figure 5.6.2: Structure of Nag3E138A homotetramer

A. Nag3E138A monomers are shown in ribbon form and coloured by chain. The surface is shown as a transparent background, coloured to match each respective monomer. The black square represents the monomer metal binding site, zoomed to in part B.
B. Close up view of the metal binding residues and the mutated catalytic E138A residue in the A subunit. Metal ligand interactions shown with dashed lines to two zinc atoms denoted as grey spheres. The 2Fo-Fc density map is shown in blue wire (1.5 Å RMSD). The Fo-Fc (+ve) density is shown as green (3 Å RMSD).

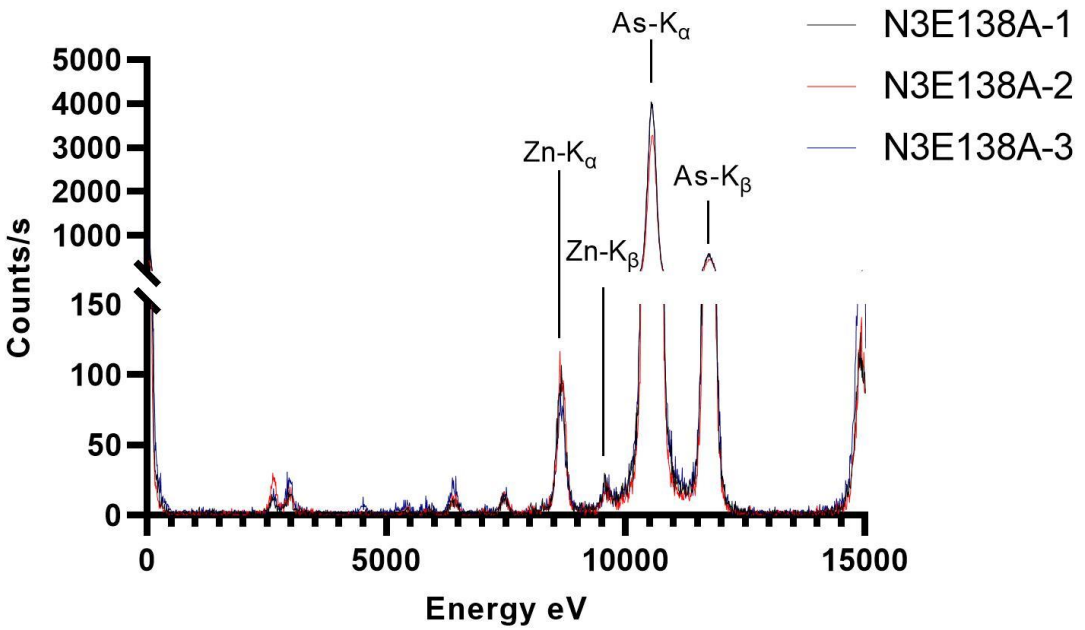
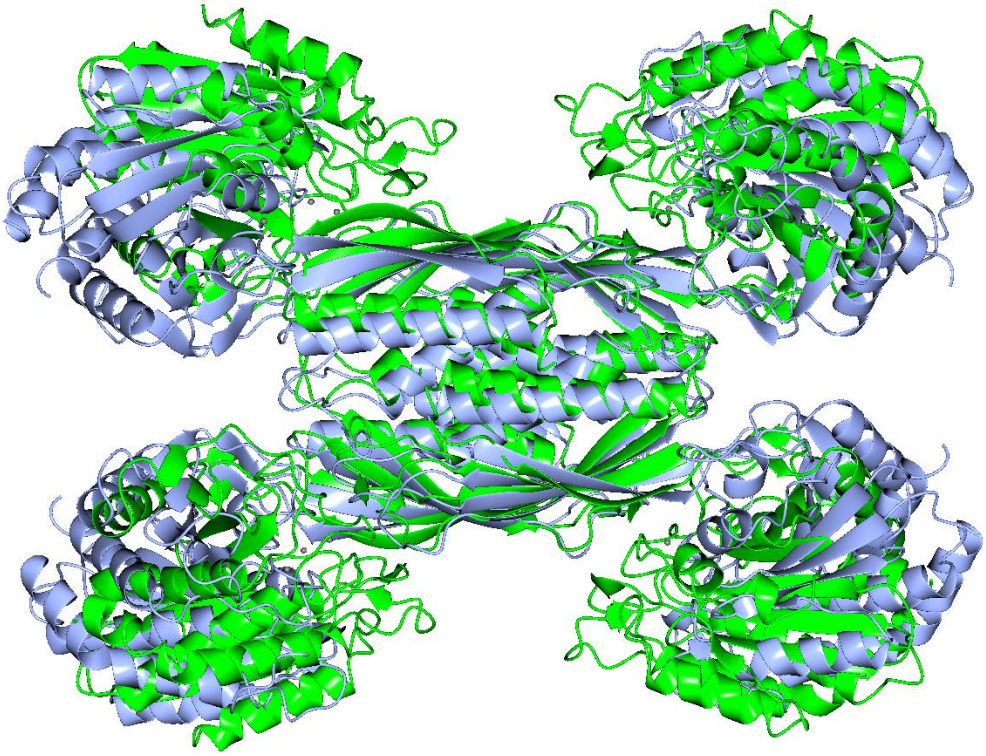


Figure 5.6.3: X-ray fluorescence spectra of Nag3E138A crystals

XRF spectra collected from three separate N3E138A crystals at Diamond, including the crystal used for structural determination (N3E138A-1). Peaks matching K_α and K_β fluorescence spectra of specific elements are labelled.

A



B

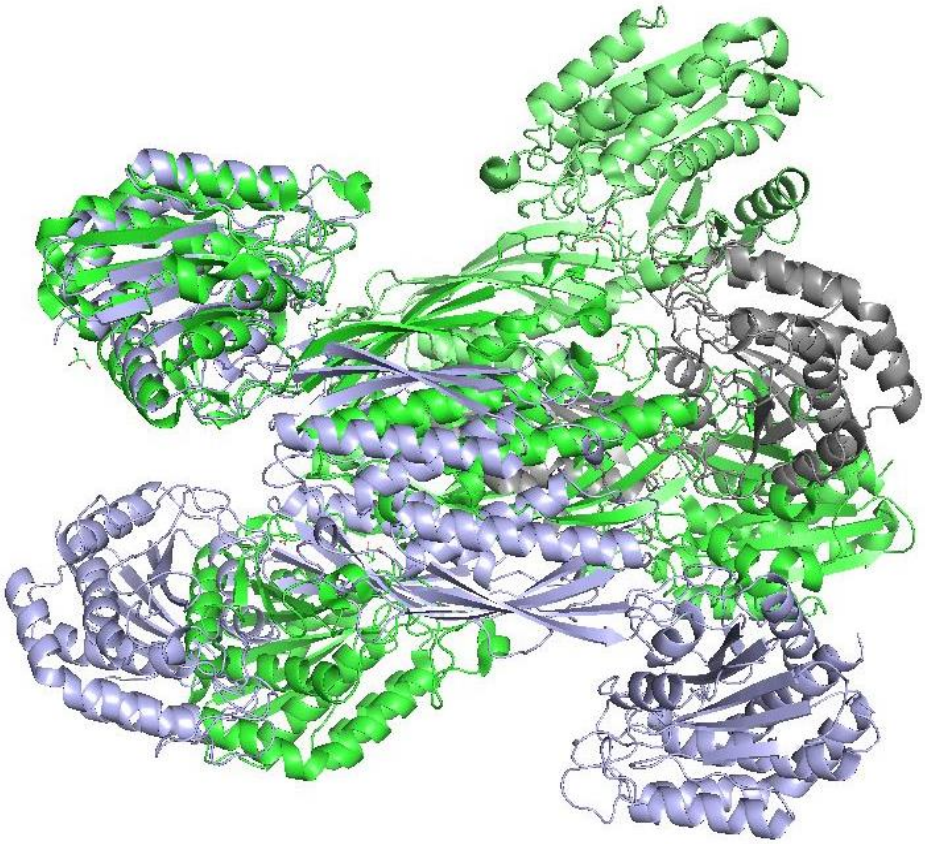


Figure 5.6.4: Superimposition of Nag3E138A and N-AGA.

A. N-AGA (PDB code: 6SLF) homotetramer was superimposed upon Nag3E138A tetramer structure using Gesamt. Nag3E138A is shown in ice blue while N-AGA is denoted as green.
B. N-AGA (PDB code: 6SLF) subunit A was superimposed upon Nag3E138A subunit A. Nag3E138A is shown in ice blue while N-AGA is denoted as green. Chain B monomers of both structures are shown in grey and light green to highlight difference.

The identified substrate interacting residues in N-AGA display significant differences in positioning within Nag3E138A (Fig 5.6.5 B), particularly Asn247 and His197 from the other monomer's oligomerisation domain which display 5.7 and 4.8 Å shifts respectively**.** This movement is caused by an overall movement in the oligomerisation domain positioning with the domain in the Nag3E138A structure being much more distant to the binding site than N-AGA (Fig 5.6.6 A). This change leads to major variation in binding site opening between the two with Nag3E138A displaying a binding site cleft (Fig 5.6.6 B) which is occluded in N-AGA by the oligomerisation domain to instead create a small opening (Fig 5.6.6 C and Fig 5.2.1 B). The competitive inhibitor, LJ8, found in the N-AGA structure does fit within the binding site cleft present in Nag3E138A after superimposition of the two structures, suggesting that this would be the position of substrate binding in Nag3 as well. It was previously noted that inhibitor addition to N-AGA leads to a change of oligomerisation state from dimer to tetramer in solution and this was proposed to also occur during substrate interaction (Natsch, 2015). This conformational shift upon binding is a likely cause of the large variation observed in monomer positioning between apo Nag3E138A and holo N-AGA structures.

Other mutated putative substrate interacting residues also display positional differences to those observed in N-AGA with Gln295 being flipped away from the binding site and Arg260 being orientated towards the oligomerisation domain rather than the substrate binding site (Fig 5.6.5 B). As the Nag3R260A mutant showed complete inactivity and this residue has been highlighted as key for substrate binding in N-AGA and others (Lindner et al., 2008; Natsch and Emter, 2020), this residue cannot assume this position during catalysis in solution. Both Arg260 and Gln295 positioning could be caused by the solid state of the Nag3E138A crystal structure which is not a definitive representation of its in solution state. Residue positioning could also be subject to change upon substrate binding in a similar mechanism as suggested earlier for oligomerisation domain positioning. Although, such a conformational change would require initial substrate binding which Gln295 and Arg260 appear to contribute to significantly (Fig 5.4.2).

The residue Phe186 is also distant to its aligned residue in N-AGA and its difference to Asn191 would prevent it fulfilling a similar role in coordination of glutamine R group moiety. The role of N-AGAAsn191 could instead be fulfilled by Asn184 in Nag3E138A which is positioned to form part of the substrate binding site (Fig 5.6.5B). Given the cleft opening observed in Nag3E138A it would appear that Phe165 has little role in providing substrate specificity by effecting the binding site opening as initially hypothesised, instead just influencing the acyl binding pocket.

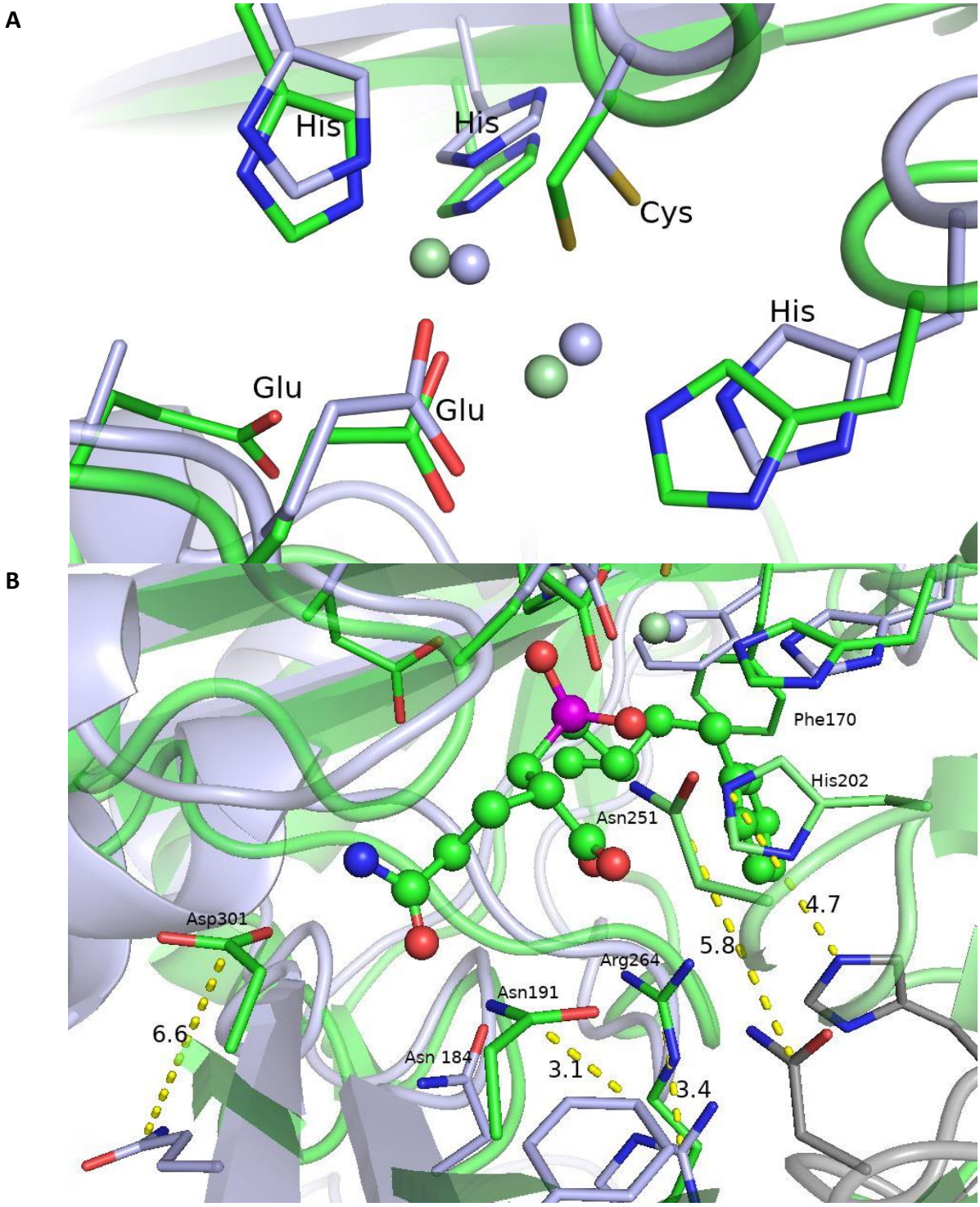
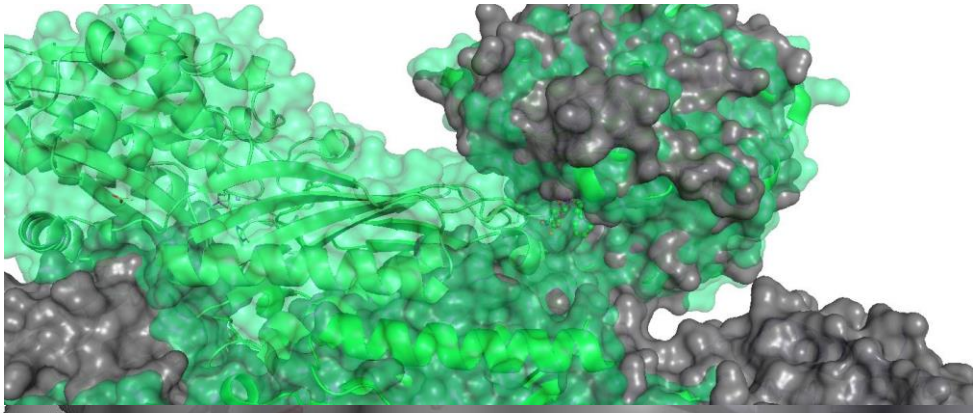


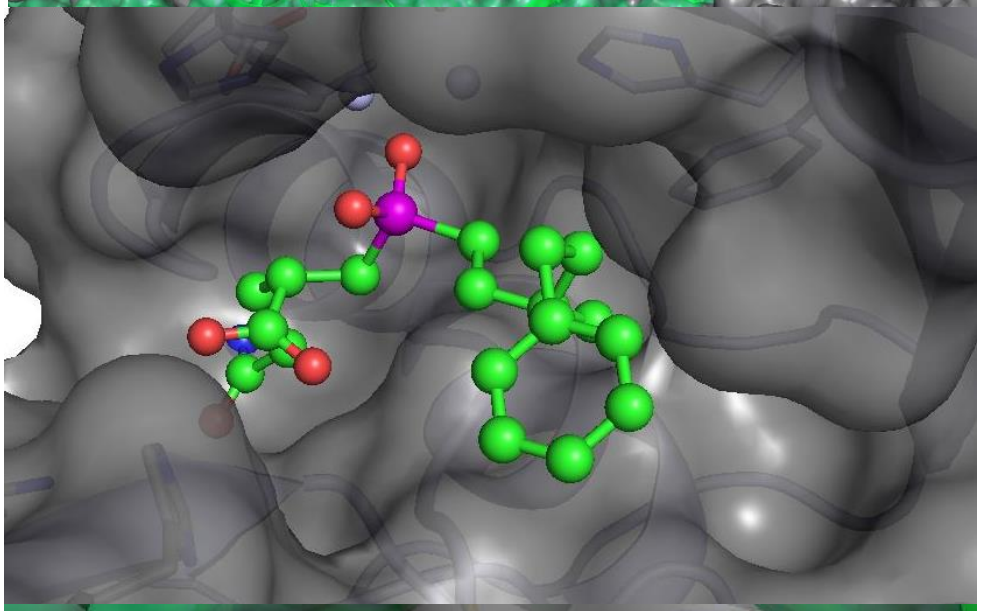
Figure 5.6.5: Nag3E138A and N-AGA substrate binding

A. Zinc ligands and coordinating residues in Nag3E138A (grey) and N-AGA (green).
B. Putative substrate interacting residues in N-AGA (green) and Nag3E138A (grey). N-AGA inhibitor LJ8 shown as green ball and stick. Distance between aligned residues in the two structures are highlighted by yellow dashed lines and listed in angstroms. Residue numbering corresponds to N-AGA primary sequence.

A



B



C

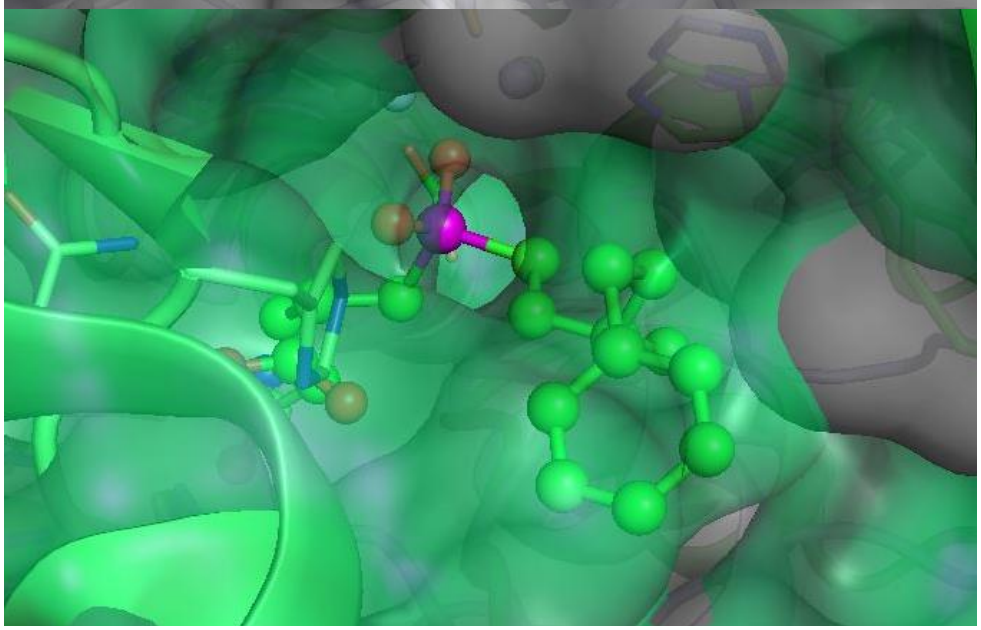


Figure 5.6.6: Surface comparison of Nag3E138A and N-AGA

A. Zoomed out view of surface differences between Nag3E138A (grey) and N-AGA (green).
B. N-AGA inhibitor LJ8 positioning within Nag3E138A binding site surface after superimposition.
C. Overlay of N-AGA surface (green) and Nag3E138A (grey) at the binding site.

5.7 X-ray crystallisation and structural determination Oct3

Alongside work on Nag3, crystallisation of Oct3 was attempted to enable structural determination by X-ray diffraction. Purified Oct3 was put into crystallisation trials using a 96 well-plate sitting drop format with commercially available crystallisation screens. The commercial screens used in these trials were JCSG+, Crystal screen HT and PEG/ION with 30 mg/ml and 15 mg/ml concentrations of Oct3. Crystals were obtained in one crystallisation condition of 0.2 M sodium thiocyanate, 20 % PEG 3350 with 15 mg/ml Oct3 after 1 week. Two crystals from this condition were fished out using nylon loops and preserved in cryoprotectant (20 % glycerol + mother liquor) before flash cooling and storage in liquid nitrogen for diffraction on the in-house X-ray diffractor. One of these crystals diffracted and was subsequently sent to Diamond for X-ray diffraction data collection on an IO4 beamline in 2019 which was processed with Xia2.

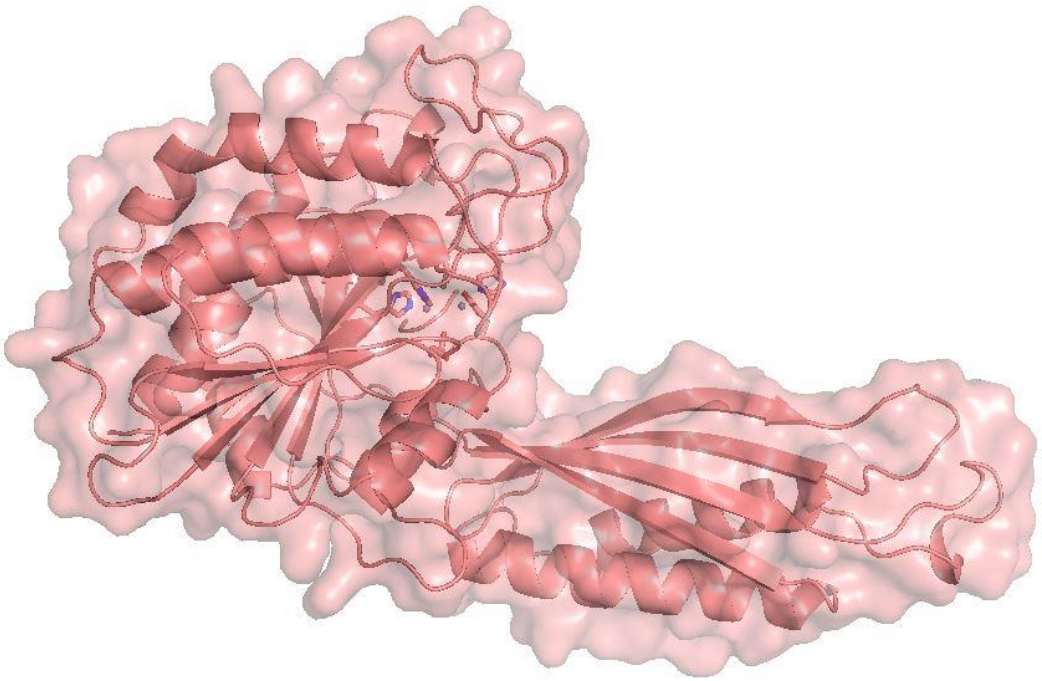
Diffraction data collected on this crystal was relatively poor and insufficient to allow structure determination by molecular replacement using the, at the time, best candidate of an amidohydrolase from *Staphylococcus aureus* COL0085 (PDB code: 4EWT) which only shared 29.9 % sequence identity. Attempts using the structure on N-AGA (PDB code: 6SLF) was also made upon its publishing in 2020 (Natsch and Emter, 2020) but also failed to yield a usable model. Success was later achieved following the structural determination of Nag3E138A and its subsequent use in molecular replacement as the search model (Table 5.6.1). Oct3 was determined to have an $I4_122$ space group and have a single protomer in the asymmetric unit. Iterative rounds of refinement and manual model building in REFMAC and COOT were performed with assistance from Reyme Herman. The structure was judged to be solved with an R value of 0.2 and Rfree of 0.25 at 2.59 Å resolution.

The Oct3 monomer assumes the same conformation of large catalytic domain and smaller oligomerisation domain as observed for Nag3E138A and N-AGA (Fig 5.7.1 A). While only a single chain was noted within the asymmetric unit, this does not indicate that Oct3 is a monomer in solution, a fact already shown by SEC-MALLS of Oct3 which revealed it as a dimer (Fig 4.4.1). A view of the metal binding residues shows the same metal coordination necessary for catalysis as observed for N-AGA and Nag3E138A (Fig 5.7.1 B). The modelled zinc atoms were visible in a contoured 2Fo-Fc map up to 9 Å RMSD. No XRF spectra were collected for Oct3 crystals at the time of data collection although anomalous diffraction consistent for zinc was found within the diffraction data collected, suggesting it is the metal cofactor present. Zinc coordination distances within Oct3 are noticeably shorter than that observed within Nag3E138A with average distances of 2.2 Å for histidine nitrogen to zinc atoms and 2.5 Å for glutamate oxygen to zinc (compared to 2.7 Å and 3 Å in Nag3E138A). The shorter

interaction distances in Oct3 are more similar to those seen in N-AGA which would support the hypothesis that the E138A mutation in Nag3 is affecting metal coordination.

In the process of refining the structure Oct3, extended electron density was observed close to the binding site of Oct3 which could not be accounted for by the primary sequence of Oct3 (Fig 5.7.2 A). This density was situated within the substrate binding site of Oct3 within the putative acyl pocket observed in Nag3E138A and N-AGA (Fig 5.7.2 B). The size and shape of this density made it difficult to identify the ligand present. It did not match thiocyanate or glycerol, present in the crystallisation condition or used as the cryoprotectant. Instead, the shape could suggest either a shortened polymer of PEG from degradation of PEG 3350 or the main chain of a peptide. Modelling in of a tetra glycine peptide does fit this density suggesting it could be a peptide although it is highly unlikely that this is the exact peptide as the only source for this ligand must be from protein degradation and no such repeat of glycine resides within the primary sequence of Oct3 (Fig 5.7.2 C). There is also no clear density to indicate amino acid side chains in which to model other amino acids, although this could be due to high flexibility in the peptide. A tetraethylene glycol molecule was also modelled and followed by a round of refinement which did suggest this as a plausible ligand to match the density with no changes to the R values of the Oct3 structure (Fig 5.7.2 D). Based on the context of crystallisation condition used it would support the ligand being a polymer of ethylene glycol rather than a peptide although it is difficult to confidently assign this ligand to either option. As the exact nature of this ligand is unknown and is likely not a substrate for this enzyme, later structural analysis exploring the binding site omits any modelled ligand at this position.

A



B

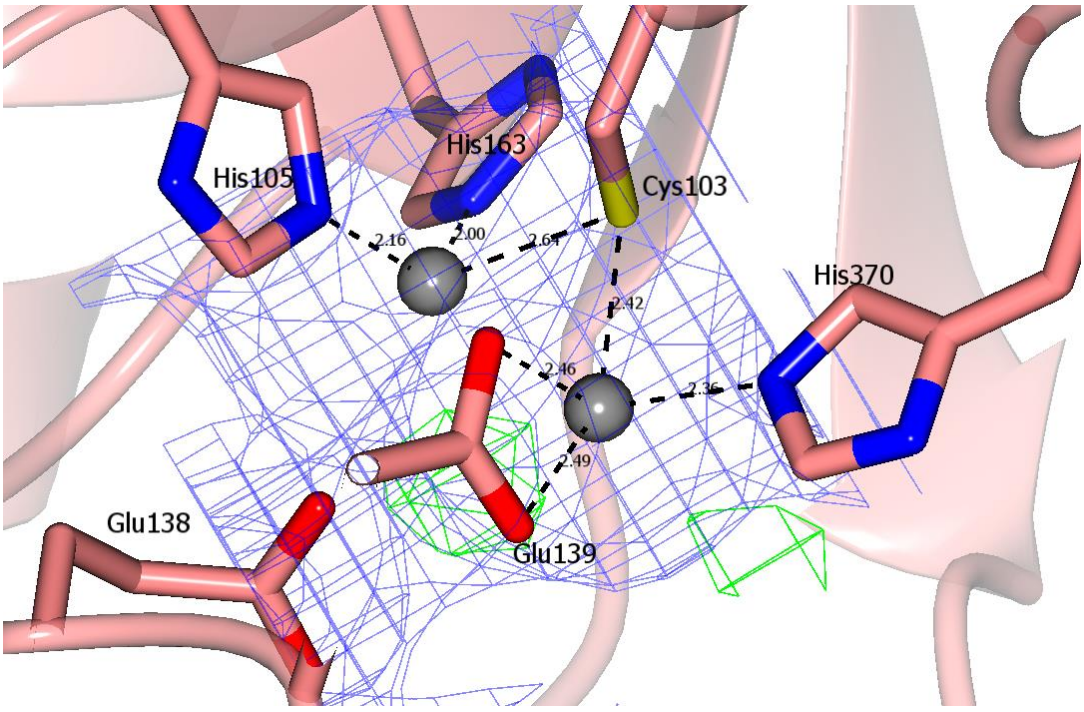


Figure 5.7.1: Structure of Oct3 monomer

A. View of single Oct3 protein found within asymmetric unit. The surface of Oct3 is shown in transparent pink. **B.** Close view of the metal binding residues and catalytic Glu138. Zinc atoms denoted as grey spheres and interactions with residues shown as dashed lines along with distances (Å). 2Fo-Fc map is shown in blue wire (1.5 Å RMSD) while F0-Fc map is shown in green (3.0 Å RMSD).

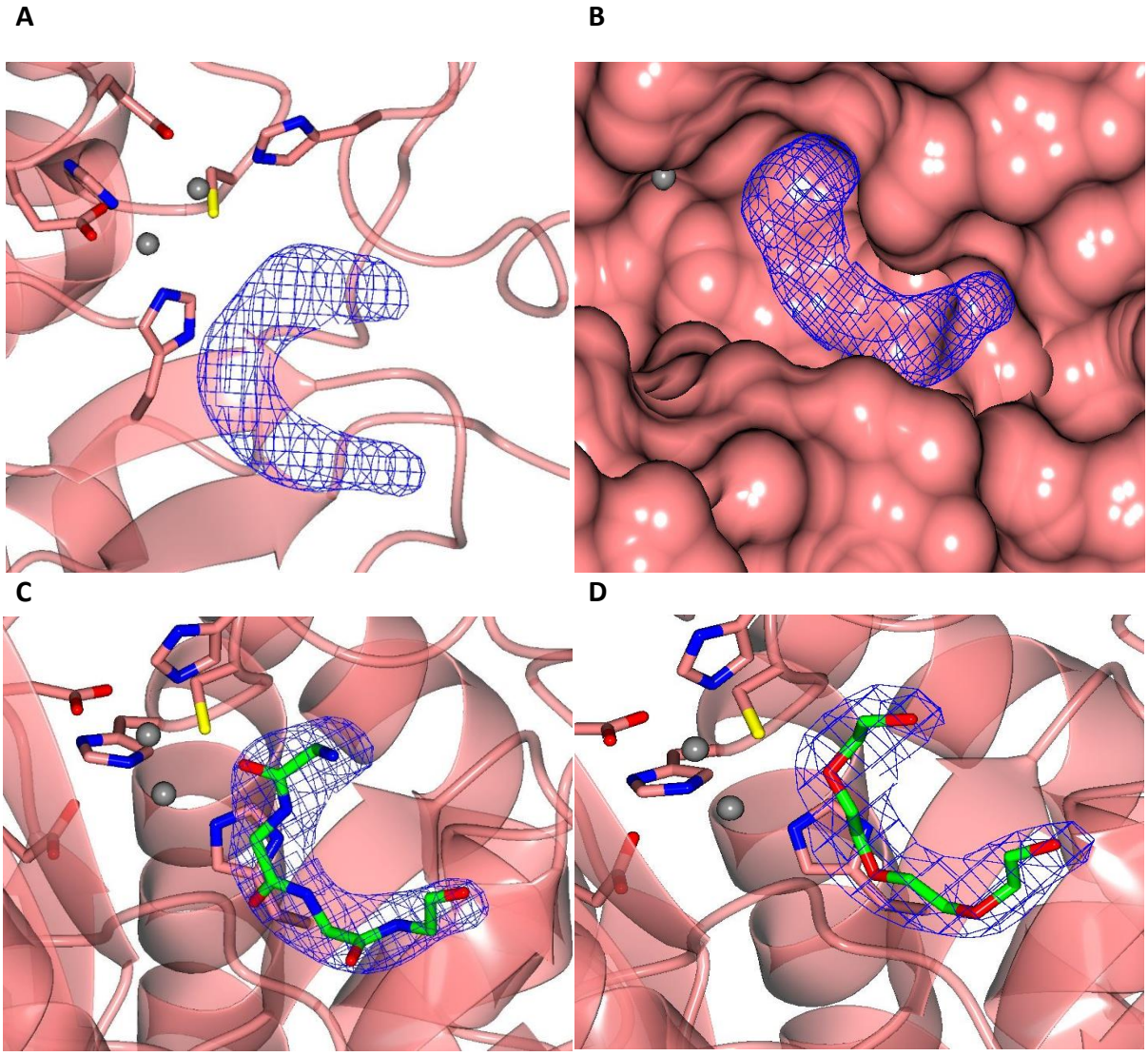


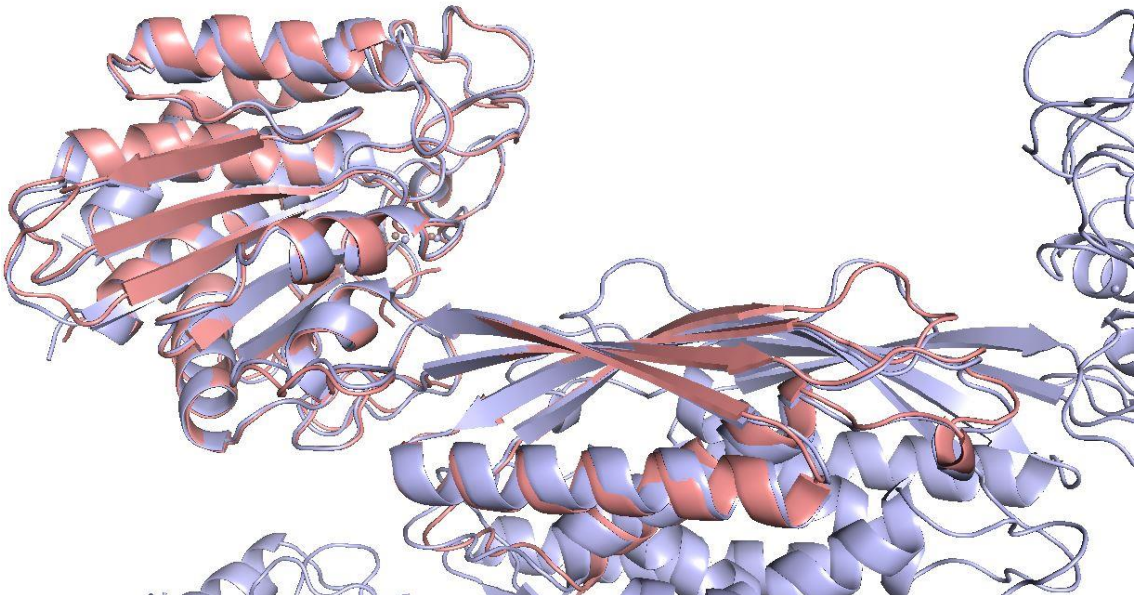
Figure 5.7.2: Electron density of unknown ligand within Oct3 binding site

A. Electron density (2Fo-Fc map) of unknown ligand shown in blue mesh in context to metal binding residues. Density maps contoured to 1.5 RMSD.
B. Ligand density shown with surface of Oct3 substrate binding site. Density map contoured to 1.5 RMSD.
C. Electron density fitted to a tetra-glycine peptide and refinement. Density map contoured to 1.5 RMSD.
D. Electron density fitted to tetraethylene glycol and refinement. Density map contoured to 1.5 RMSD.

The Oct3 protomer shows high similarity to a Nag3E138A monomer upon superimposition with an RMSD of 0.88 Å across 394 residues (Fig 5.7.3 A). This similarity also extends to the positioning of substrate interacting residues observed in Nag3E138A, indicating that substrate interactions are likely matched in the two enzymes. The presence of Asn184 in both enzymes pointing towards the binding site would suggest that this residue likely assumes the role fulfilled by Asn191 in N-AGA in both *Anaerococcus* aminoacylases. Arg260 displays a similar orientation away from the binding site in Oct3 as seen in Nag3E138A but must assume a different position towards the binding site to enable catalysis. Likewise, Gln295 in Oct3 assumes the same position with its side chain pointed away from the binding site. While His165 does show a difference in side chain orientation compared to Phe165, the presence of an accessible binding site cleft in Oct3 like Nag3E138A makes it difficult to ascertain any exact role for this residue (Fig 5.7.3 B).

As the asymmetric unit of Oct3 only contains a single monomer the residues Asn247 and His197 are not visible when looking at the binding site due to the absence of a second monomer (Fig 5.7.4 A). Due to the similarity of the structures of Oct3 and Nag3E138A it was assumed that the Oct3 would assume a similar structure to that of Nag3E138A in a tetramer oligomeric state. The Oct3 monomer was copied and superimposed onto each monomer of Nag3E138A to create the same tetramer arrangement. Visualisation of the Oct3 substrate binding site in its tetrameric arrangement shows both Asn247 and His197 in the same positions as that observed in NagE138A indicating that despite the difference in oligomeric state the two enzymes' monomers form a similar structure and replicate the likely substrate interactions (Fig 5.7.4 B). This would also indicate that formation of tetramer, at least during crystallisation, does not require major changes in monomer conformation. Instead, it would appear that extensive protein movement only occurs upon inhibitor binding as seen in N-AGA.

A



B

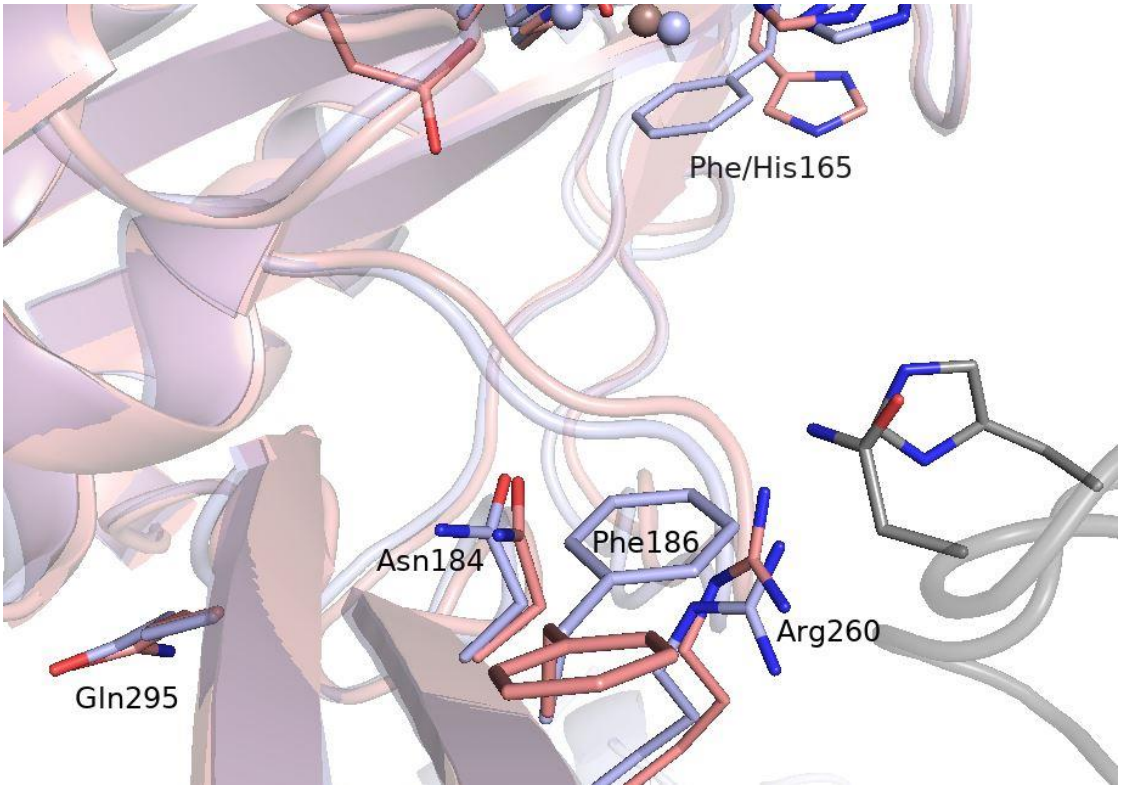
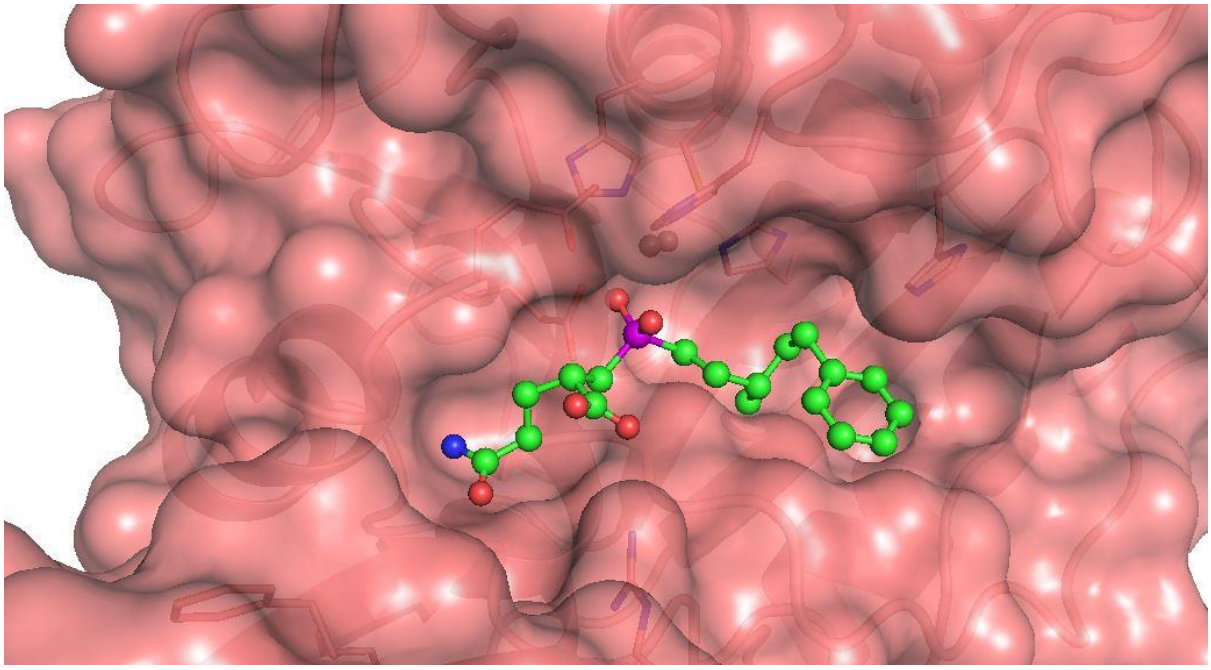


Figure 5.7.3: Comparison of Oct3 monomer and Nag3E138A

A. Superimposition of Oct3 monomer (pink) onto a single monomer of Nag3E138A (grey).
B. View of substrate binding site residues in Oct3 and Nag3E138A. A second Nag3E38A monomer is shown in dark grey and is the origin of Asn247 and His197 (depicted but unlabelled). Residue numbering corresponds to both Oct3 and Nag3 primary sequence.

A



B

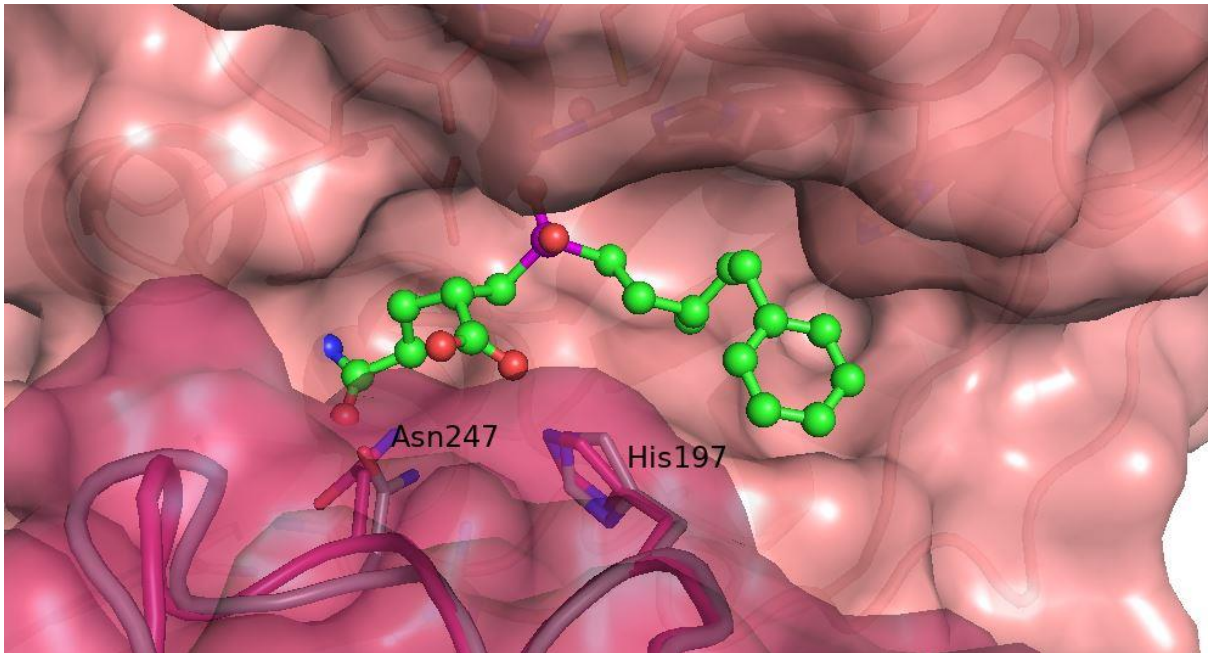


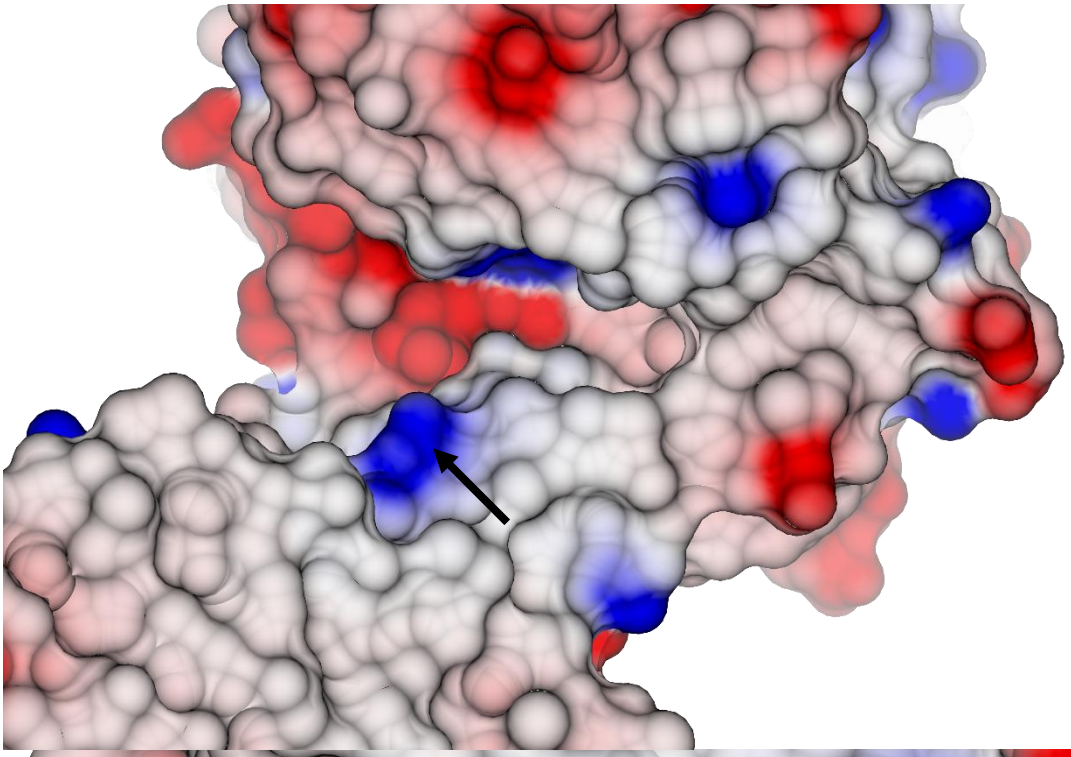
Figure 5.7.4: Oct3 binding site in monomer and tetrameric state

A. View of Oct3 monomer binding site surface. Inhibitor LJ8 is denoted in green and placed in the binding site. **B.** View of Oct3 tetramer binding site surface. The Oct3 chain B (second monomer) is shown in pink. Nag3E138A chain B in grey after superimposing. Residue numbering refers to Oct3 and Nag3 sequence.

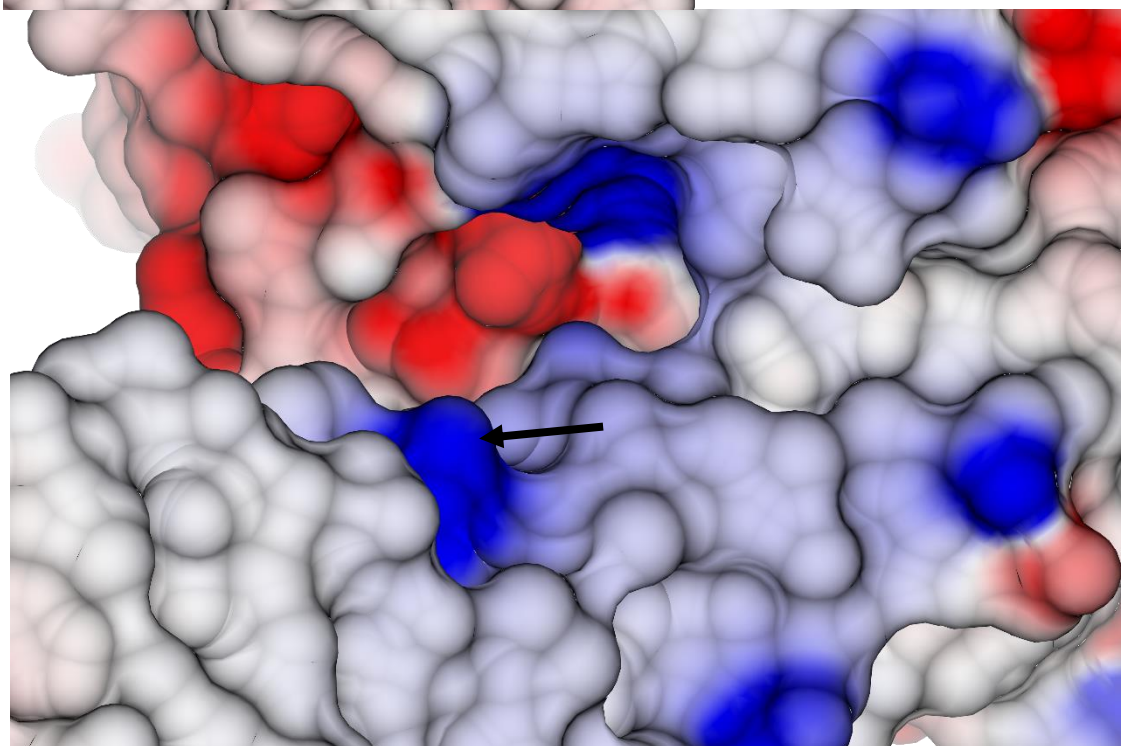
5.8 The acyl binding pocket of N-AGA, Oct3 and Nag3

The structures of Oct3, Nag3E138A and N-AGA were analysed to allow comparison between substrate binding sites to understand the difference in substrate specificity. All three enzymes displayed a charged binding site focused on the glutamine and amide bond interacting residues as well as the metal binding site (Fig 5.8.1). The positive charge of Arg260 is also prominent in all three structures, highlighting its importance in substrate binding to interact with carboxylate of glutamine. The large acyl pocket within each protein is primarily neutral due to the hydrophobic residues that are present, enabling accommodation of the hydrophobic acyl chain of subs (Fig 5.8.1). The acyl pocket residues in both Oct3 and Nag3 show a high level of similarity and are comprised of (using Nag3 residue numbering): Tyr352, Phe165 (His165 in Oct3), Asn167 (Ala167 in Oct3), Leu327, Gly331, Met181 and Ser332. This pocket is similar in size and contributing residues in N-AGA although the proximity of the oligomerisation domain from a second monomer allows a view of additional residues such as His202 and Arg205 which point into the binding pocket (Fig 5.8.2B). While His202 is fully conserved in both *Anaerococcus* enzymes and predicted to interact with the amide bond in the VFA precursors, Arg205 has no binding interactions with the LJ8 inhibitor in the N-AGA structure and aligns to a threonine in Oct3 and Nag3. This arginine does form part of a binding site opening observed in N-AGA (Fig 5.2.1B) but the smaller size of threonine and the absence of structural data for Anaerococcal enzymes in this bound conformation makes it difficult to assign any role to these residues. The rest of the acyl pocket in N-AGA is comprised of many of the same residues as seen in Oct3 and Nag3 including: Met354, Phe170, Leu335, Phe185, Ser336, Gly337, Leu186 and Ser338. The large size of this acyl pocket in all three enzymes suggests that they would be able to fit various substrates with of differing acyl moiety lengths, something which was observed with N-AGA (Natsch and Emter, 2020; Natsch et al., 2003). However, as Oct3 does display a substrate specificity against HMHA-Gln compared to 3M3H-Gln this region of the substrate binding site must confer some level of specificity. The differences in conformation between inhibitor bound N-AGA and apo Nag3E138A and Oct3 would suggest that specificity is conferred in Oct3 when the second monomer oligomerisation domain is proximal to the binding site during interaction with substrates. The lack of structural data for either Nag3 or Oct3 in this bound conformation makes it difficult to ascertain exactly how this specificity is achieved.

A



B



C

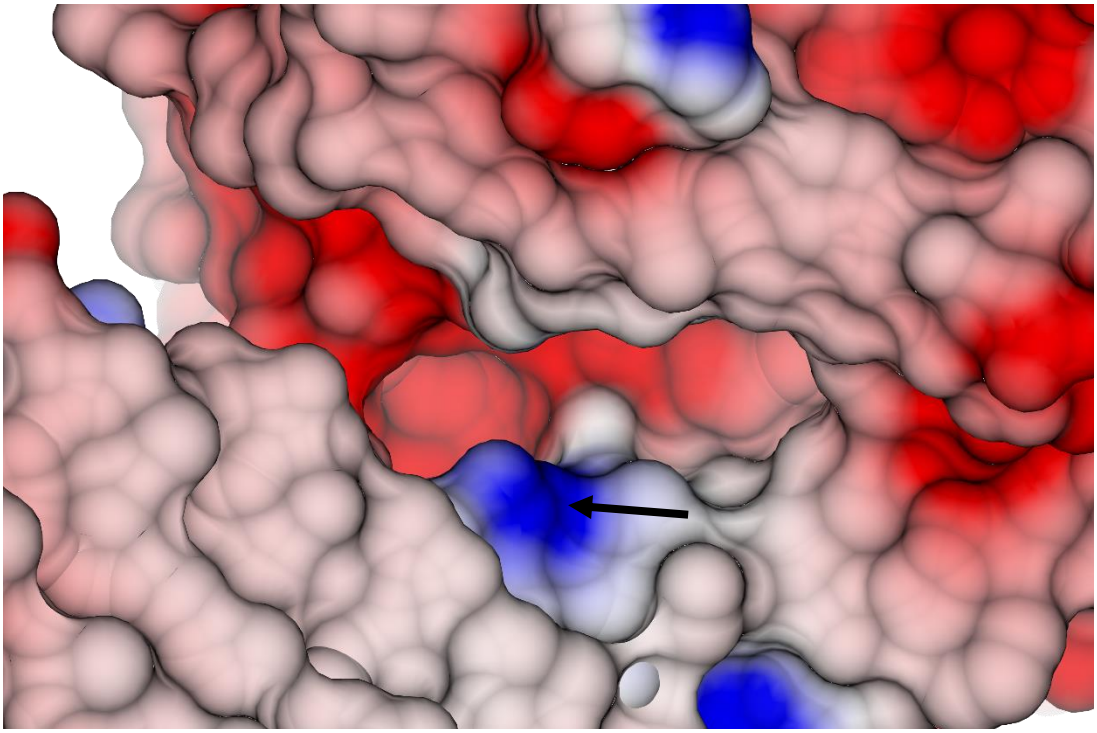


Figure 5.8.1: Electrostatic potential of VFA liberating aminoacylase binding sites

A. The electrostatic potential of Oct3 is shown. Red indicates negatively charged areas, blue denotes positively charged and white is neutral. Black arrow highlights positive charge of Arg260.
B. The electrostatic potential of Nag3E138A single monomer is shown. Red indicates negatively charged areas, blue denotes positively charged and white is neutral. Black arrow highlights positive charge of Arg260.
C. The electrostatic potential of N-AGA single monomer is shown. Red indicates negatively charged areas, blue denotes positively charged and white is neutral. Black arrow highlights positive charge of Arg264.

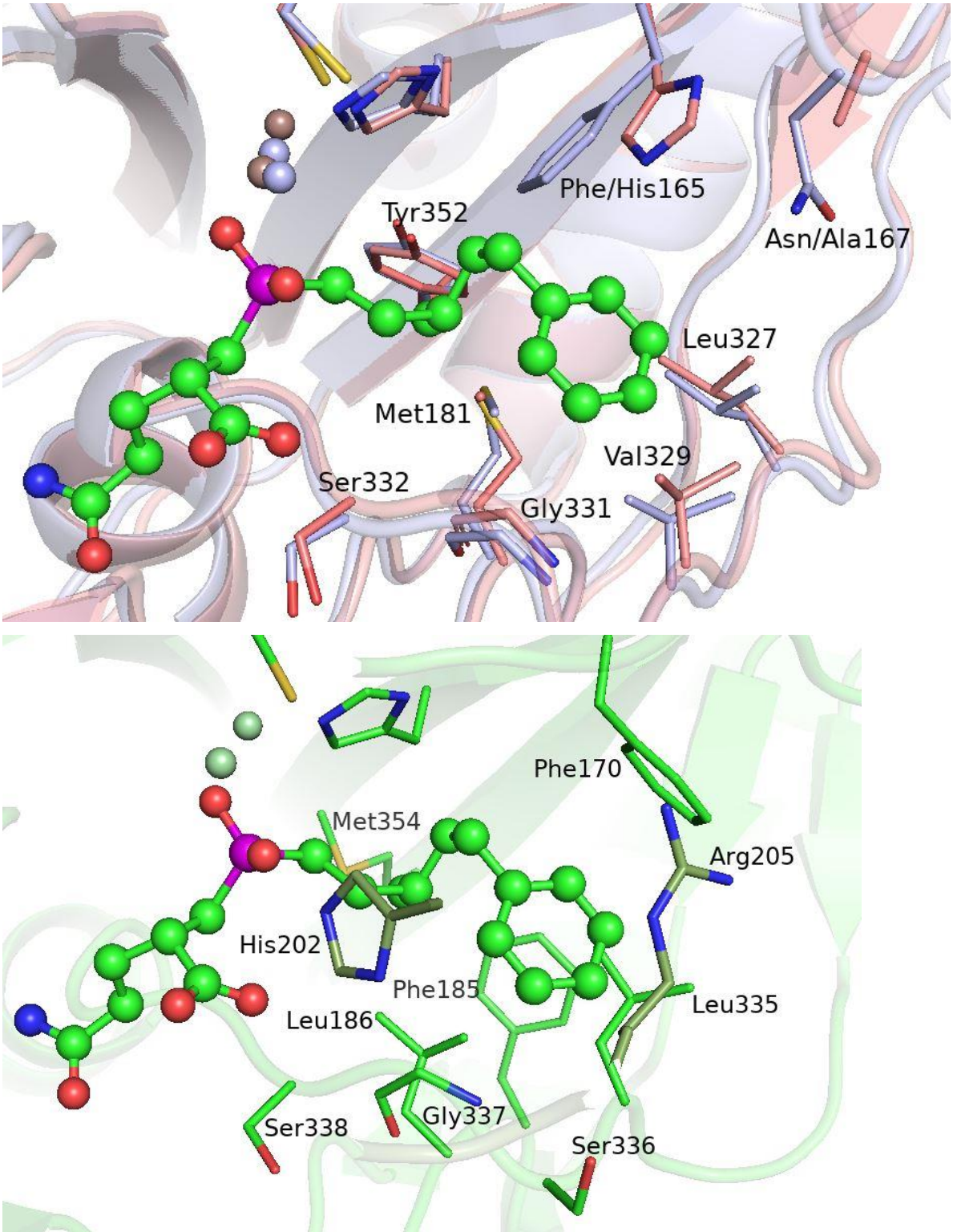


Figure 5.8.2: Acyl binding pocket of VFA releasing aminoacylases

A. Nag3E138A (grey) and Oct3 (pink) have been superimposed and residues that comprise the acyl binding pocket highlighted and labelled. Residue numbering is based on Nag3 primary sequence. The inhibitor LJ8 was placed in the binding site from alignment of Nag3 to N-AGA and is shown in green. Spheres denoted zinc atoms coloured for respective protein.
B. View of acyl binding pocket of N-AGA (green). Inhibitor LJ8 also shown in green. His202 and Arg205 are coloured dark green to denote their origin from a second N-AGA monomer.

5.9 VFA precursor binding

The binding of 3M2H-Gln and HMHA-Gln in the active aminoacylases was investigated by building the two substrates into the binding sites. Both Nag3E138A and Oct3 appear in an open conformation and have several key substrate interacting residues (such as the conserved Arg260) not in the correct orientation for substrate binding. Instead, the bound N-AGA structure was used to explore VFA precursor positioning within the binding site. As the inhibitor LJ8 in the structure mimics substrate binding and contains an almost complete glutamine moiety, it was used as the template to build the two substrates. The glutamine moiety was fixed in place before replacing the phosphinic acid group and acyl chain with the amide group and remaining acyl length of either 3M2H or HMHA. The carbonyl positioning of the amide group was matched to one of the oxygens attached to phosphorous as it was mimicking a transition state that would occur during catalysis. The rest of the acyl chain was allowed to freely assume the most likely conformation with Merck Molecular Force Field minimisation without any fixed interactions of the acyl chain to residues within N-AGA.

Both 3M2H-Gln and HMHA-Gln fit within the binding site of N-AGA as expected given the similarity of the competitive inhibitor (Fig 5.9.1 A&B). No clear interactions could be inferred from acyl chain positioning in pocket given the overall hydrophobic nature of both the acyl substrate and surrounding residues. HMHA-Gln is more hydrophilic but the hydroxyl group is still proximal to other polar groups in the substrate resulting in most of the acyl chain being involved in similar interactions as 3M2H-Gln. Hydroxyl positioning does bring this too close to His202 (2 Å distance) which would clash (Fig 5.9.1 B). This histidine is seen to interact with one of the oxygen atoms bound to phosphorous in LJ8 (Fig 5.2.1) and mimics the transition state that would be formed during catalysis with the carbonyl carbon becoming tetrahedrally bound upon hydrolytic attack. Assuming that Oct3 would form a similar bound state upon substrate binding this potential hydroxyl- histidine clash could be responsible for substrate specificity against HMHA-Gln if it affected carbonyl positioning. Although Nag3 also has this conserved His and would likely be positioned similar to Oct3, making it difficult to predict how this clash may only effect one orthologue and not the other.

Determination of the structure of Nag3E138A, and catalytically inactive Oct3, bound to the VFA precursors would provide the best means to determine the causes behind the differing substrate specificity exhibited by the two orthologous enzymes.

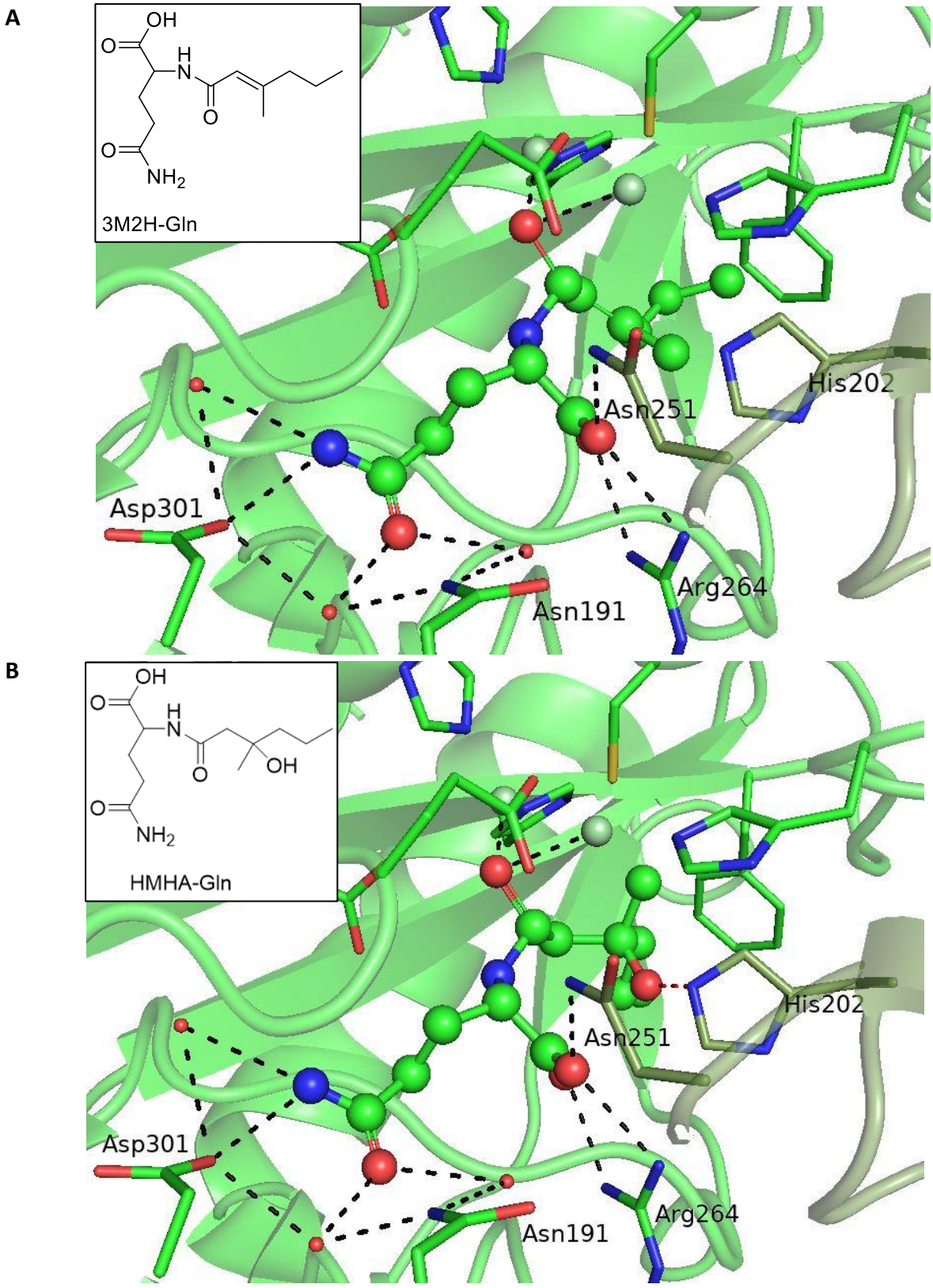


Figure 5.9.1: Building VFA precursors into N-AGA binding site

A. 3M2H-Gln built into N-AGA binding site. Waters present in LJ8 coordination of glutamine retained. N-AGA sperate monomers denoted by shading of green. Postulated interaction shown as black dashes.
B. HMHA-Gln built into N-AGA binding site. Retained water coordination of LJ8 around glutamine moiety. N-AGA monomer denoted by green shading. Postulated clashing interaction between His202 and hydroxyl group of substrate shown in red and circled.

5.10 Discussion

This chapter has centred on biochemical and structural characterisation of Oct3 and Nag3, with a focus on elucidating structure function relationships in the two proteins during VFA precursor catalysis. Initial trouble in the structural determination of Nag3 and Oct3 made the publication of the N-AGA structure co-crystallised with a competitive inhibitor fortuitous and allowed an insight into VFA precursor binding. The structure of several other M20D aminoacylases have been determined (Bitto et al., 2009; Botelho et al., 2011; Girish et al., 2013; Minasov et al., 2005) although none within this specific subfamily have been observed with a bound ligand within the binding site. The similarity of the LJ8 inhibitor allowed the structure of N-AGA to inform target residues for site directed mutagenesis and the selection of several substrate binding residues alongside conserved catalytic residues.

Both Nag3 and Oct3 were shown to be M20D aminoacylases with mutation of the catalytic Glu138 and metal coordinating Cys103 to alanine resulting in either complete abolishment or severely reduced activity with both VFA precursors. It was surprising that Oct3E138A was active given the complete abolishment observed for the same mutation in Nag3 and the designation of this enzyme as key for catalysis in M20 metalloenzymes (Rawlings and Barrett, 2013; Rawlings et al., 2018). This same mutation was shown in PhoACY to also lead to complete loss of activity (Tanimoto et al., 2008) although a Glu to Ala mutation in hACY1 led to a thousand fold decrease in activity rather than complete abolishment (Lindner et al., 2003). It therefore appears that loss of this key residue does not automatically cause complete catalytic inactivity and can instead potentially allow a very low level of activity if the catalytic water used for nucleophilic attack can be coordinated and activated by nearby residues. The catalytic glutamate is positioned next to the metal binding centre which could allow coordination of a water by the metal binding residues Glu139 and His105 to facilitate minor activity. This region also contains other nearby negatively charged residues in Glu333 and Asp78 which could fulfil the role although both are distant to Glu139 (4 and 5 Å away measuring carboxylic functional groups) which would make this unlikely without significant residue movement within the E138A mutant. No movement is observed by these residues in Nag3 E138A compared to Oct3 which would suggest this is unlikely although this could be an explanation for the difference in activity displayed between the two mutants.

The mutant Nag3 C103D showed that an aspartate can fulfil the role of bidentate zinc coordination and allow catalysis despite differences in metal binding residues between M20D proteins and those of the wider M20 family. Interestingly while this mutant did affect activity with both substrates, HMHA-Gln activity was practically abolished leading to an activity profile similar to that of Oct3. Nag3

N247A and F186A/N also resulted in much lower HMHA-Gln hydrolysis compared to 3M2H-Gln which was seen in both endpoint and kinetic assays. All these residues are too distant within the binding site to directly interact with hydroxyl group present on HMHA-Gln indicating that loss of HMHA-Gln activity in these mutants is likely due to effects on substrate positioning within the active site. Changes in substrate positioning could have a disproportionate effect on HMHA-Gln due to hydroxyl group accommodation either by steric hindrance size or specific charged interactions. Alteration in metal coordination is well known to effect M20 enzyme function but studies of its perturbation have invariably focused on change in activity with a single substrate rather than a range, preventing investigation of any substrate specific effects (Lindner et al., 2003, 2005; Tanimoto et al., 2008). Characterisation of PhoACY by Tanimoto *et al* showed that a change in metal cofactor to manganese altered substrate specificity rather than inhibit activity, although whether this was caused by a change in metal coordination, binding site shape or manganese specific was not investigated.

Nag3 N247 is also situated close to the amide bond in the precursors and mutation to alanine led to a greater reduction in HMHA-Gln hydrolysis than 3M2H-Gln. This would further suggest that substrate positioning is more stringent for HMHA-Gln release than 3M2H-Gln. This would also match the evidence that 3M2H-Gln is the preferred substrate of Nag3 and Oct3. The effect of Nag3F186A/N on activity is particularly interesting given that Phe186 has no direct role in substrate binding, with its effects upon catalytic efficiency suggesting a secondary role in orientation of substrate interacting residues such as Asn184 and Arg260. This region of the binding site interacts with the glutamine moiety of the VFA precursors and indicates that the tight binding of the amino acid moiety is key for catalysis and especially HMHA-Gln activity, likely due to effects upon the scissile bond positioning in the presence of a hydroxyl group on the acyl chain.

While much focus has been put upon glutamine moiety and amide bond binding being critical for activity and especially allowing HMHA-Gln hydrolysis, the reciprocal Phe/His165 mutations within the acyl binding pocket did influence activity. The initial hypothesis that Phe165 and His165 form part of the binding site opening in their respective proteins and may affect HMHA-Gln activity seems unlikely given the open conformation of both Anaerococcal enzymes, which creates an accessible binding cleft instead. Surprisingly Nag3F165H led to a twofold improvement in catalytic efficiency with HMHA-Gln despite initial hypothesis that the mutation would reduce efficiency. This improvement, caused by a better affinity, would suggest an effect on substrate binding is altering activity rather than the binding site opening as initially proposed. Whether the introduced histidine aids affinity by affecting other binding site residue orientation or direct substrate interaction is unclear but this beneficial role is not shared by Oct3 where mutation to phenylalanine led to minor improvement in hydrolysis of both substrates. The acyl binding pocket and its effects on various different substrate affinities has been

explored for hACY1 where a preference for short acyl chains was noted due to pocket size, with acyl branching at the α- and β-positions being particularly unfavourable due to steric clashes with binding site residues (Lindner et al., 2008). While both VFA precursors are branched at the β-position, the hydroxyl group of HMHA-Gln clearly does cause reduced affinity in both Oct3 and Nag3. Further experiments to characterise activity of both Anaerococcal enzymes with unbranched hexanoyl-Gln and 3-dimethyl hexanoyl-Gln would allow the investigation of both the effect of branching and the hydroxyl group itself upon activity. N-AGA has been shown to hydrolyse several different acylated-glutamine substrates of various lengths and branching, including other hydroxylated acyl chains, that are present in human sweat at lower concentrations than that of 3M2H-Gln and HMHA-Gln (Natsch and Emter, 2020; Natsch et al., 2006). This promiscuity was ascribed to the large acyl pocket being able to accommodate differing acyl chains but both Oct3 and Nag3 appear to share this large binding site pocket. Structural characterisation of Oct3 in a substrate bound state could reveal a more accurate picture as to binding site size during catalysis. Experimentation of VFA release from axilla sweat treated with purified Nag3 and Oct3 in a similar manner to that performed by Natsch *et al* would reveal the extent of substrate range both Anaerococcal aminoacylases have and allow a clearer picture as to their substrate specificity.

The proximity of His197 to the scissile amide bond in VFA precursors suggested a role in oxygen coordination after formation of the tetrahedral intermediate and potentially carbonyl oxygen interaction prior to hydrolytic attack. This role of carbonyl oxygen coordination has been suggested to act as an oxyanion hole in DapE, stabilising the negative charged oxygen upon formation of the tetrahedral intermediate during nucleophilic attack (Nocek et al., 2018). Mutation of this residue in both DapE and hACY1 led to almost complete loss of function with 3 % and 0.1 % activity relative to their respective WT proteins (Lindner et al., 2003, 2005; Nocek et al., 2018) which would support these residues playing such critical roles. Nag3 H197A did not display as severe a loss of activity, with catalytic efficiency reduction relative to WT of 11-fold for 3M2H-Gln and 62-fold for HMHA-Gln compared to 560-fold for hACY1 H206A with Ac-Met. This would suggest this interaction has lesser importance within Nag3, possibly with charge stabilisation being primarily achieved by the zinc cofactor instead. The positioning of this histidine within N-AGA does allow interaction with a phosphorous bound oxygen atom indicating this interaction does likely still occur. It is tempting to speculate that His197 may play a lesser, but still important, role in intermediate stabilisation in VFA releasing aminoacylases due to the accommodation of the branching acyl chains of VFA precursors in the binding site and potential clashing interaction in the case of hydroxylated acyl substrates. Given the noted adverse effects upon affinity of branching at the β-position of acyl moieties in hACY1 (Lindner et al., 2008) due to steric clashes and the prevalence of this branching in sweat derived acyl-

glutamine conjugates (Natsch et al., 2006), His197 positioning has likely adapted to facilitate activity by more than just ideal coordination of the transition state. Crystallisation of Nag3E138A bound to the VFA precursors would provide invaluable information on substrate binding and specifically acyl chain accommodation in the pocket.

Structural characterisation of Oct3 and Nag3E138A by X-ray diffraction provided much greater insight into their respective binding sites that would have been possible with solely the mutagenesis work and N-AGA. With the inherent flexibility of monomers of the M20 family due to the unstructured linking region between the two domains it is perhaps unsurprising that difficulties in crystallisation were encountered for both proteins. This is also evident in N-AGA structural determination whereby a high resolution structure was determined for an inhibitor bound state which was noted to increase protein rigidity and promote tetramer formation in solution that was later observed in the crystal (Natsch, 2015; Natsch and Emter, 2020). Subsequent data collection for both proteins also displayed generally poor data quality, necessitating structure determination at a higher resolution than that achieved for N-AGA and creating difficulties in initial structure solving by molecular replacement using the best candidate models at the time. The arrival of AlphaFold greatly enabled structural determination of Nag3E138A, with the use of a confident model of Nag3 as the candidate model (Jumper et al., 2021). The use of an AlphaFold model in molecular replacement is of particular benefit to M20 proteins due to the large family displaying such low sequence similarity to each other, even between proteins of the same function as evident with N-AGA to Oct3 and Nag3.

Both Oct3 and Nag3E138A display a similar monomer conformation that results in an open conformation in the tetrameric state in Nag3E138A which is contrasted by the closed, inhibitor bound state of N-AGA. Such an open binding site would certainly be beneficial in allowing binding of di- and tripeptides, which some members are able to hydrolyse alongside displaying aminoacylase activity (Botelho et al., 2011; Jamdar et al., 2015). Formation of a closed state must occur within Nag3 to enable the interaction of several residues that affect activity upon mutation and were shown to interact with the competitive inhibitor in N-AGA, such as Arg260, His197, Asn247, Gln295 and Asn184. Measurement of the distance of movement of these resides between Nag3 E138A and N-AGA indicates an approximate 5 Å shift of the subunit B oligomerisation domain towards the binding site to enable interaction of His197 and Asn247 with a bound substrate. This protein movement was investigated for DapE and showed a larger 10 Å movement of the oligomerisation domain residues towards the binding site while bound to product (Nocek et al., 2018). The suggestion of protein movement between an open and bound state has been made for several M20 enzymes including hACY1 (Lindner et al., 2005), an auxin amidohydrolase (Simunovic et al., 2011), a β-alanine synthase (Lundgren et al., 2007), an allantoate amidohydrolase (Agarwal et al., 2007), the dimeric dipeptidase

PepD (Chang et al., 2010) and the monomeric PepV (Jozic et al., 2002). PepV is noteworthy in being an active monomer in solution due to an effective duplication of its oligomerisation domain (termed a lid domain) to fulfil the same residue substrate interactions seen in dimeric counterparts, indicating that protein movement is not dimer dependent. Similar proposals for monomer activity and domain movement have been made for several auxin amidohydrolases such as the M20D protein BrILL2 although this protein does not share a duplication of the monomeric domain to mimic the interactions of an active dimer (Savić et al., 2009; Simunovic et al., 2011). Instead, it appears that these may have been incorrectly assumed to be monomeric due to the crystal structure of the homologue *Arabidopsis thaliana* ILL2 having a single monomer within the asymmetric unit (Bitto et al., 2009). This structure was used as the basis for protein homology modelling and subsequent molecular dynamic (MD) simulations using BrILL2 as a monomer despite both AtILL2 and BrILL2 having conservation of His197 and Asn247 and these residues being situated at the end of the oligomerisation domain, suggesting interaction with another monomer. Despite this, MD simulations did suggest that Arg coordination of the amino acid carboxylic group was the major driver of conformation change in BrILL2 which would support the inactivity exhibited by the Arg260Ala mutation in this work. The conformation change to a closed binding site implicates the oligomerisation domain size in acyl chain binding alongside the large acyl pocket found in the catalytic domain. The binding site residues close to the scissile amide bond of the substrate when closed appears to be the most likely cause of the difference in substrate specificity between Nag3 and Oct3. As noted, earlier attempts to crystallise Nag3 E138A in the presence of 3M2H-Gln and HMHA-Gln have so far failed to produce diffracting crystals but if successful would allow a greater insight into substrate binding. Structural determination of Oct3 bound to substrate would also be highly beneficial to allow direct comparison between the two although the fact that Oct3 E138A is not completely inactive may complicate crystallisation.

Using the mutant activity data discussed in this chapter alongside the solved structures of Oct3, Nag3 and N-AGA, a rational mechanism of catalytic activity can be postulated for the enzymes with the VFA precursors (Fig 5.10.1). The catalytic mechanism is shared by both substrates, with the differences in activity observed by the enzymes being caused by substrate specificity rather than catalytic inactivity with either substrate. A water molecule is coordinated between a zinc cation and a conserved glutamate. This Glu deprotonates the water molecule, activating the hydroxyl which initiates a nucleophilic attack on the carbonyl carbon. A tetrahedral intermediate is formed and stabilised by both divalent metal ions before cleavage of the amide bond releases a glutamine and VFA molecule. The mechanism would be restarted upon influx of another water molecule and efflux of the products. This catalytic mechanism is believed to occur in both *Anaerococcus* enzymes and N-AGA given the

conservation of residues and their spatial similarity and is also the likely general action of most M20D aminoacylases.

Substrate positioning to bring the carbonyl and scissile amide bond close to the zinc atoms and catalytic water would be paramount for catalysis to occur. This is achieved primarily with close binding of the glutamine moiety of the substrates with interactions with the R group, likely through water coordination achieved by Asn184 and Gln295. An essential interaction is also made between Arg260 and the carboxylic group of glutamine to allow substrate binding and catalysis. Protein movement and conformational change is likely initiated by substrate binding to give a bound, closed state whereby a second monomer's oligomerisation domain moves towards the binding site, bringing two key substrate interacting residues (Asn247 and His197) proximal to the substrate. This movement also leads to a change in the binding site from an open cleft to a closed binding site with a small opening that could play role in product efflux. Asn247 from the other monomer oligomerisation domain strengthens glutamine coordination with interaction with the carboxylic acid of glutamine. The glutamine moiety coordination allows the scissile amide bond to be close to the metal centre and enables catalysis. The tetrahedral intermediate formed during catalysis is principally stabilised by the two zinc atoms but may also involve His197 from the other monomer oligomerisation domain. This histidine could also be involved in carbonyl oxygen coordination prior to catalytic attack. In HMHA-Gln catalysis this same residue may also interact with the hydroxyl group present in the acyl moiety.

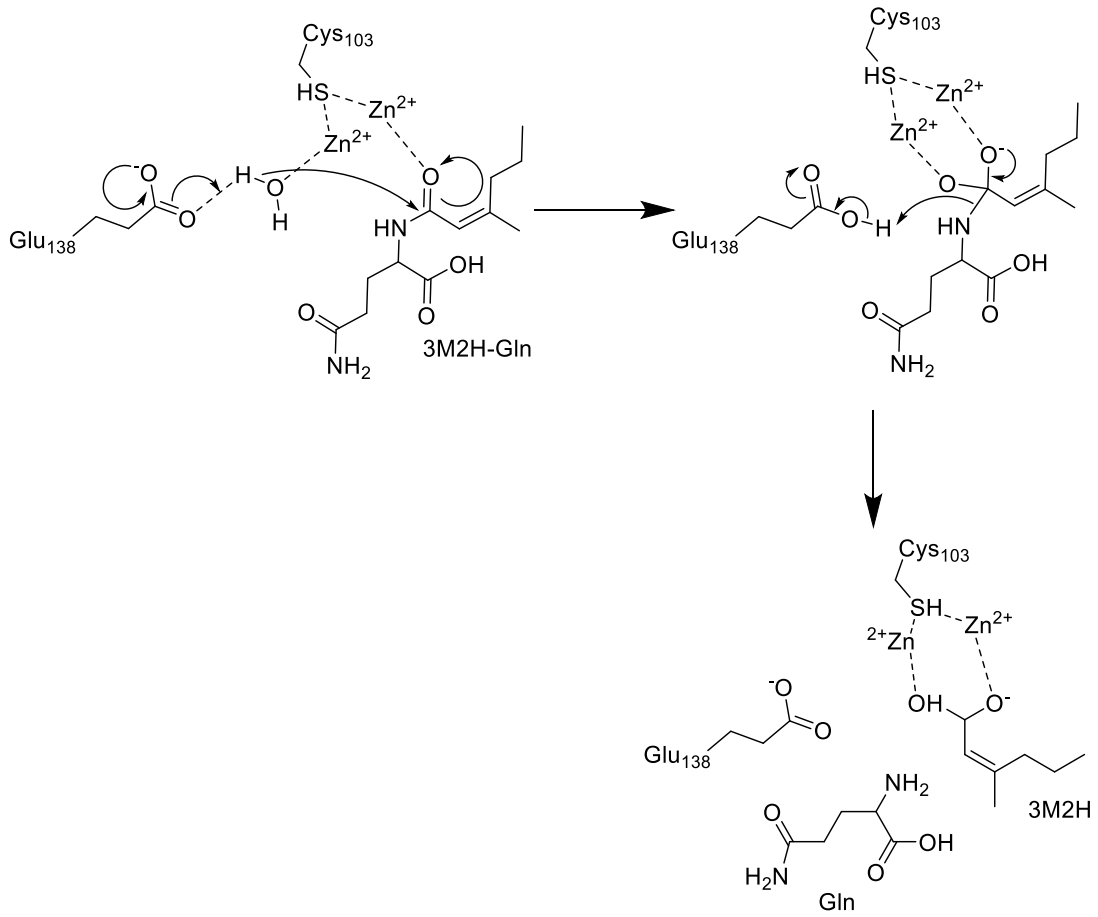


Figure 5.10.1: Proposed reaction mechanism for 3M2H-gln hydrolysis

A water molecule is coordinated between a divalent metal cation (here denoted as Zn^{2+}) and a conserved glutamate. This Glu deprotonates the water molecule, activating the hydroxyl which initiates a nucleophilic attack on the carbonyl carbon. A tetrahedral intermediate is formed and stabilised by both zinc ions before cleavage of the amide bond releases a glutamine and 3M2H molecule. Residue numbering shown are for Nag3 amino acid sequence. Catalysis functions with both Z (depicted here) and E isomers of 3M2H-Gln.

Chapter 6 – VFA production *in vivo*

6 VFA production *in vivo*

6.1 Introduction

All work detailed so far has focused on the *in vitro* characterisation of Oct3 and Nag3 and gaining insight into their action. To complement the *in vitro* findings, experiments were conducted *in vivo* to investigate VFA malodour generation in whole bacterial cell systems. The ninhydrin assay would be unsuitable for an *in vivo* assay due to the prevalence or reactive amine groups in media and cells, resulting in the use of mass spectrometry (MS) to directly detect the released VFA products. Detection of malodorous compounds by MS has been routinely used in axilla odour studies (Troccaz et al., 2004, 2015; Zeng et al., 1991) including for 3M2H and HMHA generation by N-AGA (Natsch et al., 2003, 2004). A screen of HMHA generation in a selection of Anaerococcus strains has also been published utilising GC with flame ionisation detection (GC-FID) which resulted in the identification of *A. nagyae* as a HMHA releasing microorganism (Fujii et al., 2014; Veloo et al., 2016).

Gram-positive anaerobic cocci (GPAC) have been shown to be a prevalent member of the axilla microbiome and *Anaerococcus* abundance has been linked to underarm malodour intensity (Troccaz et al., 2015), but the exact contribution *in* vivo of GPAC species to malodour is an understudied area. Aside from the aforementioned HMHA screen, no other published work has screened individual GPAC species for production of VFAs and not at all for 3M2H generation, making this a novel area of research.

Work in this chapter focuses on two main themes, first, the screening of a selection of GPAC for VFA malodour activity, followed by a secondary area of a gain of function experiment with *E. coli* expressing an active aminoacylase and its use in a VFA precursor transport screen. The screening of GPAC species for VFA activity allowed investigation into the prevalence of malodour generation within *Peptoniphilaceae*, the family containing *Anaerococcus* as well as *Peptoniphilus* which is another prevalent axilla genus that has been speculated to contribute to BO (Egert et al., 2011; James et al., 2013b; Troccaz et al., 2015). With Oct3 and Nag3 being labelled the primary VFA liberating enzymes from their respective species, the presence of an Oct3/Nag3 homologue within GPAC species was hypothesised to be a predictor of VFA malodour activity. The use of *E. coli* expressing Nag3 for screening was performed initially to validate the active aminoacylase is responsible for VFA generation in an *in vivo* setting with a gain of function experiment, given the lack of genetic tools for *Anaerococcus* making the more commonly used genetic knockout and complementation experiments unavailable. *Escherichia coli* expressing Nag3 was then used in a small, targeted experiment to try identify the transporter responsible for VFA precursor uptake by utilising *E. coli* strains deficient in one or more transporters.

6.2 Predicting VFA liberating activity via a bioinformatic approach

It was hypothesised that the presence of an Oct3/Nag3 homologue within a strain could be used as a marker for the ability to generate VFA malodour. To investigate this hypothesis a phylogenetic tree was created for putative aminoacylases within the entirety of the Peptoniphilaceae family. A Blastp search within Peptoniphilaceae using Oct3 as the query sequence was performed and all 383 sequences showing sequence similarity were collected. An additional Blastp search using N-AGA as the query sequence was also used to identify any other potential sequences which may have been missed from the Oct3 search, given the low sequence similarity (30.8 %) between the two aminoacylases. This search identified 358 sequences, including N-AGA, which were collated with the initial results from the Oct3 search before duplicates were removed to leave 395 unique sequences. The list was then further reduced to include protein sequences from a single representative genome of a taxonomically classified species. The resultant 190 sequences were then used in a MSA before generation of a maximum likelihood phylogeny tree to reveal the distribution of aminoacylases within Peptoniphilaceae (Fig 6.2.1)

The Peptoniphilaceae family is relatively understudied with most sequenced genomes belonging to the *Anaerococcus* and *Peptoniphilus* genera preventing much to be concluded for aminoacylase distribution for the family in its entirety. *Anaerococcus* sequences principally clustered with others of their genus suggesting no widespread horizontal genetic exchange occurs for aminoacylase genes. As such the sequences closely clustering within the Oct3/Nag3 clade are all *Anaerococcus* species with *A. pacaensis, A. degeneri, A. murchdochii* and *A. lactolyticus* being the closest members. Only 2 species of Peptoniphilus appear be in the vicinity of this clade with *Peptoniphilus stercorisuis* and *P. duerdenii* although both have low sequence identity to Oct3 (~35 %) especially when compared to the Anaerococcus sequences within the cluster that range from 60-80 % similarity. These two genera were the only ones to cluster close to Oct3/Nag3 within the Peptoniphilaceae family. No clustering was observed with the N-AGA sequence suggesting as proposed from the *Anaerococcus* phylogeny tree that sequence similarity to the Corynebacterial enzyme cannot be used to identify active VFA liberating enzymes from the other aminoacylase homologues.

Based on this bioinformatic analysis it was hypothesised that the four *Anaerococcus* species which contain a close homologue to Oct3/Nag3 would have the intracellular biochemical capability to generate VFAs. The lack of obvious close homologue to Oct3 or Nag3 within *Peptoniphilus* would either suggest that this genus is not able to produce VFA malodour or, more likely, that sequence similarity to Oct3/Nag3 is not an accurate predicator of activity within genera outside of *Anaerococcus*.

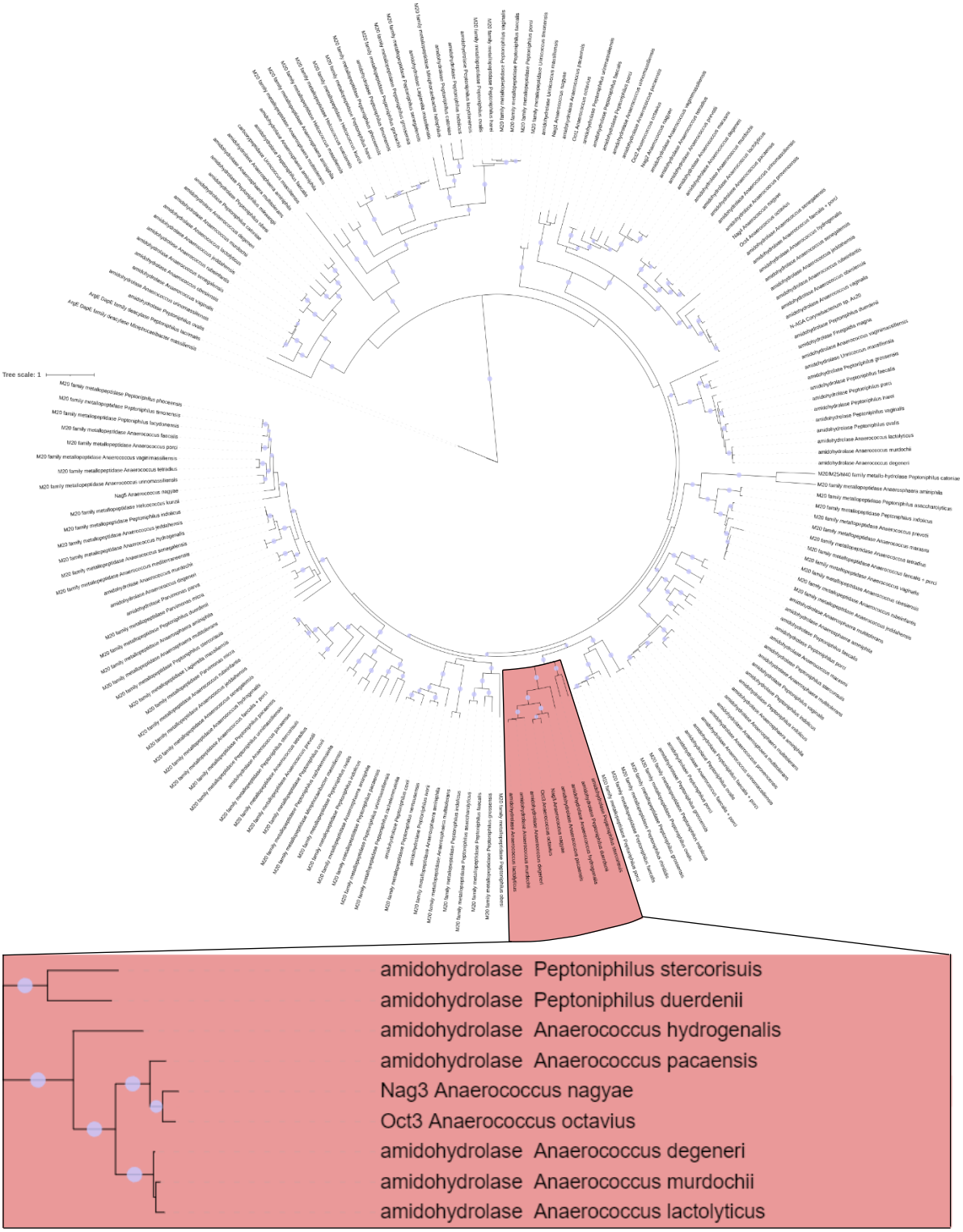


Figure 6.2.1: Aminoacylase phylogeny within Peptoniphilaceae

Sequences obtained from protein BLAST within Peptoniphilaceae using N-AGA and Oct3 as the primary sequences for the query. Multiple sequence alignment generated in Clustal Omega and used in IQ-TREE for maximum likelihood phylogeny tree generation. Circles denote node support from 1000 ultrafast bootstrap values from IQ-TREE (>90). Red colouring denotes suspected Oct3/Nag3 clade.

6.3 Procurement and preparation of strains for screening VFA biotransformation activity

The findings from the phylogenetic analysis were used in conjunction with discussions with Unilever to devise an experimental plan to screen select GPAC species for VFA biotransformation activity. This screen would focus on testing the hypothesis proposed in section 6.2 and attempt a more widespread screening for activity within Peptoniphilus. All work on Peptoniphilaceae strains, including sample generation for the activity screen, was performed at Unilever's Port Sunlight R&D site over a three-week period.

Twenty-five strains of Peptoniphilaceae were procured from DSMZ for the planned activity screening. An additional two microorganisms were added to this list from Unilever's internal strain libraries to total 27 strains for testing (Table 6.3.1). This collection of strains was comprised of 16 *Anaerococcus* isolates, 9 *Peptoniphilus* isolates, 1 *Finegoldia* strain and an *E. coli* strain to use as a negative control. *Finegoldia* defines another genus within Peptoniphilaceae, with the type species *Finegoldia magna* being a common GPAC skin isolate, that was of interest to Unilever. All DSM isolates were revived in the same manner, with colonies appearing on fastidious anaerobe agar plates after at least 2 days growth in an anaerobic chamber (37 °C) for all but one of the strains. Multiple attempts at revival of this strain, *A. tetradius* DSM 2951, failed to yield any viable growth and so could not be used in subsequent VFA screening. The other *A. tetradius* strain procured did grow suggesting the failure to revive was not due an inadequate anaerobe revival and growth methodology. The isolate *Peptonihilus* sp. 022 was obtained from Unilever's internal library and represents the only strain to have been isolated from the human axilla. Subsequent species identification of this strain was performed after the VFA biotransformation screen by Unilever and identified this strain as *P. lacydonensis*, matching the species of one of the strains procured from DSM (*P. lacydonensis* DSM 100661).

Table 6.3.1: Peptoniphilaceae VFA biotransformation assay strains		
Strain	**Isolate environment**	**additional information**
Anaerococcus octavius DSM 10024	Human nose, nasal polyp	
Anaerococcus octavius DSM 11663	Human nose	Host strain of Oct proteins
Anaerococcus hydrogenalis DSM 7454	Human faeces	
Anaerococcus lactolyticus DSM 7456	Human vaginal discharge	
Anaerococcus pacaensis DSM 26346	Human Blood	Limited biomass – 1x replicate, OD 2
Anaerococcus nagyae DSM 101193	Human blood	Host strain of Nag proteins
Anaerococcus vaginalis DSM 7457	Human ovarian abscess	
Anaerococcus vaginalis DSM 25446	Human faeces	
Anaerococcus prevotii DSM 20473	Human nose	
Anaerococcus prevotii DSM 20548	Human plasma	Limited biomass – 2x replicates
Anaerococcus tetradius DSM 2951	Human vagina	Unable to revive from DSM ampoule
Anaerococcus tetradius DSM 20358	Human abscess	
Anaerococcus senegalensis DSM 25366	Human faeces	
Anaerococcus degeneri DSM 29674	Human abdominal abscess	
Anaerococcus murdochii DSM 21462	Human foot ulcer	
Anaerococcus rubeinfantis DSM 101186	Human faeces, infant kwashiorkor	
Peptoniphilus porci DSM 104947	Pig intestinal tract	
Peptoniphilus lacydonensis DSM 100661	Human nose	
Peptoniphilus asaccharolyticus DSM 20463	Unknown	
Peptoniphilus harei DSM 10020	Human sacral sore	
Peptoniphilus gorbachii DSM 21461	Human groin abscess	
Peptoniphilus grossensis DSM 25475	Human faeces	Insufficient biomass for assaying
Peptoniphilus pacaensis DSM 101839	Human vagina	
Finegoldia magna DSM 20470	Unknown	Limited biomass – 1x replicate, OD 2
Peptoniphilus sp. 022	Human axilla	16s rRNA identified as *P. lacydonensis*
Peptoniphilus stercorisuis DSM 27563	Pig faeces	
Escherichia coli 10536	Laboratory	Negative control

6.4 *Anaerococcus in vivo* VFA biotransformation screen

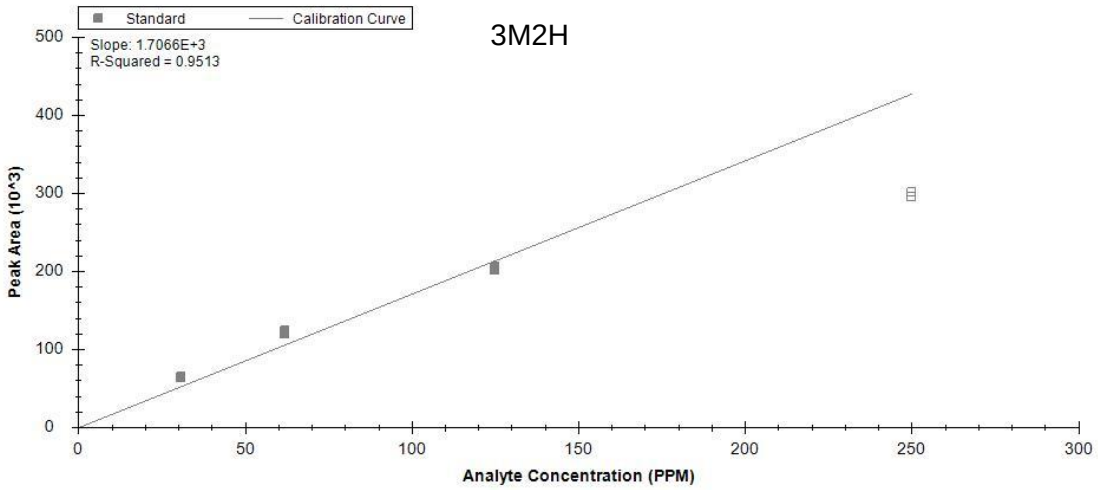
The 26 viable strains were used in a VFA biotransformation screen with both VFA precursors. Each strain was used in biological triplicate experiments with each replicate being incubated in the presence of 3M2H-Gln, HMHA-Gln or no substrate before solvent extraction and MS analysis of VFA presence. Due to the large number of samples this generated, the sample generation and MS analysis was split into batches to aid throughput.

During the biotransformation screen each strain was incubated, at a final OD 2.6, in a minimal media supplemented with 2 mM 3M2H-Gln or HMHA-Gln for 40 h under anaerobic conditions while shaken (37 °C). Upon 40 h being reached a solvent extraction was performed using ethyl acetate and hydrochloric acid to acidify any VFA present and allow its extraction into the top organic layer. An aliquot of the organic layer was then taken for analysis using a Unilever proprietary method utilising hydrophilic interaction liquid chromatography (HILIC) coupled to an ARC-QDA for sample resolution and MS detection. Mass spectrometry data was processed in Skylines with VFA levels quantified using calibration curves of known standards of the two VFA products generated at the start of each analysis run (Fig 6.4.1). Standards at 31, 62.5, 125 and 250 ppm 3M2H and HMHA were used to generate the standard curves although a clear loss in sensitivity at the highest VFA concentrations necessitated omission of the 250 ppm standards to ensure trendline linearity through zero. Alongside the solvent extraction a growth spot assay was performed on the cell incubations to verify cell viability of each strain after the 40 h incubation, allowing discernment between lack of VFA liberation activity and dead cells. For this a 1 µl aliquot of cell suspensions from each reaction was spotted onto Oxoid blood agar plates and incubated in an anaerobic chamber for 2 days (37 °C).

VFA release was observed for several Anaerococcus strains, with the highest levels being observed for *A. nagyae* with both substrates (Fig 6.4.2). *Anaerococcus pacaensis* and *A. tetradius* DSM 20358 gave the second and third highest levels of 3M2H and HMHA release respectively. Two strains of *A. octavius* were tested but only one displayed clear activity, primarily with 3M2H and matching the *in vitro* substrate specificity displayed by Oct3. The lack of activity for *A. octavius* DSM 10024 could be attributed to cell death during the incubation reaction due to the absence of growth in the cell viability spot assay, especially compared to *A. octavius* DSM 11663 (Table 6.4.1). Low levels of 3M2H release were detected for *A. murdochii, A. lactolyticus* and *A. degeneri,* which along with *A. pacaensis* all contain an Oct3/Nag3 homologue and match the hypothesis postulated from the bioinformatics analysis. Screening of *A. pacaensis* and *A. prevotii* DSM20548 could only be accomplished for single and duplicate replicates respectively due limited available biomass for assay set up after anaerobe

growth, with a further caveat that *A. pacaensis* screening could only be prepared using a lower optical density (OD 2). These strains could not be screened again due to the time constraints of the industrial placement. *Escherichia coli* 10536 showed no activity with either precursor molecules as expected in its use as the negative control strain.

A



B

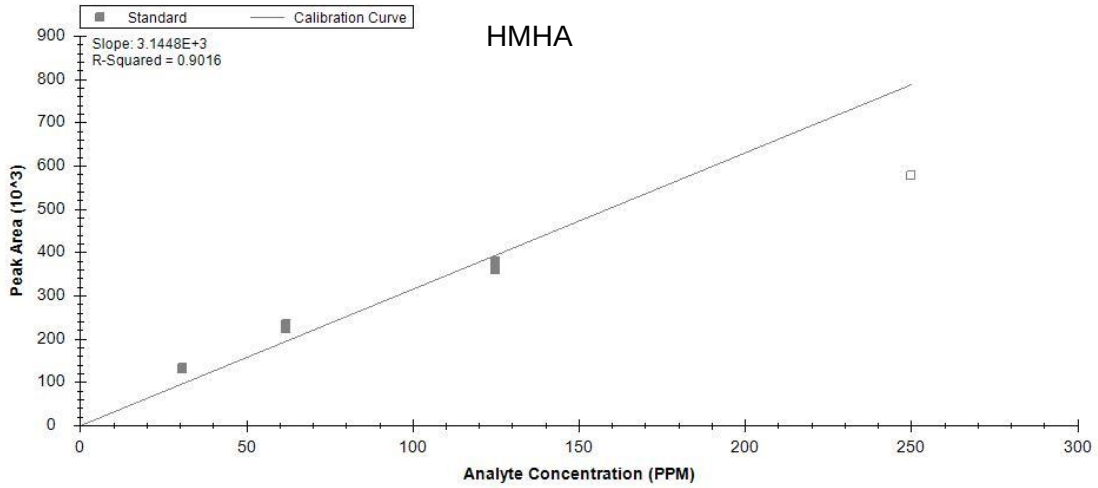


Figure 6.4.1: VFA HILIC ARC-QDA MS calibration curves

A. Calibration curve of 3M2H standards generated. Trendline uses standard concentrations 125, 62.5 and 31 ppm just to loss in accuracy at 250 ppm measurements. Solid squares show standard used for trendline generation; open square shows omitted 250 ppm standards.
B. Calibration curve of HMHA standards generated. Trendline uses standard concentrations 125, 62.5 and 31 ppm just to loss in accuracy at 250 ppm measurements. Solid squares show standard used for trendline generation; open square shows omitted 250 ppm standards

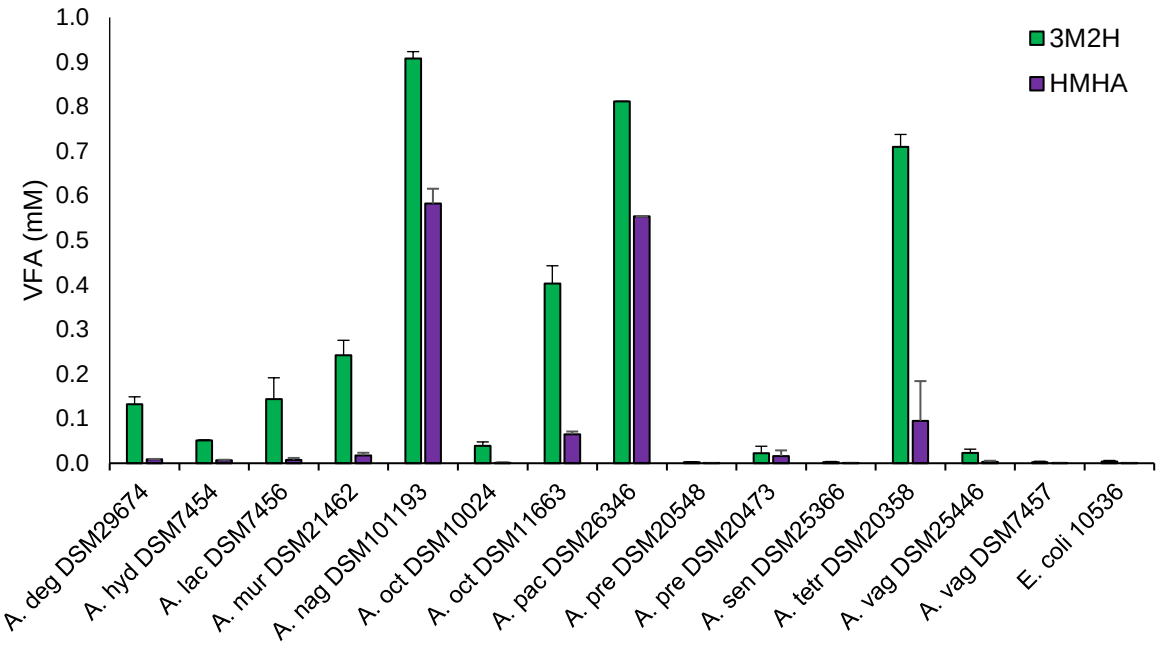
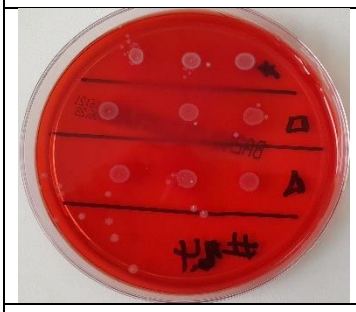
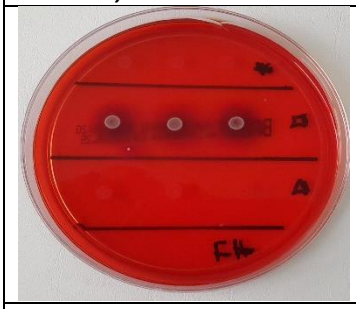
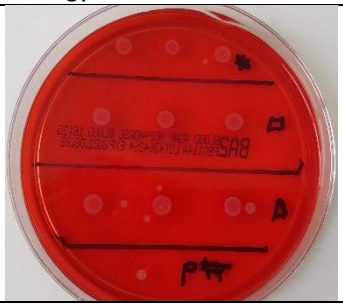


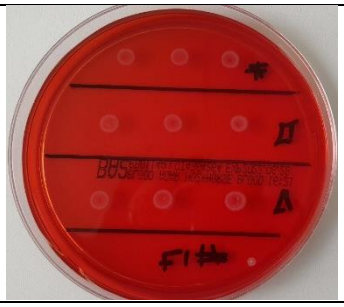
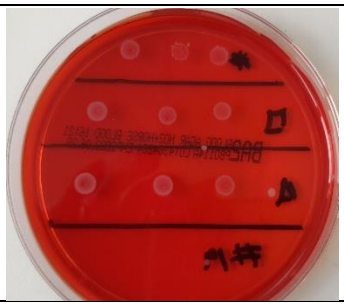
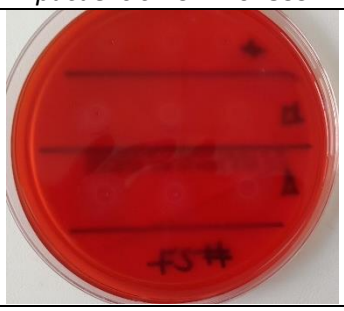
Figure 6.4.2: *Anaerococcus in vivo* VFA biotransformation screen

Screen of 3M2H and HMHA release from incubations of stated strain with respective 2 mM VFA precursors after solvent extraction and LC-MS. Quantification of VFA liberated achieved with calibration curves of each VFA. All experiments performed in biological triplicate except for *A. prevotii* DSM20548 and *A. pacaensis* which used duplicate and singular replicates, respectively, due to limited biomass. Error bars denote standard deviation.

Table 6.4.1: Peptoniphilaceae cell viability spot assay

Assay of strain viability on blood agar plates post *in vivo* VFA transformation assay. Each spot represents a separate sample of up to 9 reactions performed for each strain. Rows denote the same biological replicate while columns indicate VFA precursor each sample was incubated with (left = negative control, middle = 3M2H-Gln, right = HMHA-Gln).

		
A. octavius DSM 10024	*A. octavius* DSM 11663	*A. hydrogenalis* DSM 7454
		
A. lactolyticus DSM 7456	*A. pacaensis* DSM 26346	*A. nagyae* DSM 101193
		
A. vaginalis DSM 7457	*A. vaginalis* DSM 25446	*A. prevotii* DSM 20473
		
A. prevotii DSM 20548	*A. tetradius* DSM 2951	*A. senegalensis* DSM 25366

		
A. degeneri DSM 29674	*A. murdochii* DSM 21462	*P. stercorisuis* DSM 27563
		
P. porci DSM 104947	*P. lacydonensis* DSM 100661	*P. asaccharolyticus* DSM 20463
		
P. harei DSM 10020	*P. gorbachii* DSM 21461	*P. pacaensis* DSM 101839
		
F. magna DSM 20470	*P. sp 022*	*E. coli 10536*

6.5 Peptoniphilaceae *in vivo* VFA biotransformation screen

The exact same screening conditions were performed for the *Peptoniphilus* and *Finegoldia* strains procured, with VFA release being quantified using the same LC-MS method.

During assay preparation a single Peptoniphilus strain, *P. grossensis* DSM 25475, failed to grow adequately enough for use in the screen preventing its screening. Additionally, *Finegoldia magna* DSM 20470 grew poorly and resulted in only a single replicate being screened at a lower OD than that used in other incubations (OD 2 instead of 2.6). All other strains tested were tested in biological triplicate for both VFAs.

Only *P. lacydonensis* 022 displayed significant VFA release to comparable levels as the highest Anaerococcus strains tested (Fig 6.5.1). This strain was identified post screening as *P. lacydonensis* and shows a striking contrast to the *P. lacydonensis* DSM 100661 strain which displayed no activity with either VFA precursor and cell viability spot tests confirmed the presence of live cells post VFA incubation (Table 6.4.1). This indicates clear strain differences within the species for VFA generation, which could be attributed the isolate environments of the two strains with DSM 100661 being recovered from a human nasal swab while *P.* sp 022 was from the human axilla. The matching of VFA activity to the isolate recovered from the axilla would suggest that native environmental exposure to VFA precursors has caused strain specific adaptation to facilitate high activity, although it is surprising that DSM100661 appears completely inactive given both should have the biochemical capability to hydrolyse 3M2H-Gln and HMHA-Gln. Both *P. porci* and *P. pacaensis* did generate small amounts of principally 3M2H although to a much lesser degree than *P. sp* 022 or many of the Anaerococcus strains. This screening has shown that VFA malodour generation is not a widespread trait of the entire genus and instead appears to be limited to a select few strains.

Bioinformatic analysis placed *P. stercorisuis* and *P. duerdenii* as containing the closest Peptoniphilus homologues to Oct3/ Nag3 but the *P. stercorisuis* strain tested showed negligible activity with both precursors. Hence it appears that Oct3/Nag3 homology searching is insufficient to identify VFA malodour generating species outside of the Anaerococcus genus.

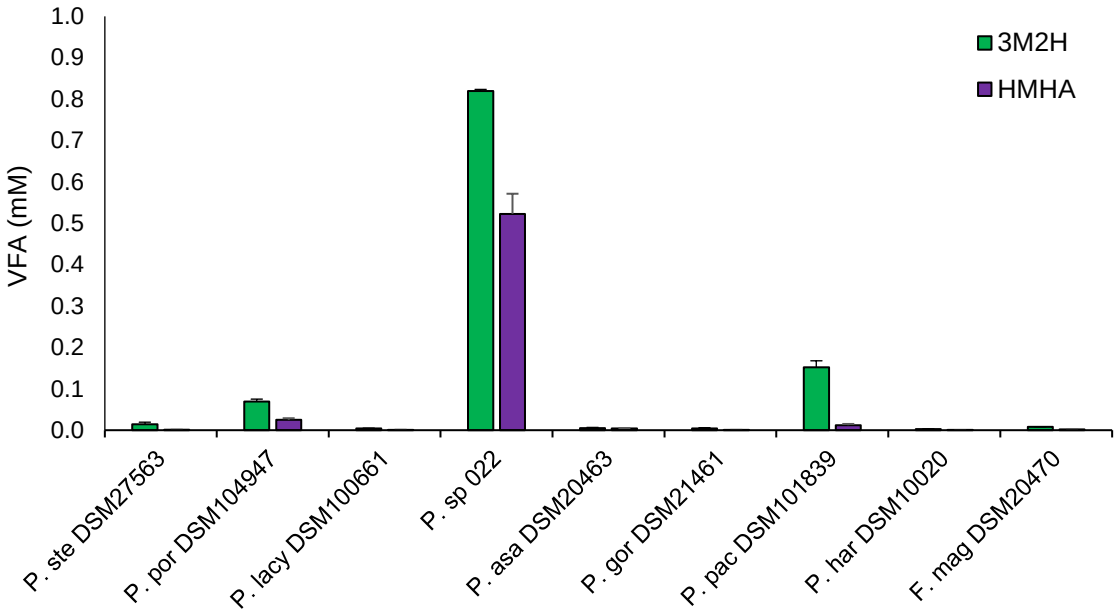


Figure 6.5.1: *Peptoniphilus* and *Finegoldia in vivo* VFA biotransformation screen

Screen of 3M2H and HMHA release from incubations of stated strain with respective 2 mM VFA precursors after solvent extraction and LC-MS. Quantification of VFA liberated achieved with calibration curves of each VFA. All experiments performed in biological triplicate except for *F. magna* where a single replicate was screened due to limited biomass. Error bars denote standard deviation.

6.6 MS detection from *in vitro* reactions

All current *in vitro* characterisation of aminoacylases has involved detection of the amino acid product from substrate hydrolysis rather than direct detection of the VFA products. To verify that the active VFA liberating enzymes are successfully generating the proposed VFAs, *in vitro* incubation of the active aminoacylases were analysed by the same Unilever MS system for detection and quantification of 3M2H and HMHA.

Incubations of 2.5 µM Oct3, Nag3 and N-AGA with 2 mM 3M2H-Gln or HMHA-Gln were prepared and incubated for 1 h at 37 °C aerobically, using the 2 ml reaction volume used during the *in vivo* experiments. After 1 h timepoint a 0.9 ml aliquot of the reaction was solvent extracted into ethyl acetate to generate samples for MS analysis. VFA levels were then detected and quantified from calibration curves of 3M2H and HMHA standards. Due to time constraints with use of the Unilever MS system, sample generation for this experiment was performed once in triplicate using the only available enzyme stocks at the time.

All three aminoacylases were observed to produce 3M2H after incubation with its glutamine conjugated precursor, validating that they do produce the volatile fatty acid they were posited to (Fig 6.6.1). However, HMHA detection appeared lower or, in the case of Nag3, absent than expected from previous *in vitro* characterisation of the enzymes. The amount of 3M2H detected with Nag3 incubations was also strikingly low compared to typical activity levels. While the substrate concentration used was lower than that used in the standard ninhydrin screen (2 mM rather than 2.5 mM) the most parsimonious explanation for this activity drop is degradation of the enzyme stocks used in the assay. Neither VFA was detected in substrate only controls indicating enzymatic action is required for VFA release rather than spontaneous chemical hydrolysis.

This *in vitro* experiment has verified the activity of N-AGA, matching previous literature (Natsch et al., 2003), and shown 3M2H generation for the two *Anaerococcus* orthologues. However, a repeat of this experiment would be necessary for Nag3 to validate its ability to generate HMHA as observed in the ninhydrin assays. Likewise, some level of activity for Oct3 with HMHA-Gln would have been expected based on previous *in vitro* characterisation making a repeat experiment for this enzyme a necessity. As mentioned earlier though the time limitation upon the Unilever placement and specifically use of the LC-MS equipment prevented these repeats from being conducted.

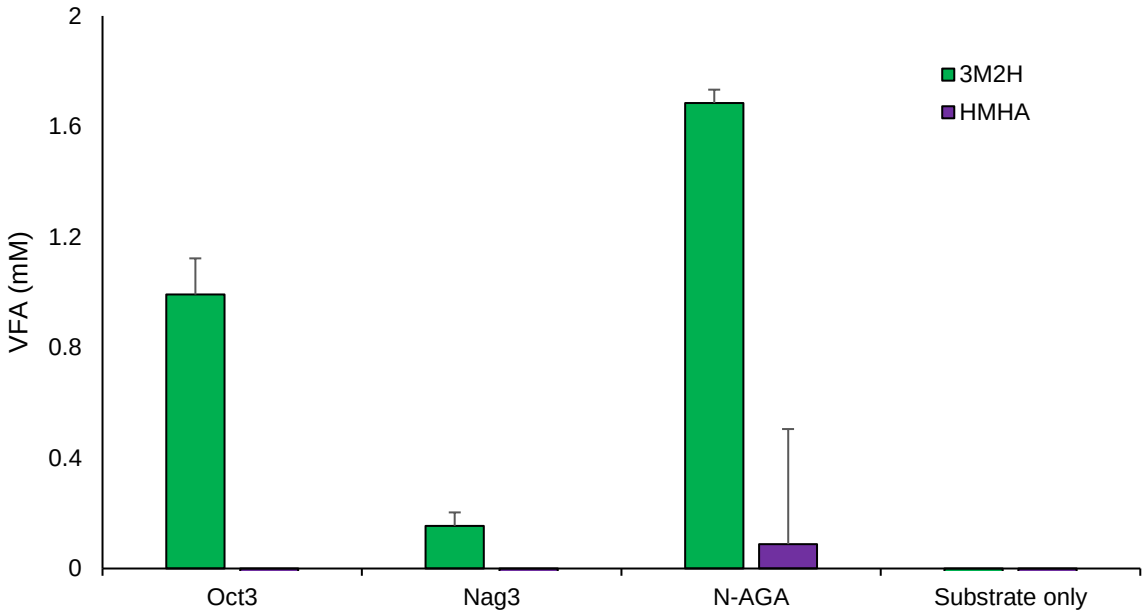


Figure 6.6.1: *In vitro* aminoacylase VFA MS detection

Screen of 3M2H and HMHA release from incubations with malodour generating aminoacylases with their respective 2 mM VFA precursors after solvent extraction and LC-MS. Quantification of VFA liberated achieved with calibration curves of each VFA. Error bars denote standard deviation (n=3).

6.7 *E. coli* VFA generation gain of function experiment

The lack of genetic manipulation tools for *Anaerococcus* prevents the use of a genetic knockout and complementation experiment to validate that Oct3/Nag3 are the principal causers of VFA malodour. To address this a gain of function experiment was instead attempted in *E. coli* due to its ease of manipulation and continuous use throughout the project.

The Peptoniphilaceae *in vivo* biotransformation screen clearly showed that wild type *E. coli* is not able to hydrolyse either precursor (Fig 6.4.2). The same experiment as that performed in sections 6.4 and 6.5 was conducted with *E. coli* BL21 (DE3) containing pET-21b(+) encoding Nag3 and the same strain without the plasmid. Both were grown in the presence of 0.5 mM IPTG aerobically before being used in the same biotransformation assay as detailed previously, but without the anaerobic requirements.

Escherichia coli BL21 (DE3) not expressing Nag3 matched the result for *E. coli* 10536 with no VFA release from either precursor after 40 h incubation (Fig 6.7.1). Expression of Nag3 showed clear detection of both 3M2H and HMHA indicating a gain of function for the *E. coli* strain and indicating that Nag3 can generate HMHA, something that could not be shown in the previous *in vitro* MS experiment. This result would also support the conclusion that protein degradation was the cause for lack of activity observed in those experiments given the high activity displayed in this *in vivo* context. The concentration of VFA detected showed a similar pattern of activity displayed as to Nag3 *in vitro* ninhydrin assay characterisation and all tested GPAC strains with 3M2H release being higher than HMHA, suggesting that *E. coli* can efficiently take up both precursors. In fact, VFA release was the highest observed for any species tested *in vivo* with the full hydrolysis of 3M2H-Gln (2.06 mM) and approximately 50% of HMHA-Gln (0.97 mM). This can likely be attributed to Nag3 expression from the use of the pET-21b vector that has consistently given high yield in this strain and will have likely surpassed the expression levels of any native aminoacylase protein in the Peptoniphilaceae strains which would be constrained by host expression feedback systems.

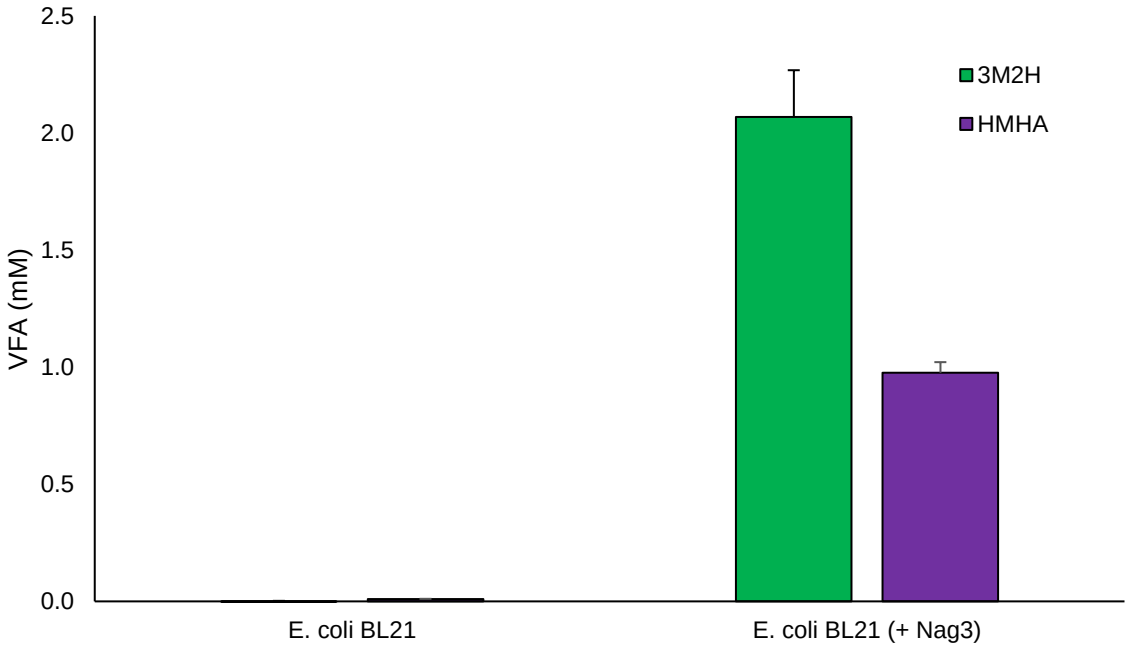


Figure 6.7.1: *E. coli* VFA malodour gain of function

Screen of 3M2H and HMHA release from incubations of *E. coli* BL21 (DE3) with respective 2 mM VFA precursors after solvent extraction and LC-MS. Quantification of VFA liberated achieved with calibration curves of each VFA. Experiments performed in biological triplicate. Error bars denote standard deviation (n=3).

6.8 Screen for *E. coli* VFA transporter

The ability of *E. coli* expressing Nag3 to generate VFA malodour offers a potential tool to identify the transporter responsible for precursor uptake in *E. coli* with the use of genetic mutants developed for *E. coli*. Such a level of genetic manipulation is not available for *Anaerococcus*, necessitating the use of *E. coli*. Discovery of the transporter(s) responsible for VFA precursor uptake in *E. coli* would hopefully enable the identification of homologous uptake systems in *Anaerococcus* which would become the most likely transporter(s) responsible for precursor uptake. Such identification would have to be caveated with the vast differences between *E. coli* and *Anaerococcus* species, most importantly including the difference in cell membrane between the Gram-negative Enterobacteriaceae and the Gram-positive Peptoniphilaceae.

Escherichia coli expressing Nag3 exhibits a gain of VFA generating activity which was hypothesised is a phenotype that could be lost upon deletion of the transporter responsible for substrate uptake. Therefore, screening of *E. coli* strains deficient in the necessary transporter(s) would exhibit an abolishment, or potentially lowering in the case of multiple uptake systems, in VFA generation compared to wild type *E. coli* when both are expressing Nag3 and supplied with VFA precursors. While no evidence points to the idea that uptake is mediated by multiple transporters, it could be a possibility which could be missed in this screen depending on the level of redundancy.

The screen utilised the *E. coli* KEIO collection (Baba et al., 2006) for transport gene deletion strains to screen, alongside two *E. coli* strains developed and used in the Thomas lab that contain multiple transport gene deletions, named DB1 and DB50 (Bawdon, 2014). Focus was placed upon amino acid and peptide transporters due to their wide prevalence and characterisation in *E. coli*, with examples of highly efficient uptake systems and more promiscuous scavenging systems. Fatty acid uptake systems are less well characterised but in *E. coli* is primarily mediated by the outer membrane protein (OM) FadL and the inner membrane (IM) associated fatty acyl-CoA synthetase FadD which removes the IM associated fatty acid and concomitantly catalyses the thioesterification to fatty acyl-CoA. Such a process appears highly unlikely to occur for either VFA precursor due to the presence of hydrophilic glutamine and the amide bond, making the fatty acid carboxylic group unavailable for interaction with coenzyme A. A small list of transport mutants was selected, representing the most likely routes for uptake of glutamine containing molecules, for screening with the following amended genotypes from an *E. coli* BW25113 parental strain: wildtype (WT), *ΔglnP::kan, ΔglnQ::kan, ΔgltP::kan,* DB1 (*ΔdppC, ΔoppC*) *,* DB50 (*ΔdtpA-D*). The mutants *glnP* and *glnQ* have deletions in the transmembrane and ATPase subunits of the glutamine ATP-binding cassette (ABC) transporter respectively. The *gltP* strain

has a deletion of the glutamate/aspartate-proton symporter. The DB1 strain is deficient in the transmembrane subunits of the dipeptide and oligopeptide ABC transporters, *dppC* and *oppC*. DB50 has had all four of its proton-dependent oligopeptide transporters (POTs) removed.

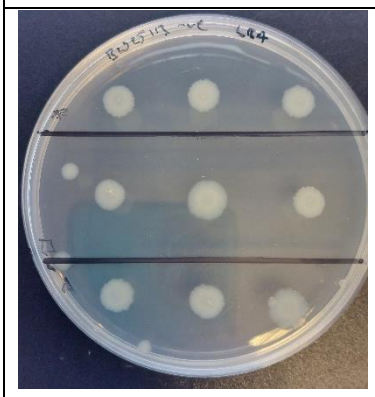
As the *E. coli* background of these mutants does not contain the DE3 prophage to allow T7 polymerase transcription, Nag3 was cloned into a pBADcLIC vector under arabinose control (Appendix A)**.** This plasmid was then transformed into each of the six strains listed earlier and used with an untransformed wildtype *E. coli* BW25113 strain to act as a negative control. Each of the seven strains were screened in the same *in vivo* VFA biotransformation screen (under aerobic conditions) as used previously with solvent extraction of potential VFA products into ethyl acetate. All strains grew sufficiently to enable biological triplicate experiments to be performed with each strain.

Due to limited access for the Unilever HILIC-MS system used for all VFA product detection a local MS system at York was used in its place, with a chromatography method and MS analysis performed by Swen Langer in the York Technology Facility. VFA were analysed by GC-MS using a Zebron ZB-5HT inferno column for chromatographic resolution before molecule detection by secondary-ion mass spectrometry (SIMS). The pure VFA standards used at Unilever were not available for this work at York, resulting in standards being created from solvent extractions of *in vitro* incubations of Nag3 with the VFA precursors similar to that collected in section 6.6. Interestingly, the standards generated did detect 3M2H and HMHA from incubation with their respective substrate (Fig 6.8.1), further supporting the theory of Nag3 degradation in previous *in vitro* MS detection. Due to the impurity of these standards and that this screen is looking for a relative loss in VFA levels between wildtype expressing Nag3 and mutant strains, exact quantification of VFA levels using calibration curves was not performed. Instead, results were collected as the integrated peak area of the identified VFA molecules. All samples were analysed in a single GC-MS run removing the possibility of inter-run differences. All strains were also tested for viability after incubation using a spotting viability assay on LB plates and showed clear growth (Table 6.8.1).

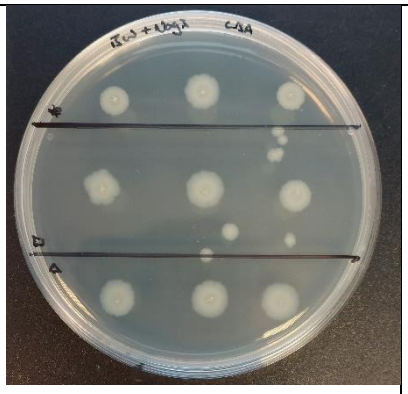
Screening showed clear production of VFAs by *E. coli* BW25113 expressing Nag3, matching that seen previously for expression in the BL21 (DE3) strain and verifying protein expression in the pBADcLIC vector (Fig 6.8.2). All six transport deletion strains still showed some level of VFA production with both precursors, suggesting that none of the deleted transporters are solely responsible for substrate uptake. A significant difference was observed for the DB50 strain with both 3M2H and HMHA compared to WT expressing Nag3, while DB1 displayed only significant difference with HMHA. Structural similarity of precursors would suggest that both are translocated by the same transport system making the result of DB1 surprising and could warrant further replication of this result to verify.

The reduction in activity observed for both VFAs with DB50 would suggest that in *E. coli* their uptake is partially mediated by a POT transporter. Further experimentation with single POT deletion strains, which have been developed within the Thomas lab (Bawdon, 2014), would enable identification of which transporter(s) is responsible. As evidenced by the presence of activity, there appears to be clear levels of redundancy for VFA substrate uptake in *E. coli* which may be present in *Anaerococcus* as well. Although as *E. coli* does not contain the biochemical capability to generate these VFAs and is not typically exposed to the precursors it could be possible that *Anaerococcus* has a more specific transporter responsible that is not found within *E. coli*.

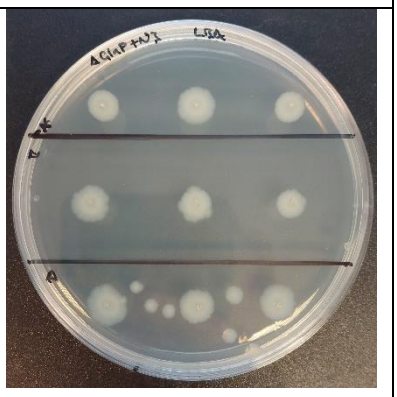
Table 6.8.1: *E. coli* transport screen cell viability test



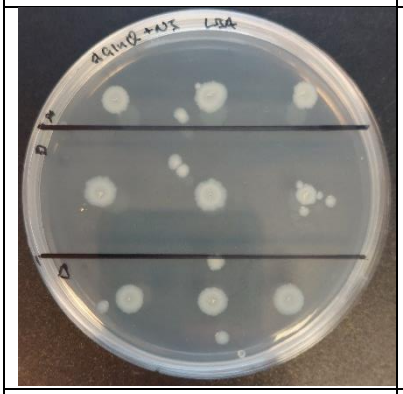
E. coli BW25113



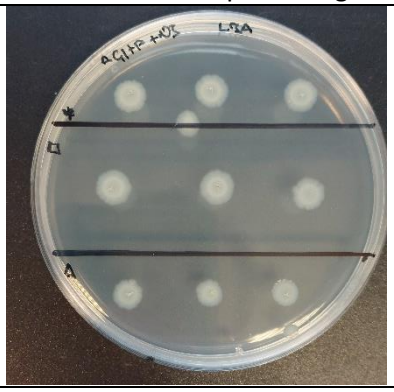
E. coli BW25113 + pBAD-Nag3



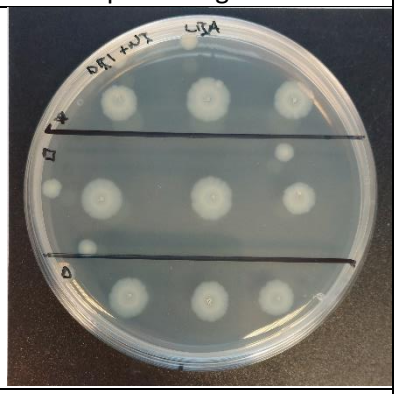
ΔGlnP + pBAD-Nag3



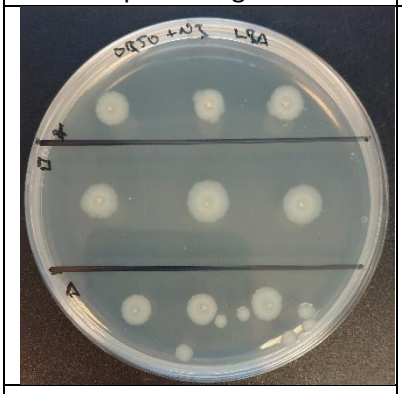
ΔGlnQ + pBAD-Nag3



ΔGltP + pBAD-Nag3



DB1 + pBAD-Nag3



DB50 + pBAD-Nag3

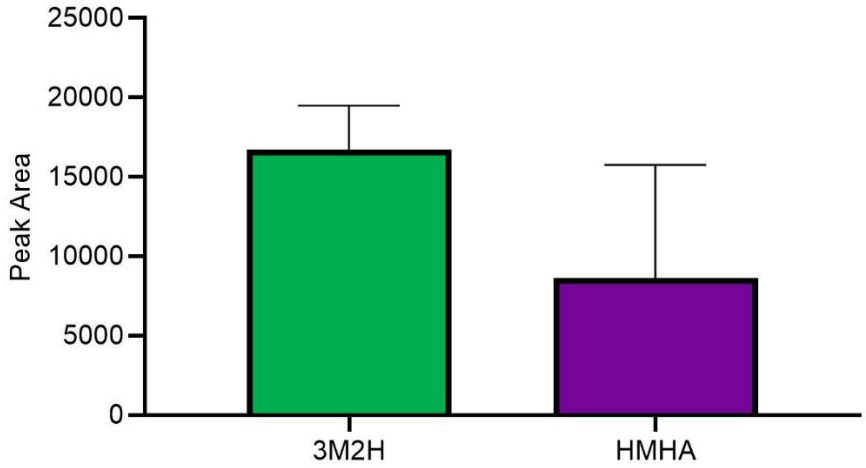


Figure 6.8.1: *In vitro* Nag3 VFA GC-MS detection

Detection of 3M2H and HMHA from *in vitro* incubations of Nag3 with respective glutamine conjugated precursor. VFAs quantified based on peak area of detected VFA molecules after GC resolution. Error bars denote standard deviation (n=6).

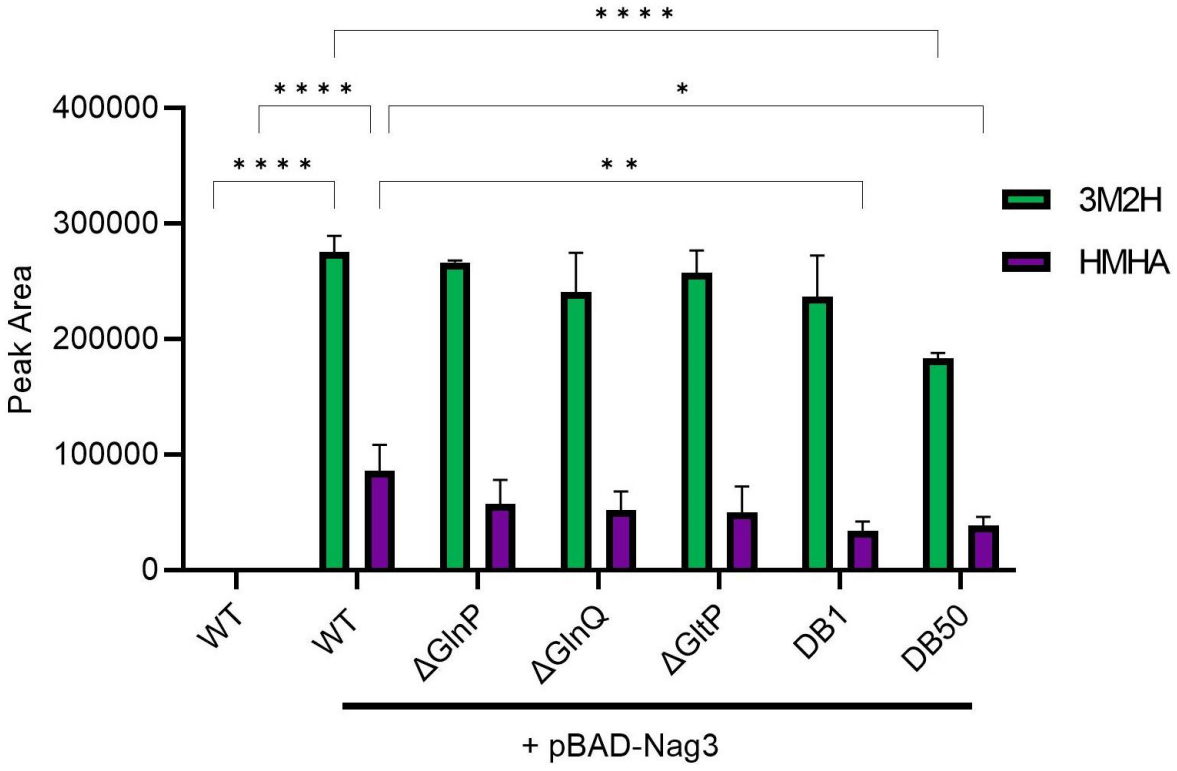


Figure 6.8.2: *E. coli* transporter mutant screen

Screen of 3M2H and HMHA release from incubations of *E. coli* strains with respective 2 mM VFA precursors after solvent extraction and GC-MS. VFAs quantified based on relative peak area of detected VFA. Asterisks show statistical analysis from two-way ANOVA comparing to WT + pBAD-Nag3 for each VFA. Error bars denote standard deviation (n=3).

6.9 Discussion

This chapter has focused on the use of mass spectrometry to directly detect the generation of VFAs and enabled investigation in an *in vivo* setting to complement *in vitro* characterisation.

Bioinformatic analysis for Oct3/Nag3 homologues within Peptoniphilaceae revealed a clade of *Anaerococcus* species that were hypothesised to be able to generate VFA malodour due to their sequence similarity to the VFA liberating aminoacylases. This trend was observed upon screening of *Anaerococcus* strains with activity observed with *A. nagyae, A. octavius, A. pacaensis*, *A. murdochii, A. lactolyticus* and *A. degeneri* which all contain a sequence in the supposed Oct3/Nag3 malodour clade. Both *A. nagyae* and *A. octavius* showed noticeable levels of activity that also matched the substrate preference for 3M2H-Gln hydrolysis seen from *in vitro* characterisation suggesting that Oct3 and Nag3 are the likely major drivers of VFA generation and that the precursor uptake mechanism in *Anaerococcus* displays no major bias for either molecule. As seen from *in vitro* characterisation, *A. nagyae* showed higher activity than *A. octavius* which would implicate this strain as being a bigger contributor to axilla malodour generation in a cell-to-cell comparison. *Anaerococcus pacaensis* showed a very similar activity profile to *A. nagyae* and contains the closest homologue to Oct3 and Nag3 within the clade (78 % identity to Oct3 and 77 % to Nag3) clearly supporting the hypothesis that this is a clade of malodour liberating aminoacylases. To date only one other screening of *Anaerococcus* strains for VFA malodour generation has been performed but only for HMHA release and using cell lysate rather than a whole cell assay (Fujii et al., 2014). That study did show the highest % HMHA release was observed with *A.* sp A20, later identified as *A. nagyae* (Veloo et al., 2016), at 79.8 % while *A. octavius* DSM11663 (also tested in this work) showed only 1.3 %, broadly matching results seen here. These two species represented the first and second highest HMHA activity levels from a total of 8 *Anaerococcus* species screened by Fuji *et al* and would support the hypothesis that HMHA release is less common than that of 3M2H. A surprising result was the high activity displayed by *A. tetradius* despite it not containing a close homologue within the supposed malodour clade. This strain contains 5 putative aminoacylases within its genome, like *A. nagyae,* with one sequence clustering close to Nag5 in the Peptoniphilaceae aminoacylase phylogeny tree. As Nag5 was shown during *in vitro* screening to be able to hydrolyse both precursors it could be possible for the homologous sequence in *A. tetradius* to be responsible for the observed *in vivo* activity. This clearly shows though that sequence similarity to Oct3 and Nag3 cannot be used as a definitive predictor of malodour generation. This is already evident within *Corynebacterium* where even strains containing close homologues of N-AGA fail to generate anything close to the amount of VFAs released by *C.* sp. Ax20 (James et al., 2013b).

Screening of *Peptoniphilus* species from DSMZ revealed a general lack of activity for both VFAs with very low levels of activity being observed for *P. porci* and *P. pacaensis* that would have suggested that this genus is not important for VFA malodour. Fortuitously the testing of an axilla isolate from Unilever's collection did show a high level of 3M2H and HMHA release from a strain that was subsequently identified as *P. lacydonensis,* making it the same species as one of the DSMZ procured strains (*P. lacydonensis* DSM100661) which gave no activity with either substrate. As both strains appeared viable before and after incubation with VFA precursors it would suggest that there are significant adaptations within the axilla isolate to cause this gulf in activity. Whole genome sequencing of the axilla isolate would allow comparison with the genome present on Genbank for DSM 100661 for any genomic differences that could explain this difference and provide a novel area of future research. There are two putative aminoacylases from the genome of *P. lacydonensis* DSM100661 present in the phylogeny tree with one clustering close to Nag5 (identity of 73 %) which, like as speculated for *A. tetradius*, could suggest that this protein is responsible for VFA liberation. Alignment of both these sequences to Oct3, Nag3 and Nag5 reveals no major alteration of key residues indicating that the inactivity of DSM100661 does not appear to be due to the protein being inactive, although this cannot be completely ruled out without *in vitro* characterisation of the enzymes. Previous screening of *Anaerococcus* isolates and *C.* sp A20 for VFA release (Fujii et al., 2014; Natsch et al., 2003) would suggest that VFA liberating aminoacylases are expressed constitutively given that VFA release was observed from cell lysates without prior incubation with VFA precursors. Hence to have such an abolishment it is likely that DSM100661 contains mutations that cause protein inactivity and/or disruption of gene transcription. The genomic context of each aminoacylase gene in its host species provides little evidence as to transcriptional regulation with no obvious operon organisation for the genes, necessitating further experimentation to determine the cause in differing activity. Future experiments could include western blot analysis with primary antibody targeting native aminoacylases to verify protein expression, although care would have to be taken with potential cross reactivity to homologous aminoacylases both inter and intraspecies. Alternatively, protein LC-MS of lysed cells could be performed to detect the relative abundance of aminoacylase proteins within samples. The single axilla strain tested also demonstrated that screening of axilla isolates would likely provide a much more complete picture of GPAC contribution to axilla VFA malodour generation. High levels of variability in malodour production have already been observed with thioalcohol generation in strains of *S. hominis,* even between axilla isolates (Bawdon et al., 2015). It therefore seems likely that malodour production within *Peptoniphilus* and *Anaerococcus* could have similar levels of inter-species complexity, making it difficult to confidently ascribe species as positive and negative VFA malodour generators.

The same LC-MS system utilised for Peptoniphilaceae screening was also used upon *in vitro* incubations of Oct3, Nag3 and N-AGA with both 3M2H-Gln and HMHA-Gln to allow direct detection of VFAs and validate these enzymes do catalyse the reactions proposed. Generation of 3M2H was observed for all enzymes with N-AGA displaying release of both 3M2H and HMHA, matching previous MS characterisation of this enzyme and host species (Natsch et al., 2003, 2004). While the lack of activity with HMHA-Gln was expected for Oct3 given an incubation of just 1 h, mirroring that performed for Oct3 from *in vitro* ninhydrin VFA screening, it was surprising that Nag3 also displayed this. Combined with the low release of 3M2H Nag3 displayed it was concluded that the enzyme stock used had likely been affected by degradation. Later use of a GC-MS system at York was developed using standards generated from *in vitro* incubation of Nag3 with VFA precursors and standards used throughout consistently detected both 3M2H and HMHA, validating that Nag3 can hydrolyse both substrates and supporting prior assumption of Nag3 stock degradation.

Peptoniphilaceae screening used a strain of *E. coli* as a negative control, which did prove unable to hydrolyse either precursor molecule during incubation. This led to the use of *E. coli* BL21 (DE3) expressing Nag3, a strain typically used in recombinant expression, in a similar *in vivo* VFA screen to show a gain of function with the ability to release both VFAs from their respective substrates. The success of this allowed use of this organism in a screen for the transporter responsible for precursor uptake in *E. coli* using transport deficient strains in an *E. coli* BW25113 background. Five transport deletion strains were tested *in vivo* focusing on glutamine and peptide transporters, but none displayed absolute abolishment of VFA generation indicating they are not responsible solely for precursor uptake. The ABC transporter GlnPQ was targeted due to its ability to uptake free glutamine as well as evidence of uptake of asparagine and glutamate via an atypical four substrate binding domain proteins (Fulyani et al., 2013, 2016), suggesting a promiscuity in substrate binding which could have extended to acylated-glutamine molecules but this strain showed no significant decrease in VFA levels upon screening. GltP is a cation coupled glutamate/ aspartate symporter that is a member of the dicarboxylate/ amino acid cation symporter family (Rahman et al., 2017) which has no prior evidence of glutamine uptake and screening revealed it has no effect of VFA generation. The size of HMHA-Gln and 3M2H-Gln is similar to that of a di- and tripeptide which could implicate a peptide transporter as responsible for uptake. Screening of DB1 with deletions in a transmembrane subunit of the ABC dipeptide and oligopeptide transporters, DppC and OppC, did show a reduction in HMHA release but not 3M2H. While this result would seem promising, if surprising to only be effecting one VFA, further exploration of the literature notes that translocation in these transporters requires a positive charge at the N-terminal amino group which is lost upon acetylation and prevents transport

(Payne, 1974; Smith et al., 1999). Therefore, Dpp and Opp should be unable to effect VFA generation, throwing doubt on the validity of this result for DB1.

DB50 is deficient in four POT transporters (DtpA-D), which provide another mechanism of peptide import. This family has been observed to exhibit high levels of substrate promiscuity (Ito et al., 2013), with this promiscuity being suggested to be due to the presence of specificity pockets that can accommodate a range of chemical properties displayed by potential substrates (Newstead, 2017). The uptake of extensively modified peptides and amino acids has been observed by this family in eukaryotes with the uptake of prodrugs conjugated to amino acids or peptides by the human PepT1 and PepT2 (Smith et al., 2013). Study of rabbit PepT1 substrate specificity also revealed high promiscuity for di- and tripeptides and affinity for *N*-acetyl-L-phenylalanine (Meredith et al., 2000), indicating the ability of this family to translocate acylated amino acids. Closer examples in bacteria for transport of modified peptides have also been observed, including for axilla malodour generation with uptake of Cys-Gly-3M3SH for thioalcohol generation in *S. hominis* by the POT transporter $PepT_{SH}$ (also known as DtpT) (Minhas et al., 2018). While the conjugation of the aliphatic alcohol moiety within Cys-Gly-3M3SH is different to that seen for the acyl group of the VFA precursors, it clearly indicates an adaptation of this transporter to accommodate a large, generally hydrophobic group. This would suggest that a POT transporter could mediate VFA precursor import and screening of DB50 did show a reduction in release of both VFAs. The lack of abolishment in VFA release points to multiple mechanisms of import that must include a transporter not tested in this screen. Single POT deletion strains have already been generated (Bawdon, 2014) and screening of these four would allow identification of those responsible for translocation. While this evidence does point to uptake being mediated, at least partially, by a POT transporter in *E. coli* this does not guarantee an orthologous system within *Anaerococcus*. In fact, use of the predictive transport protein pipeline database TransportDB 2.0 (Elbourne et al., 2017) shows that *A. prevotii,* the only *Anaerococcus* genome annotated in its database, lacks any POT transporters. However, BlastP searches for homologues to each *E. coli* Dtp within *Anaerococcus* does reveal one predicted POT transporter in *A. octavius* with highest similarity to DtpB (23.5 % identity). Interestingly this protein is not common for the genus with only a couple of unspecified *Anaerococcus* strains also containing a homologue. The absence of *A. nagyae*, or any of the identified VFA releasing GPAC species, would suggest that this protein is not involved in precursor import in *Anaerococcus* and implicating another mechanism.

Failure to identify all transporters required for *E. coli* VFA biotransformation in this screen warrants further exploration of this area. The activity displayed by the positive control *E. coli* strain expressing Nag3 indicates that the transporter is expressed under the conditions used in the *in vivo* screen, although the potential of modulation of transporter expression should not be entirely discounted. The

genes *glnPQ* are transcriptionally repressed in low nitrogen/ ammonium environments by the small RNA GlnZ and ultimately the nitrogen regulator NtrC (Walling et al., 2022). The symporter GltP is regulated by ArgR as it feeds into the arginine biosynthesis pathway (Ishihama et al., 2016). The peptide transporters Dpp and Opp exhibit expression during late exponential and stationary phase growth, speculated to be due to a shift from growth on free amino acids to larger peptides (Baev et al., 2006). The Dtp family is less well characterised but is likely to be regulated similarly to the *dpp* and *opp* operons with upregulation in the presence of peptides (Bawdon, 2014). Given the biomass used in the transporter screen was grown using LB medium for 1 day it can be assumed that most of these transporters will have been expressed in the rich medium.

The requirement of recombinant expression of an active malodour generating aminoacylase in *E. coli* to enable VFA release prevents a high throughput screening of many different KEIO transport mutants and necessitates a rationalised approach as used earlier. Screening of WT BW25113 expressing Nag3 for VFA release while in the presence of an ionophore such as nigericin or carbonyl cyanide m-chlorophenyl hydrazone (CCCP) would disrupt the proton motive force, inhibiting symport translocation and allow determination as to whether import is mediated by secondary transporters. A similar experiment could also feasibly be performed upon *A. nagyae* and *A. octavius* to aid identification of the *Anaerococcus* transporter. Experimentation with varied VFA precursor concentrations could also yield informative results as to the transporter affinity without requiring identification. Even with the use of ionophores when screening this would still require further rationalised choice in which *E. coli* transport knockout mutants to test. Recent work in *S. aureus* identified a glutamine secondary transporter, AlsT, which was previously annotated as an alanine/glycine Na^+ symporter (Zeden et al., 2020). Sequence similarity searches within *Anaerococcus* does identify putative transporters across the genus, including *A. nagyae* and *A. octavius*, and homologues are also found within *E. coli*. This transporter family should be screened to ascertain any role in VFA release and demonstrates the potential for novel transporters to be identified within previously misannotated protein families.

In summary, screening of Peptoniphilaceae strains revealed VFA liberation is present in several *Anaerococcus* species that contain a close homologue to Oct3 and Nag3, although this is not an absolute predictor of activity as evidenced by *A. tetradius* activity. Nevertheless, all species which did contain a close homologue were shown to be active. *Peptoniphilus* VFA generation appears much rarer than in *Anaerococcus* but screening of an axilla *P. lacydonensis* isolate demonstrated that native axilla flora has clear differences in VFA generation likely caused by adaptation to the environment and exposure to the VFA precursors. All three active VFA liberating aminoacylases were validated as being able to release the proposed VFA products matching *in vitro* characterisation. Expression of Nag3 in

E. coli was shown to gain VFA malodour generation in this strain and enabling its use to identify the *E. coli* transporter responsible for precursor uptake. A small screen suggests that uptake is not mediated by a single transporter but likely involves a POT transporter as one route of uptake. While *A. octavius* contains a putative POT transporter this appears to be unique to this species, suggesting that POT mediated precursor transport is not the main mechanism of precursor uptake in *Anaerococcus.* Further work on other transport mutants in *E. coli* is needed to confidently identify the proteins responsible for VFA precursor translocation.

Summary and future directions

7 Summary and future directions

In this study, the enzymes responsible for VFA malodour generation in *A. octavius* and *A. nagyae* were identified as the orthologous aminoacylases, Oct3 and Nag3. Both displayed N^{α}-acylglutamine activity, and share on average 34 % sequence identity to the previously identified N-AGA from *C.* sp. Ax20 (Natsch et al., 2003). These orthologous anaerococcal proteins were shown to be zinc containing M20D metalloenzymes and were observed *in vitro* to generate VFA malodour by detection of both the L-glutamine product with a colourimetric ninhydrin assay and direct detection of VFAs by chromatography and mass spectrometry. Nag3 was also shown to confer VFA releasing activity to *E. coli* in a gain of function experiment indicating that the identified enzymes are responsible for VFA liberation. Interestingly the orthologous enzymes display a distinct difference in substrate specificity for HMHA-Gln with Oct3 showing much lower activity than either Nag3 or N-AGA suggesting differences in substrate binding sites.

Structural determination of Oct3 and the inactive mutant Nag3E138A by x-ray diffraction showed both assume the characteristic domain and fold architecture of dimeric M20 proteins and that observed for N-AGA (Natsch and Emter, 2020; Rawlings and Barrett, 2013). The *Anaerococcus* structures were observed in a similar open conformation to one another despite difference in oligomerisation within the asymmetric unit of the crystals. Both Nag3E138A and Oct3 show conserved glutamine moiety interacting residues within their binding site, suggesting both form key interactions with glutamine containing substrates which was shown with site directed mutagenesis in Nag3. These binding site residues are also very similar in N-AGA indicating they will have the same interactions with VFA substrates (Natsch and Emter, 2020). All three VFA liberating enzymes were observed to have high specificity for acylglutamine substrates suggesting specific adaptation to utilise the VFA precursors present in their host organism environment.

The homotetrameric structure of Nag3E138A showed notable difference in protomer positioning compared to the closed, bound N-AGA structure with an approximately 5 Å shift for substrate interacting residues residing on loops on the end of another monomer's oligomerisation domain. This difference seems evident of clear inter- and intra-monomer movement upon substrate binding, which has been studied in detail for the M20 enzyme DapE (Nocek et al., 2018). This movement appears to be a common feature of the family due to the flexible linking regions between the catalytic and oligomerisation domain and has been evidenced in several M20 proteins (Chang et al., 2010; Jozic et al., 2002; Lindner et al., 2005; Lundgren et al., 2007). However, it appears in some literature that this feature may have been overlooked (Savić et al., 2009; Simunovic et al., 2011) due to the incorrect

assumption that oligomer state within a structure is representative of its in-solution state (Bitto et al., 2009).

The large-scale protein movement makes it difficult to accurately find the cause of the differing HMHA-Gln specificity between Oct3 and Nag3 without structures in their bound state but the location of His197 close to the potential positioning of the hydroxyl in HMHA could likely form adverse interactions if too close. The generation of catalytically inactive mutants of Nag3 and Oct3 makes further attempts to crystallise both bound to the VFA precursors a logical further experiment to conduct that would yield informative results. The screening of Nag3 and Oct3 with other branched acylglutamine substrates would also be beneficial in understanding the full specificity of each enzyme. Screening for activity of Oct3 and Nag3 with axilla sweat in a similar method to that performed by Natsch *et al.* (2006) with purified N-AGA on axilla sweat would allow access to a variety of unusual branched acylglutamine substrates (Natsch et al., 2006). Additionally, the treatment of axilla sweat with a mixture of Nag3, Oct3 and N-AGA would act as a more complete representation of VFA releasing activity in the axilla and potentially allow the discovery of other minor VFAs.

Nag3E138A forms the similar homotetramer arrangement observed in structures of several other M20 proteins, including N-AGA (Botelho et al., 2011; Girish et al., 2013; Natsch and Emter, 2020). This tetramer state was assumed to be an effect of crystallisation but has been reported in solution upon competitive inhibitor binding in N-AGA, which was suggested may occur during substrate binding (Natsch, 2015). The interface between Nag3E138A oligomerisation domains was not explored in this work but mutagenesis to disrupt these interactions could be performed to assess if they have any role in activity.

The observation of Nag3 and Oct3 affinity values in the millimolar range for VFA precursors, would indicate the need for an efficient transporter to increase intracellular concentrations to suitable levels of significant activity. HMHA-Gln was observed to be present in axilla sweat at an averaged concentration of 113.5 mg/ L (Troccaz et al., 2009), which is equal to approximately 0.4 mM. This is almost 10-fold lower than affinity of Nag3 for this substrate further supporting the requirement for an effective transporter to enable substrate concentration. Screening of *E. coli* strains deficient in expression of the most common peptide or glutamine transporters in *E. coli* did not lead to dramatic loss of VFA production, although the potential role of a POT transporter(s) in uptake was implicated. While this screen failed to identify the major transporter of VFA precursors, the work and methodology detailed provides a clear mechanism to search for this transporter. Bacterial uptake of substrates during VFA biotransformation remains the single most important unknown process in axilla malodour generation with the process in thioalcohol production having already been well

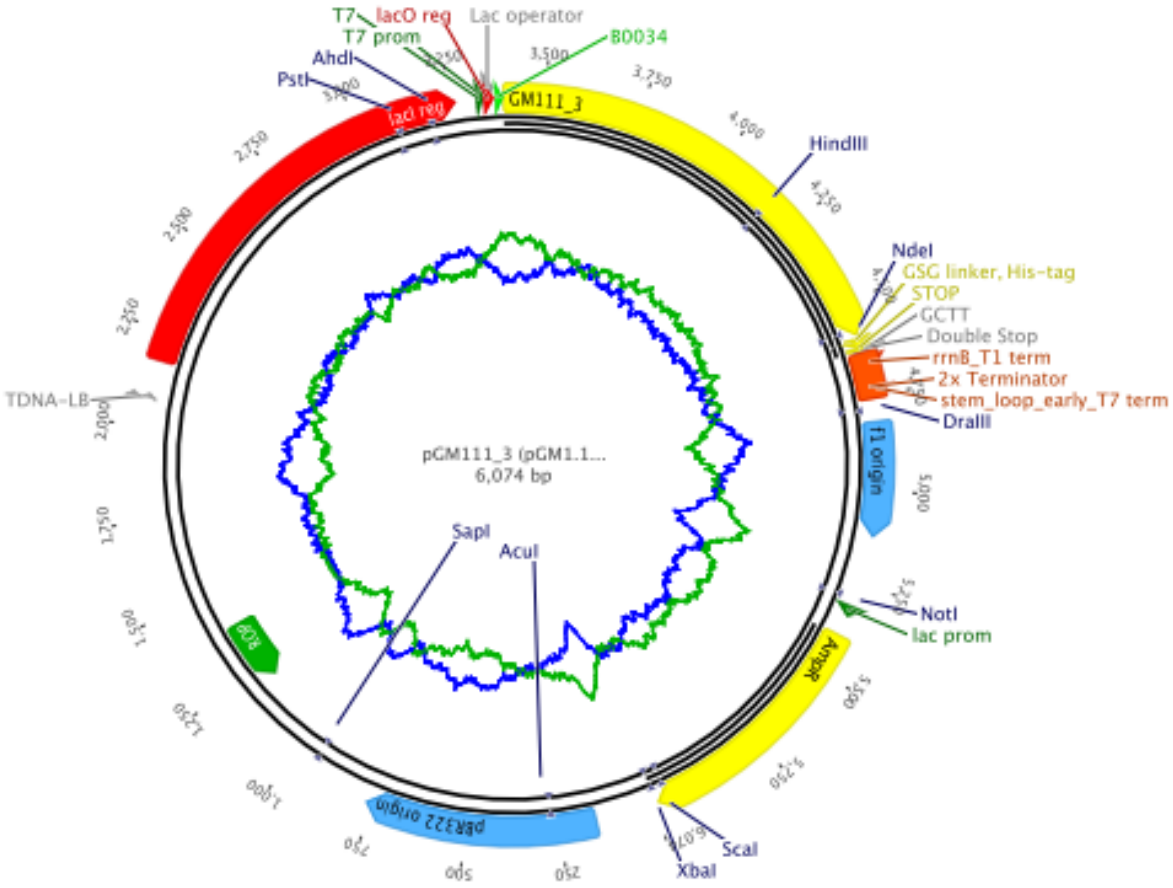
characterised (Minhas et al., 2018; Rudden et al., 2020). Therefore, the mechanism of import remains an exciting area of research.

Screening of *Anaerococcus* species for VFA release identified many that had at least a low level of activity with 3M2H-Gln, with HMHA-Gln activity being rarer. These species, except for *A. tetradius*, all contained a close homologue of Nag3/ Oct3 suggesting those enzymes were responsible for VFA generation in their host species. All *Anaerococcus* species tested were not axillary isolates which could indicate that the observed activity may be even higher if screened with the appropriate strains that have adapted to the axilla environment. Such adaptation could be clearly seen in the *Peptoniphilus* screen with the axilla isolated strain of *P. lacydonensis* showing high VFA release and the human nose isolate having no activity. Such stark difference in activity highlights the importance in screening to be performed on axilla isolate strains to enable a more accurate picture of VFA generation by species. Future experiments performing this would be ideal as well as the screening of axilla bacteria isolated from both high and low axilla BO individuals. This would allow the study of any association with specific species and BO intensity as well as inform on any strain heterogeneity in VFA production which may be present.

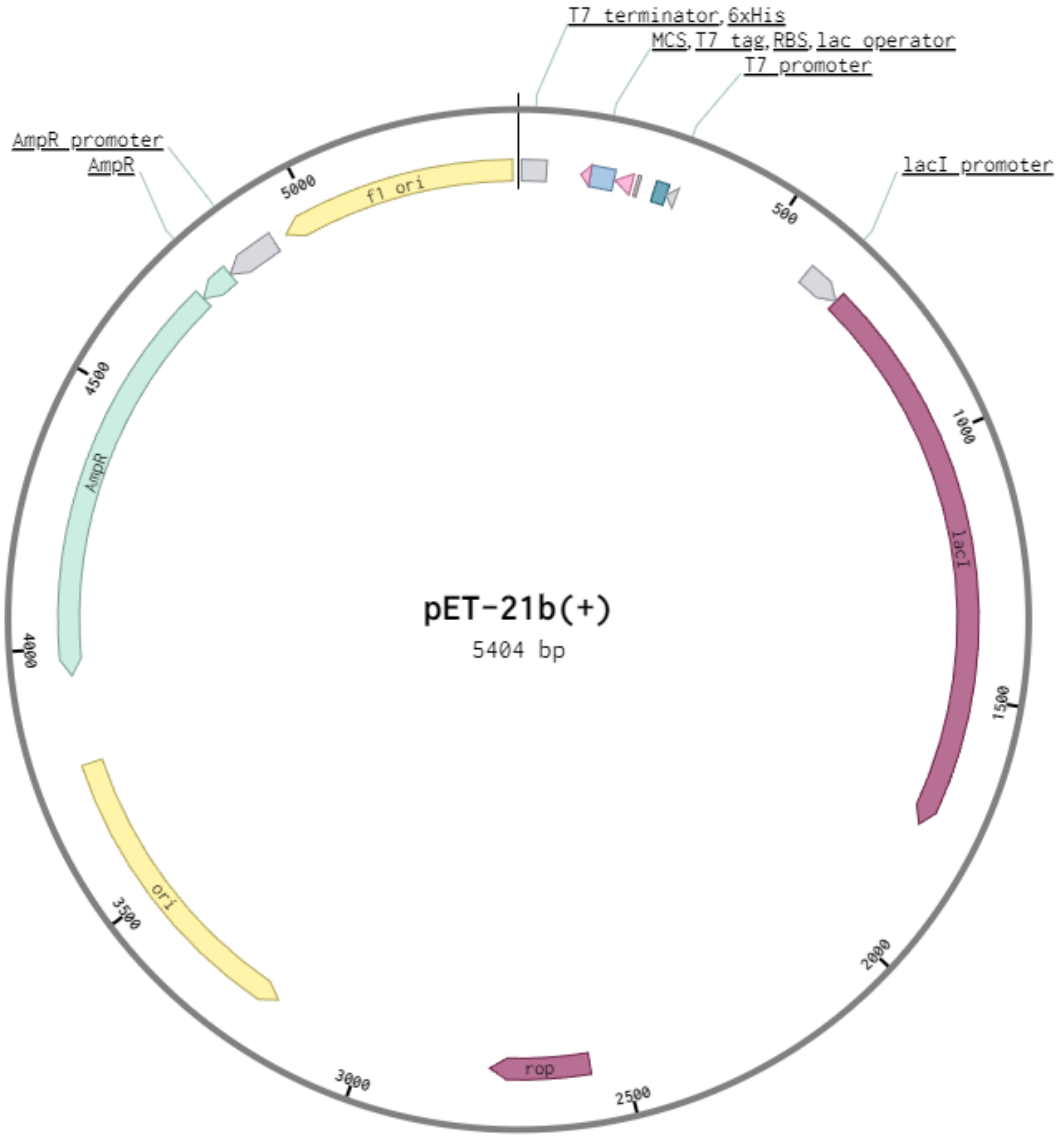
This work clearly shows a key role for *Anaerococcus* in VFA based malodour generation in the axilla which raise questions as to the importance of N-AGA and *Corynebacterium* in underarm BO. While N-AGA has been shown both here and by others (Natsch et al., 2003, 2006) to be a highly active and efficient VFA liberating aminoacylase, the host species *C.* sp. Ax20 appears to be unique within the genus for its significant activity. *Corynebacterium* have consistently been linked with axilla body odour (Leyden et al., 1981; Taylor et al., 2003; Troccaz et al., 2015), and yet their contribution to both thioalcohol and VFA malodour has been questioned (Bawdon et al., 2015; James et al., 2013b). The exact role of corynebacteria therefore remains unknown but could be linked with aiding microbial growth or malodour activity in BO generating species in *Staphylococcus* and *Anaerococcus*. The study of potential interactions between axilla bacteria and their effects on axilla body odour could provide an important insight into malodour generation in the underarm.

8 Appendices

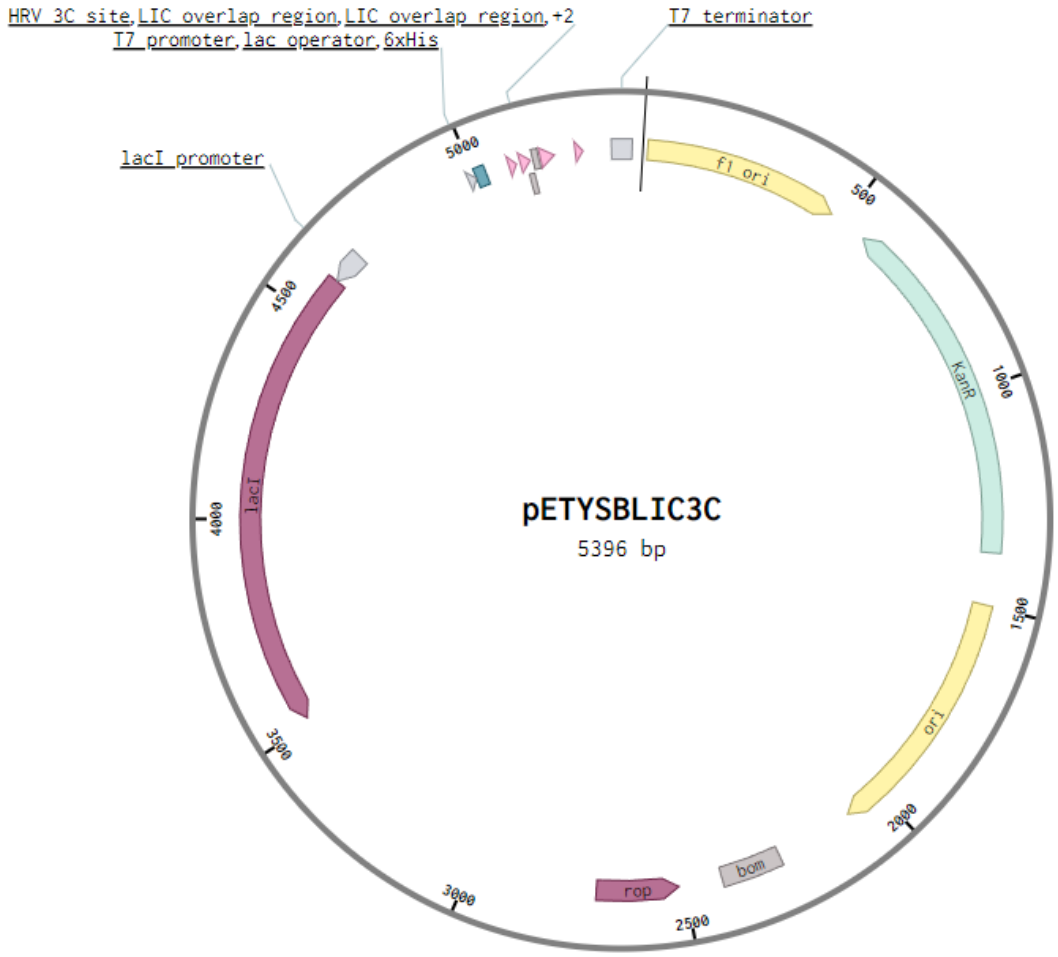
8.1 Appendix A – plasmid maps



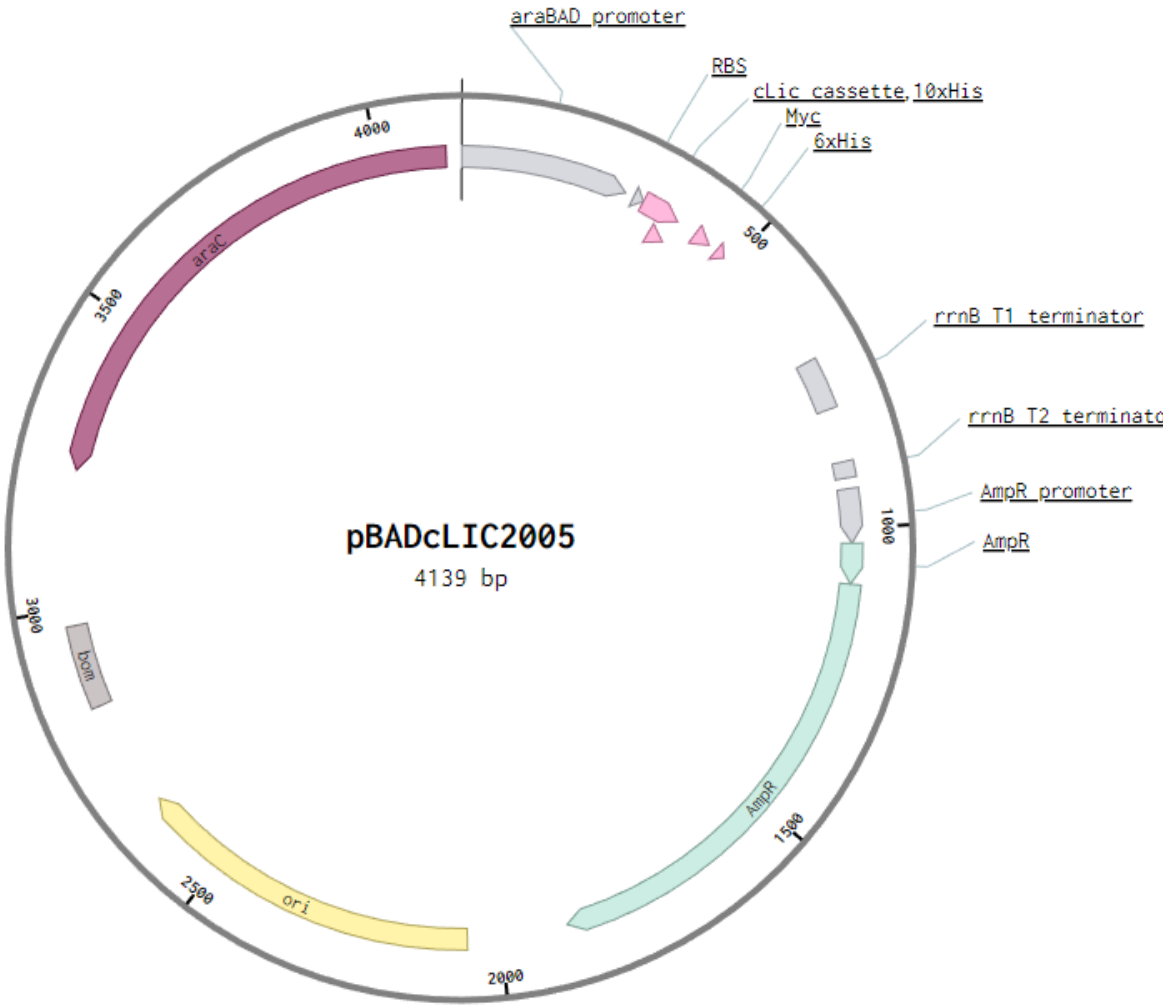
Vector map of pGM vector containing Oct3.
Ampicillin resistance conferred by AmpR gene. Upstream of Oct3 (labelled GM111-3) is the T7 promoter and lacO regulator site, enabling IPTG inducible protein expression, while downstream is a short GSG linker to a hexa-histidine tag. The origin of replication is marked as pBR322 origin and ROP denotes repressor of primer gene sequence.



pET-21b(+) empty vector map.
Ampicillin resistance conferred by AmpR gene. pBR332 origin of replication and repressor of primer gene marked as ori and rop respectively.

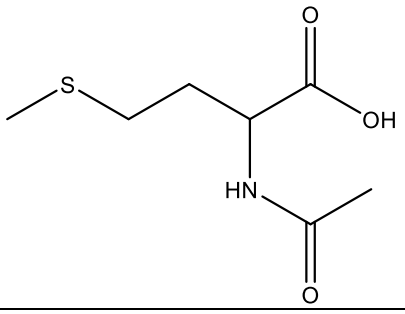
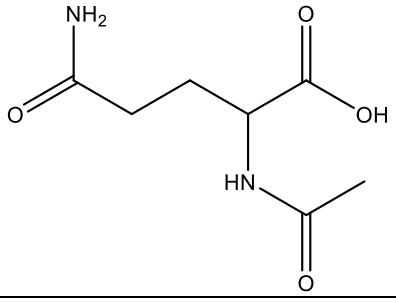
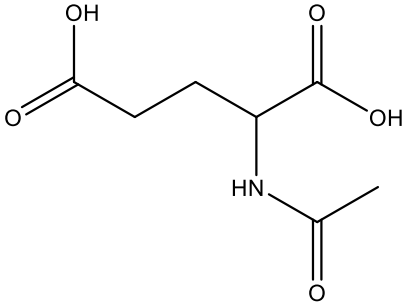
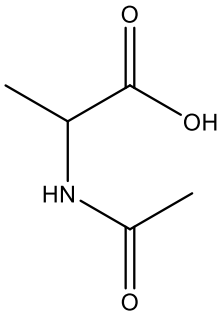
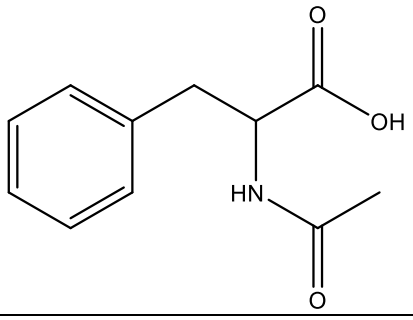


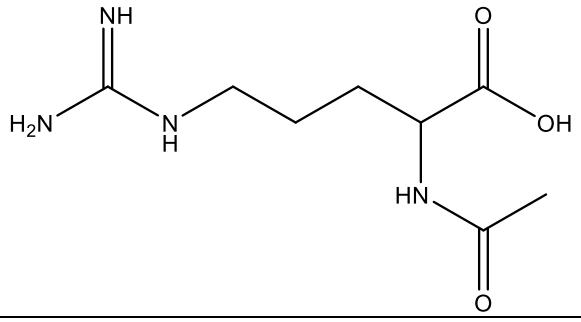
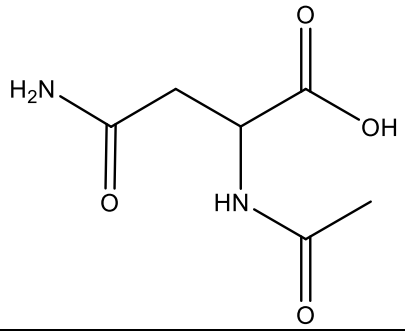
pETYSBLIC3C empty vector map.
Kanamycin resistance conferred by KanR gene. pBR332 origin of replication and repressor of primer gene marked as ori and rop respectively.



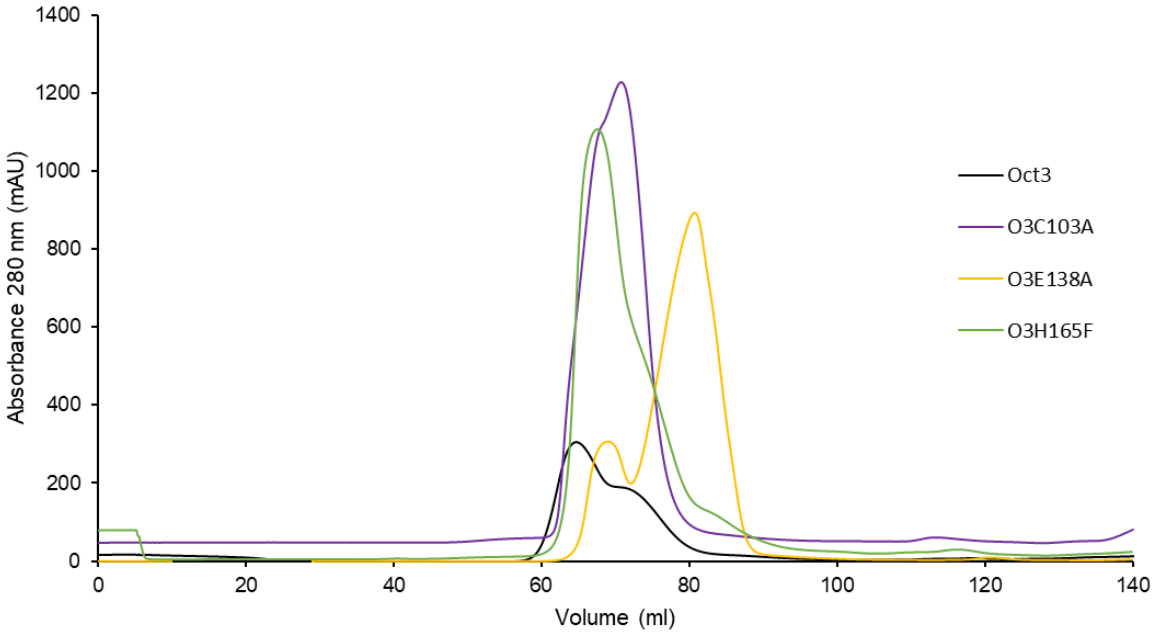
pBADcLIC2005 empty vector map.
Ampicillin resistance conferred by AmpR gene. pBR332 origin of replication marked as ori. Protein expression controlled by araBAD promoter.

8.2 Appendix B – acetylated amino acid structures

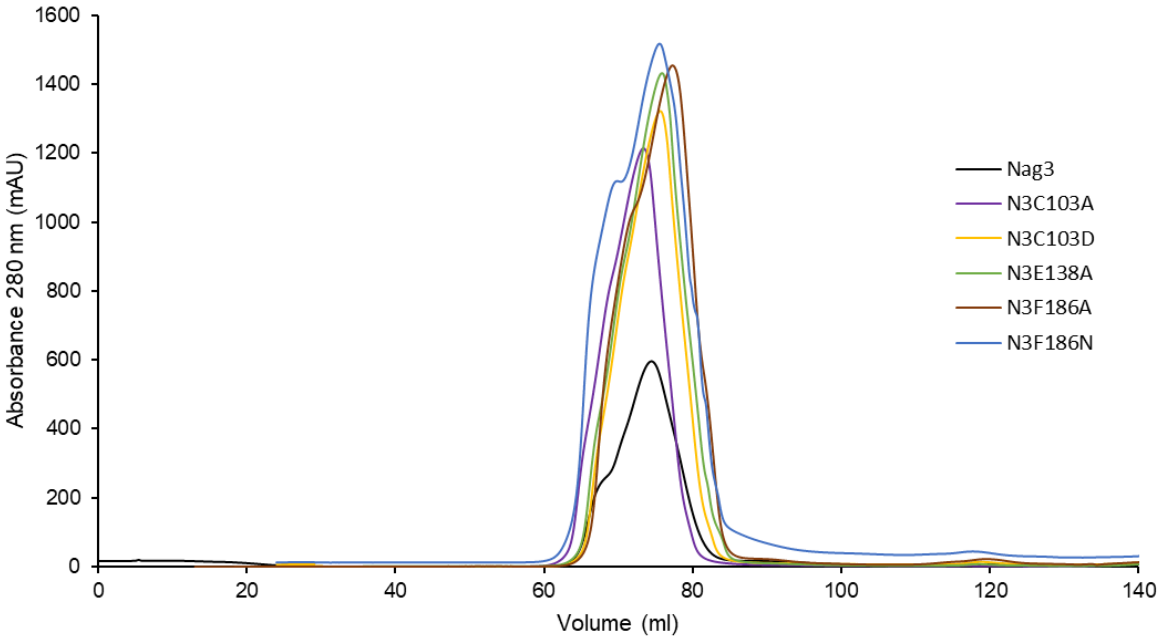
Acetylated amino acid chemical structures	
Name	Structure
N-acetyl-methionine	
N^{α}-acetyl-glutamine	
N-acetyl-glutamate	
N-acetyl-alanine	
N-acetyl-phenylalanine	

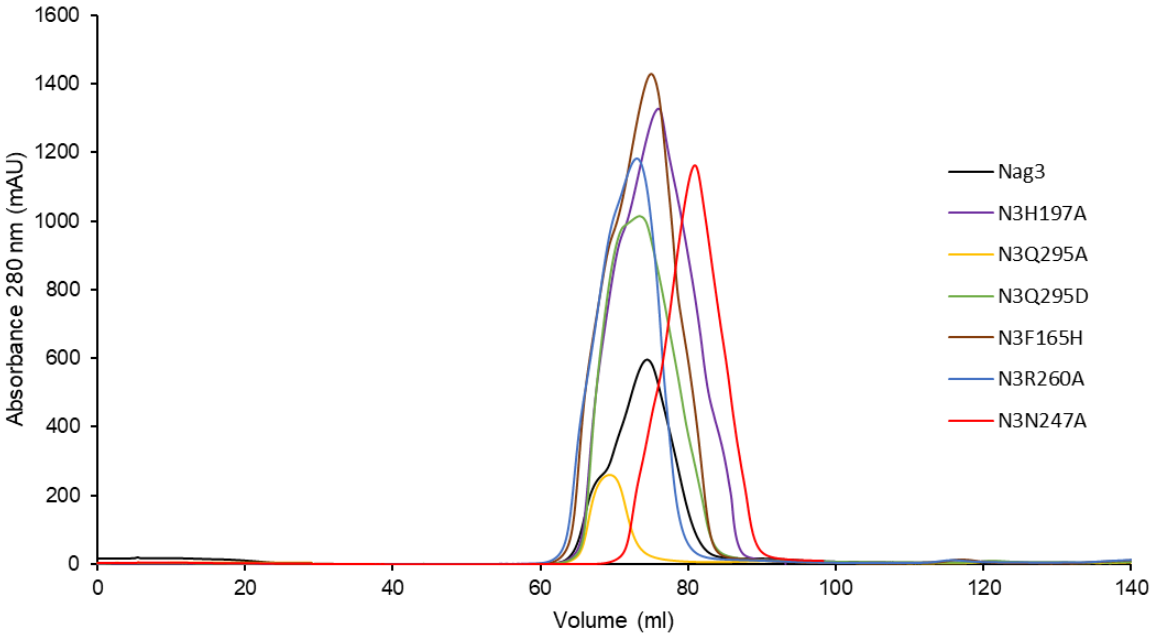
N-acetyl-arginine	
N-acetyl-asparagine	

8.3 Appendix C – Mutant size exclusion chromatography traces



Oct3 mutant size exclusion chromatography absorbance traces





Nag3 mutants size exclusion chromatography absorbance traces

9 References

Adelman, S., Taylor, C.R., and Heglund, N.C. (1975). Sweating on paws and palms: what is its function? American Journal of Physiology *229*, 1400–1402.

Agarwal, R., Burley, S.K., and Swaminathan, S. (2007). Structural Analysis of a Ternary Complex of Allantoate Amidohydrolase from *Escherichia coli* Reveals its Mechanics. Journal of Molecular Biology *368*, 450–463.

Akiba, S., Arai, N., Kusuoku, H., Takagi, Y., Hagura, T., Takeuchi, K., Fuji, A., and Fujimura, T. (2011). The N-terminal amino acid of apolipoprotein D is putatively covalently bound to 3-hydroxy-3-methyl hexanoic acid, a key odour compound in axillary sweat. International Journal of Cosmetic Science *33*, 283–286.

Amoore, J.E., Pelosi, P., and Forrester, L.J. (1977). Specific anosmias to 5a-androst-16-en-3-one and co-pentadecalactone: The urinous and musky primary odors. Chemical Senses *2*, 267–281.

Andreo-Vidal, A., Sanchez-Amat, A., and Campillo-Brocal, J.C. (2018). The pseudoalteromonas luteoviolacea L-amino acid oxidase with antimicrobial activity is a flavoenzyme. Marine Drugs *16*.

Anjem, A., and Imlay, J.A. (2012). Mononuclear iron enzymes are primary targets of hydrogen peroxide stress. Journal of Biological Chemistry *287*, 15544–15556.

Asahina, M., Poudel, A., and Hirano, S. (2015). Sweating on the palm and sole: physiological and clinical relevance. Clinical Autonomic Research *25*, 153–159.

Auld, D.S. (1995). Removal and replacement of metal ions in metallopeptidases. Methods in Enzymology *248*, 228–242.

Austin, C., and Ellis, J. (2003). Microbial pathways leading to steroidal malodour in the axilla. Journal of Steroid Biochemistry and Molecular Biology *87*, 105–110.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. Molecular Systems Biology *2*, 2006.0008.

Baev, M. V., Baev, D., Jansco Radek, A., and Campbell, J.W. (2006). Growth of *Escherichia coli* MG1655 on LB medium: Monitoring utilization of amino acids, peptides, and nucleotides with transcriptional microarrays. Applied Microbiology and Biotechnology *71*, 317–322.

Baker, L.B., and Wolfe, A.S. (2020). Physiological mechanisms determining eccrine sweat composition. European Journal of Applied Physiology *120*, 719–752.

Bawdon, D. (2014). Identification of bacterial transporters for hydroxyalkylcysteinylglycines. University of York.

Bawdon, D., Cox, D.S., Ashford, D., James, A.G., and Thomas, G.H. (2015). Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. FEMS Microbiology Letters *362*, fnv111.

Beer, G.M., Baumüller, S., Zech, N., Wyss, P., Strasser, D., Varga, Z., Seifert, B., Hafner, J., and Mihic-Probst, D. (2006). Immunohistochemical differentiation and localization analysis of sweat glands in the adult human axilla. Plastic and Reconstructive Surgery *117*, 2043–2049.

Berger, B.J., English, S., Chan, G., and Knodel, M.H. (2003). Methionine regeneration and aminotransferases in *Bacillus subtilis*, *Bacillus cereus*, and *Bacillus anthracis*. Journal of Bacteriology

185, 2418–2431.

Berglund, L.G., and McNall, P.E. (1973). Human sweat film area and composition during prolonged sweating. Journal of Applied Physiology *35*, 714–718.

Bibel, D.J., and Lovell, D.J. (1976). Skin flora maps: a tool in the study of cutaneous ecology. Journal of Investigative Dermatology *67*, 265–269.

Bitto, E., Bingman, C.A., Bittova, L., Houston, N.L., Boston, R.S., Fox, B.G., and Phillips, G.N. (2009). X-ray structure of ILL2, an auxin-conjugate amidohydrolase from *Arabidopsis thaliana*. Proteins: Structure, Function, and Bioinformatics *74*, 61–71.

De Boever, E.H., De Uzeda, M., and Loesche, W.J. (1994). Relationship between volatile sulfur compounds, BANA-hydrolyzing bacteria and gingival health in patients with and without complaints of oral malodor. The Journal of Clinical Dentistry *4*, 114–119.

Botelho, T.O., Guevara, T., Marrero, A., Arêde, P., Fluxà, V.S., Reymond, J.-L., Oliveira, D.C., and Gomis-Rüth, F.X. (2011). Structural and Functional Analyses Reveal That *Staphylococcus aureus* Antibiotic Resistance Factor HmrA Is a Zinc-dependent Endopeptidase. Journal of Biological Chemistry *286*, 25697–25709.

Bottom, C.B., Hanna, S.S., and Siehr, D.J. (1978). Mechanism of the ninhydrin reaction. Biochemical Education *6*, 4–5.

Bovell, D. (2015). The human eccrine sweat gland: Structure, function and disorders. Journal of Local and Global Health Science *2015*, 5.

Bovell, D.L., Corbett, A.D., Holmes, S., MacDonald, A., and Harker, M. (2007). The absence of apoeccrine glands in the human axilla has disease pathogenetic implications, including axillary hyperhidrosis. British Journal of Dermatology *156*, 1278–1286.

Brazier, J., Chmelar, D., Dubreuil, L., Feierl, G., Hedberg, M., Kalenic, S., Könönen, E., Lundgren, B., Malamou-Ladas, H., Nagy, E., et al. (2008). European surveillance study on antimicrobial susceptibility of Gram-positive anaerobic cocci. International Journal of Antimicrobial Agents *31*, 316–320.

Brooksbank, B.W.L., Brown, R., and Gustafsson, J.A. (1974). The detection of 5 α-androst-16-en-3 α-ol in human male axillary sweat. Experientia *30*, 864–865.

Brown, C.W., Sridhara, V., Boutz, D.R., Person, M.D., Marcotte, E.M., Barrick, J.E., and Wilke, C.O. (2017). Large-scale analysis of post-translational modifications in *E.coli* under glucose-limiting conditions. BMC Genomics *18*, 1–21.

Bullard, R.W., Dill, D.B., and Yousef, M.K. (1970). Responses of the burro to desert heat stress. Journal of Applied Physiology *29*, 159–167.

Campbell, G.A., Burgdorf, W.H.C., and Everett, M.A. (1982). The immunohistochemical localization of lysozyme in human axillary apocrine glands. Journal of Investigative Dermatology *79*, 351–353.

Cao, C.-H., Cheng, F., Xue, Y.-P., and Zheng, Y.-G. (2020). Efficient synthesis of L-phosphinothricin using a novel aminoacylase mined from *Stenotrophomonas maltophilia*. Enzyme and Microbial Technology *135*, 109493.

Cao, S., Huseby, D.L., Brandis, G., and Hughes, D. (2017). Alternative evolutionary pathways for drug-resistant small colony variant mutants in *Staphylococcus aureus*. MBio *8*, e00358-17.

Cernoch, J.M., and Porter, R.H. (1985). Recognition of Maternal Axillary Odors by Infants. Child Development *56*, 1593.

Chalmers, T.M., and Keele, C.A. (1952). The nervous and chemcial control of sweating. British Journal of Dermatology *64*, 43–54.

Chang, C.-Y., Hsieh, Y.-C., Wang, T.-Y., Chen, Y.-C., Wang, Y.-K., Chiang, T.-W., Chen, Y.-J., Chang, C.-H., Chen, C.-J., and Wu, T.-K. (2010). Crystal Structure and Mutational Analysis of Aminoacylhistidine Dipeptidase from *Vibrio alginolyticus* Reveal a New Architecture of M20 Metallopeptidases. Journal of Biological Chemistry *285*, 39500–39510.

Chaplin, G., Jablonski, N.G., Sussman, R.W., and Kelley, E.A. (2014). The role of piloerection in primate thermoregulation. Folia Primatologica *85*, 1–17.

Chiang, P.K., Gordon, R.K., Tal, J., Zeng, G.C., Doctor, B.P., Pardhasaradhi, K., and McCann, P.P. (1996). S-Adenosylmetliionine and methylation. The FASEB Journal *10*, 471–480.

Classen, C., Howes, D., and Synnott, A. (1994). Aroma: The Cultural History of Smell.

Claus, R., and Alsing, W. (1976). Occurrence of 5α-androst-16-en-3-one, a boar pheromone, in man and its relationship to testosterone. Journal of Endocrinology *68*, 483–484.

Collins, K.J. (1989). Sweat Glands: Eccrine and Apocrine. (Springer, Berlin, Heidelberg), pp. 193–212.

Colombo, S., D'auria, S., Fusi, P., Zecca, L., Raia, C.A., and Tortora, P. (1992). Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*. European Journal of Biochemistry *206*, 349–357.

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. Science *326*, 1694–1697.

Craig Roberts, S., Havlíček, J., and Schaal, B. (2020). Human olfactory communication: Current challenges and future prospects. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Cramer, M.N., and Jay, O. (2016). Biophysical aspects of human thermoregulation during heat stress. Autonomic Neuroscience: Basic and Clinical *196*, 3–13.

Dahmani, M., Diatta, G., Labas, N., Diop, A., Bassene, H., Raoult, D., Granjon, L., Fenollar, F., and Mediannikov, O. (2018). Noncontiguous finished genome sequence and description of *Bartonella mastomydis* sp. nov. New Microbes and New Infections *25*, 60–70.

Daud, H., Browne, S., Al-Majmaie, R., Murphy, W., and Al-Rubeai, M. (2016). Metabolic profiling of hematopoietic stem and progenitor cells during proliferation and differentiation into red blood cells. New Biotechnology *33*, 179–186.

Delanghe, G., Ghyselen, J., Bollen, C., van Steenberghe, D., Vandekerckhove, B.N., and Feenstra, L. (1999). An inventory of patients' response to treatment at a multidisciplinary breath odor clinic. Quintessence International (Berlin, Germany : 1985) *30*, 307–310.

Derayea, S.M., and Samir, E. (2020). A review on the use of fluorescamine as versatile and convenient analytical probe. Microchemical Journal *156*, 104835.

Devos, Patte F, Rouault J, Laffort P, and Van Gemert L (1990). Human Olfactory Thresholds. 118.

Downing, D.T., Stewart, M.E., and Strauss, J.S. (1989). Changes in sebum secretion and the sebaceous gland. Clinics in Geriatric Medicine *5*, 109–114.

Egert, M., Schmidt, I., Höhne, H.M., Lachnit, T., Schmitz, R.A., and Breves, R. (2011). rRNA-based profiling of bacteria in the axilla of healthymales suggests right-left asymmetry in bacterial activity. FEMS Microbiology Ecology *77*, 146–153.

Elbourne, L.D.H., Tetu, S.G., Hassan, K.A., and Paulsen, I.T. (2017). TransportDB 2.0: A database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Research *45*, D320–D324.

Emter, R., and Natsch, A. (2008). The sequential action of a dipeptidase and a β-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. Journal of Biological Chemistry *283*, 20645–20652.

Evershed, R.P., Robertson, D.H.L., Beynon, R.J., and Green, B.N. (1993). Application of electrospray ionization mass spectrometry with maximum-entropy analysis to allelic 'fingerprinting' of major urinary proteins. Rapid Communications in Mass Spectrometry *7*, 882–886.

Ezaki, T., Kawamura, Y., Li, N., Li, Z.Y., Zhao, L., and Shu, S.E. (2001). Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov for members of the genus *Peptostreptococcus*. International Journal of Systematic and Evolutionary Microbiology *51*, 1521–1528.

Farkaš, R. (2015). Apocrine secretion: New insights into an old phenomenon. Biochimica et Biophysica Acta - General Subjects *1850*, 1740–1750.

Ferdenzi, C., Ortegón, S.R., Delplanque, S., Baldovini, N., and Bensafi, M. (2020). Interdisciplinary challenges for elucidating human olfactory attractiveness. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Fogg, M.J., and Wilkinson, A.J. (2008). Higher-throughput approaches to crystallization and crystal structure determination. Biochemical Society Transactions *36*, 771–775.

Folk, G.E., and Semken, A. (1991). The evolution of sweat glands. International Journal of Biometeorology *35*, 180–186.

Fredrich, E., Barzantny, H., Brune, I., and Tauch, A. (2013). Daily battle against body odor: Towards the activity of the axillary microbiota. Trends in Microbiology *21*, 305–312.

Friedman, M. (2004). Applications of the Ninhydrin Reaction for Analysis of Amino Acids, Peptides, and Proteins to Agricultural and Biomedical Sciences. Journal of Agricultural and Food Chemistry *52*, 385–406.

Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R.C., Giglione, C., and Meinnel, T. (2006). The proteomics of N-terminal methionine cleavage. Molecular and Cellular Proteomics *5*, 2336–2349.

Fujii, T., Shinozaki, J., Kajiura, T., Iwasaki, K., and Fudou, R. (2014). A newly discovered *Anaerococcus* strain responsible for axillary odor and a new axillary odor inhibitor, pentagalloyl glucose. FEMS Microbiology Ecology *89*, 198–207.

Fulyani, F., Schuurman-Wolters, G.K., Žagar, A.V., Guskov, A., Slotboom, D.J., and Poolman, B. (2013). Functional diversity of tandem substrate-binding domains in ABC transporters from pathogenic bacteria. Structure *21*, 1879–1888.

Fulyani, F., Schuurman-Wolters, G.K., Slotboom, D.J., and Poolman, B. (2016). Relative rates of amino acid import via the ABC transporter GlnPQ determine the growth performance of *Lactococcus lactis*. Journal of Bacteriology *198*, 477–485.

Gade, W., and Brown, J.L. (1981). Purification, characterization and possible function of α-N-acylamino acid hydrolase from bovine liver. BBA - Enzymology *662*, 86–93.

Gallo, R.L. (2017). Human Skin Is the Largest Epithelial Surface for Interaction with Microbes. Journal of Investigative Dermatology *137*, 1213–1214.

Geertsma, E.R., Groeneveld, M., Slotboom, D.J., and Poolman, B. (2008). Quality control of overexpressed membrane proteins. Proceedings of the National Academy of Sciences of the United States of America *105*, 5722–5727.

Gemmer, C.M., DeAngelis, Y.M., Theelen, B., Boekhout, T., and Dawson, T.L. (2002). Fast, noninvasive method for molecular detection and differentiation of *Malassezia* yeast species on human skin and application of the method to dandruff microbiology. Journal of Clinical Microbiology *40*, 3350–3357.

Girish, T.S., and Gopal, B. (2010). Crystal Structure of *Staphylococcus aureus* Metallopeptidase (Sapep) Reveals Large Domain Motions between the Manganese-bound and Apo-states. Journal of Biological Chemistry *285*, 29406–29415.

Girish, T.S., Vivek, B., Colaco, M., Misquith, S., and Gopal, B. (2013). Structure of an amidohydrolase, SACOL0085, from methicillin-resistant *Staphylococcus aureus* COL. Acta Crystallographica Section F: Structural Biology and Crystallization Communications *69*, 103–108.

Gower, D.B. (1972). 16-Unsaturated C19 steroids a review of their chemistry, biochemistry and possible physiological role. Journal of Steroid Biochemistry *3*, 45–103.

Gower, D.B., Nixon, A., and Mallet, A.I. (1988). The significance of odorous steroids in axillary odour. In Perfumery, (Springer Netherlands), pp. 47–76.

Gower, D.B., Holland, K.T., Mallet, A.I., Rennie, P.J., and Watkins, W.J. (1994). Comparison of 16-androstene steroid concentrations in sterile apocrine sweat and axillary secretions: interconversions of 16-androstenes by the axillary microflora--a mechanism for axillary odour production in man? The Journal of Steroid Biochemistry and Molecular Biology *48*, 409–418.

Greene, R.S., Downing, D.T., Pochi, P.E., and Strauss, J.S. (1970). Anatomical variation in the amount and composition of human skin surface lipid. The Journal of Investigative Dermatology *54*, 240–247.

Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., Bouffard, G.G., Blakesley, R.W., Murray, P.R., Green, E.D., et al. (2009). Topographical and Temporal Diversity of the Human Skin Microbiome. Science *324*, 1190–1192.

de Groot, J.H.B., Semin, G.R., and Smeets, M.A.M. (2014). Chemical communication of fear: A case of male-female asymmetry. Journal of Experimental Psychology: General *143*, 1515–1525.

Habig, W., Pabst, M., and Jakoby, W. (1974). Glutathione S-transferase. The Journal of Biological Chemistry *249*, 7130–7139.

Harker, M., and Harding, C.R. (2013). Amino acid composition, including key derivatives of eccrine sweat: Potential biomarkers of certain atopic skin conditions. International Journal of Cosmetic Science *35*, 163–168.

Harker, M., Coulson, H., Fairweather, I., Taylor, D., and Daykin, C.A. (2006). Study of metabolite composition of eccrine sweat from healthy male and female human subjects by 1H NMR spectroscopy. Metabolomics *2*, 105–112.

Hasegawa, Y., Yabuki, M., and Matsukane, M. (2004). Identification of New Odoriferous Compounds in Human Axillary Sweat. Chemistry and Biodiversity *1*, 2042–2050.

Hashimoto, K. (1965). The ultrastructure of the skin of human embryos. I. The intraepidermal eccrine sweat duct. The Journal of Investigative Dermatology *45*, 139–151.

Havlíček, J., Winternitz, J., and Craig Roberts, S. (2020). Major histocompatibility complex-associated odour preferences and human mate choice: Near and far horizons. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Herz, R.S., and Inzlicht, M. (2002). Sex differences in response to physical and social factors involved in human mate selection. The importance of smell for women. Evolution and Human Behavior *23*, 359–364.

Hochachka, P.W. (1986). Defense strategies against hypoxia and hypothermia. Science *231*, 234–241.

Hurley, H.J., and Shelley, W.B. (1960). The Human Apocrine Sweat Gland in Health and Disease. Archives of Dermatology *82*, 848.

Ishihama, A., Shimada, T., and Yamazaki, Y. (2016). Transcription profile of *Escherichia coli*: Genomic SELEX search for regulatory targets of transcription factors. Nucleic Acids Research *44*, 2058–2074.

Ishikawa, K., Ishida, H., Matsui, I., Kawarabayasi, Y., and Kikuchi, H. (2001). Novel bifunctional hyperthermostable carboxypeptidase/aminoacylase from *Pyrococcus horikoshii* OT3. Applied and Environmental Microbiology *67*, 673–679.

Ito, K., Hikida, A., Kawai, S., Lan, V.T.T., Motoyama, T., Kitagawa, S., Yoshikawa, Y., Kato, R., and Kawarasaki, Y. (2013). Analysing the substrate multispecificity of a proton-coupled oligopeptide transporter using a dipeptide library. Nature Communications *4*, 1–10.

Iyamu, E.W., Asakura, T., and Woods, G.M. (2008). A colorimetric microplate assay method for high-throughput analysis of arginase activity in vitro. Analytical Biochemistry *383*, 332–334.

Jackman, P.J.H., and Noble, W.C. (1983). Normal axillary skin in various populations. Clinical and Experimental Dermatology *8*, 259–268.

Jain, R., Hao, B., Liu, R.P., and Chan, M.K. (2005). Structures of *E. coli* peptide deformylase bound to formate: Insight into the preference for Fe2+ over Zn2+ as the active site metal. Journal of the American Chemical Society *127*, 4558–4559.

Jamdar, S.N. (2009). A novel aminopeptidase from *Burkholderia cepacia* specific for acidic amino acids. FEMS Microbiology Letters *295*, 230–237.

Jamdar, S.N., Are, V.N., Navamani, M., Kumar, S., Nagar, V., and Makde, R.D. (2015). The members of M20D peptidase subfamily from *Burkholderia cepacia*, *Deinococcus radiodurans* and *Staphylococcus aureus* (HmrA) are carboxydipeptidases, primarily specific for Met-X dipeptides. Archives of Biochemistry and Biophysics *587*, 18–30.

James, A.G. (2020). The Axillary Microbiome and its Relationship with Underarm Odor. In Skin Microbiome Handbook: From Basic Research to Product Development, (John Wiley & Sons, Ltd), pp. 79–129.

James, A.G., Hyliands, D., and Johnston, H. (2004a). Generation of volatile fatty acids by axillary bacteria. In International Journal of Cosmetic Science, (Wiley/Blackwell (10.1111)), pp. 149–156.

James, A.G., Casey, J., Hyliands, D., and Mycock, G. (2004b). Fatty acid metabolism by cutaneous bacteria and its role in axillary malodour. World Journal of Microbiology and Biotechnology *20*, 787–793.

James, A.G., Cox, D., and Worrall, K. (2013a). Microbiological and biochemical origins of human foot malodour. Flavour and Fragrance Journal *28*, 231–237.

James, A.G., Cox, D., and Worrall, K. (2013b). Microbiological and biochemical origins of human axillary malodour. Flavour and Fragrance Journal *28*, 231–237.

Janzen, D.H. (1977). Why Fruits Rot, Seeds Mold, and Meat Spoils. The American Naturalist *111*, 691–713.

Jörnvall, H. (1975). Acetylation of protein N-terminal amino groups structural observations on α-amino acetylated proteins. Journal of Theoretical Biology *55*, 1–12.

Jozic, D., Bourenkow, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., Huber, R., Bode, W., and Maskos, K. (2002). Crystal Structure of the Dinuclear Zinc Aminopeptidase PepV from *Lactobacillus delbrueckii* Unravels Its Preference for Dipeptides. Structure *10*, 1097–1106.

Ju, L.C., Cheng, Z., Fast, W., Bonomo, R.A., and Crowder, M.W. (2018). The Continuing Challenge of Metallo-β-Lactamase Inhibition: Mechanism Matters. Trends in Pharmacological Sciences *39*, 635–647.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

Kaushik, P., Curell, R.V.B., Henry, M., Barron, N., and Meleady, P. (2020). LC-MS/MS-based quantitative proteomic and phosphoproteomic analysis of CHO-K1 cells adapted to growth in glutamine-free media. Biotechnology Letters *42*, 2523–2536.

Kim, M.J., Tagele, S.B., Jo, H.W., Kim, M.C., Jung, Y.G., Park, Y.J., So, J.H., Kim, H.J., Kim, H.J., Lee, D.G., et al. (2021). Effect of a bioconverted product of Lotus corniculatus seed on the axillary microbiome and body odor. Scientific Reports *11*, 1–14.

Kloos, W.E., and Musselwhite, M.S. (1975). Distribution and Persistence of *Staphylococcus* and *Micrococcus* Species and Other Aerobic Bacteria on Human Skin. Applied Microbiology *30*, 381–395.

Koreishi, M., Asayama, F., Imanaka, H., Imamura, K., Kadota, M., Tsuno, T., and Nakanishi, K. (2005). Purification and characterization of a novel aminoacylase from *Streptomyces mobaraensis*. Bioscience, Biotechnology and Biochemistry *69*, 1914–1922.

Koreishi, M., Nakatani, Y., Ooi, M., Imanaka, H., Imamura, K., and Nakanishi, K. (2009). Purification, Characterization, Molecular Cloning, and Expression of a New Aminoacylase from *Streptomyces mobaraensis* That Can Hydrolyze N -(Middle/Long)-chain-fatty-acyl- L -amino Acids as Well as N -Short-chain-acyl- L -amino acids. Bioscience, Biotechnology, and Biochemistry *73*, 1940–1947.

Kurz, A. (2008). Physiology of Thermoregulation. Best Practice and Research: Clinical Anaesthesiology *22*, 627–644.

Kwak, J., Willse, A., Preti, G., Yamazaki, K., and Beauchamp, G.K. (2010). In search of the chemical basis for MHC odourtypes. In Proceedings of the Royal Society B: Biological Sciences, (The Royal Society), pp. 2417–2425.

Lamothe, P.J., and McCormick, P.G. (1973). Role of Hydrindantin in the Determination of Amino Acids Using Ninhydrin. Analytical Chemistry *45*, 1906–1911.

Larsen, K.S., and Auld, D.S. (1989). Carboxypeptidase A: Mechanism of Zinc Inhibition. Biochemistry *28*, 9620–9625.

Larsen, K.S., and Auld, D.S. (1991). Characterization of an Inhibitory Metal Binding Site in Carboxypeptidase A. Biochemistry *30*, 2613–2618.

Laska, M., Bautista, R.M.R., Höfelmann, D., Sterlemann, V., and Salazar, L.T.H. (2007). Olfactory sensitivity for putrefaction-associated thiols and indols in three species of non-human primate. Journal of Experimental Biology *210*, 4169–4178.

Leeming, J.P., Holland, K.T., and Cunliffe, W.J. (1984). The microbial ecology of pilosebaceous units isolated from human skin. Journal of General Microbiology *130*, 803–807.

Lepock, J.R. (2003). Cellular effects of hyperthermia: Relevance to the minimum dose for thermal damage. In International Journal of Hyperthermia, pp. 252–266.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research *44*, W242–W245.

Levy, D.R., and Yizhar, O. (2018). Stress and sociability. Nature Neuroscience *21*, 304–306.

Leyden, J.J., McGinley, K.J., Holzle, E., Labows, J.N., and Kligman, A.M. (1981). The microbiology of the human axilla and its relationship to axillary odor. Journal of Investigative Dermatology *77*, 413–416.

Leyden, J.J., McGinley, K.J., Nordstrom, K.M., and Webster, G.F. (1987). Skin microflora. Journal of Investigative Dermatology *88*, 65s-72s.

Lin, L.L., Chen, M.H., Chien, H.C.R., Kan, S.C., Chen, C.C., Hu, H.Y., and Hsu, W.H. (2007). Characterization of a bifunctional aminoacylase/carboxypeptidase from radioresistant bacterium *Deinococcus radiodurans* R1. Journal of Biotechnology *128*, 322–334.

Lindner, H.A., Lunin, V. V., Alary, A., Hecker, R., Cygler, M., and Ménard, R. (2003). Essential Roles of Zinc Ligation and Enzyme Dimerization for Catalysis in the Aminoacylase-1/M20 Family. Journal of Biological Chemistry *278*, 44496–44504.

Lindner, H.A., Alary, A., Boju, L.I., Sulea, T., and Ménard, R. (2005). Roles of dimerization domain residues in binding and catalysis by aminoacylase-1. Biochemistry *44*, 15645–15651.

Lindner, H.A., Alary, A., Wilke, M., and Sulea, T. (2008). Probing the acy1-binding pocket of aminoacylase-1. Biochemistry *47*, 4266–4275.

Lonsdale-Eccles, A., Leonard, N., and Lawrence, C. (2003). Axillary hyperhidrosis: Eccrine or apocrine? Clinical and Experimental Dermatology *28*, 2–7.

Lundgren, S., Andersen, B., Piškur, J., and Dobritzsch, D. (2007). Crystal structures of yeast β-alanine synthase complexes reveal the mode of substrate binding and large scale domain closure movements. Journal of Biological Chemistry *282*, 36037–36047.

Macfarlane, A. (1975). Olfaction in the development of social preferences in the human neonate. Parent - Infant Interaction 103–117.

Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Research *50*, W276–W279.

Makin, J.W., and Porter, R.H. (1989). Attractiveness of Lactating Females' Breast Odors to Neonates. Child Development *60*, 803.

Marples, M.J. (1969). Life on the human skin. Scientific American *220*, 108–115.

Martin, A., Saathoff, M., Kuhn, F., Max, H., Terstegen, L., and Natsch, A. (2010). A Functional ABCC11 Allele Is Essential in the Biochemical Formation of Human Axillary Odor. Journal of Investigative Dermatology *130*, 529–540.

Meredith, D., Temple, C.S., Guha, N., Sword, C.J., Boyd, C.A.R., Collier, I.D., Morgan, K.M., and Bailey, P.D. (2000). Modified amino acids and peptides as substrates for the intestinal peptide transporter PepT1. European Journal of Biochemistry *267*, 3723–3728.

Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.H., Goto, Y., Réfrégiers, M., and Kardos, J. (2018). BeStSel: A web server for accurate protein secondary structure prediction and fold

recognition from the circular dichroism spectra. Nucleic Acids Research *46*, W315–W322.

Minasov, G., Shuvalova, L., Brunzelle, J.S., Collart, F.R., and Anderson, W.F. (2005). RCSB PDB - 1YSJ: Crystal Structure of *Bacillus Subtilis* YXEP Protein (APC1829), a Dinuclear Metal Binding Peptidase from M20 Family.

Minhas, G.S., Bawdon, D., Herman, R., Rudden, M., Stone, A.P., James, A.G., Thomas, G.H., and Newstead, S. (2018). Structural basis of malodour precursor transport in the human axilla. ELife *7*, e34995.

Montagna, W., and Parakkal, P.F. (1974). The Structure & Function of Skin (Elsevier).

Moore, S., and Stein, W.H. (1948). Photometric ninhydrin method for use in the chromatography of amino acids. The Journal of Biological Chemistry *176*, 367–388.

Murdoch, D.A. (1998). Gram-positive anaerobic cocci. Clinical Microbiology Reviews *11*, 81–120.

Murphy, E.C., and Frick, I.M. (2013). Gram-positive anaerobic cocci - commensals and opportunistic pathogens. FEMS Microbiology Reviews *37*, 520–553.

Nagata, Y. (2003). Measurement of Odor Threshold by Triangle Odor Bag Method. Odor Measurement Review 118–127.

Natsch, A. (2015). What Makes Us Smell: The Biochemistry of Body Odour and the Design of New Deodorant Ingredients. CHIMIA International Journal for Chemistry *69*, 414–420.

Natsch, A., and Emter, R. (2020). The specific biochemistry of human axilla odour formation viewed in an evolutionary context. Philosophical Transactions of the Royal Society B: Biological Sciences *375*, 20190269.

Natsch, A., Gfeller, H., Gygax, P., Schmid, J., and Acuna, G. (2003). A specific bacterial aminoacylase cleaves odorant precursors secreted in the human axilla. Journal of Biological Chemistry *278*, 5718–5727.

Natsch, A., Schmid, J., and Flachsmann, F. (2004). Identification of odoriferous sulfanylalkanols in human axilla secretions and their formation through cleavage of cysteine precursors by a C-S lyase isolated from axilla bacteria. Chemistry and Biodiversity *1*, 1058–1072.

Natsch, A., Derrer, S., Flachsmann, F., and Schmid, J. (2006). A broad diversity of volatile carboxylic acids, released by a bacterial aminoacylase from axilla secretions, as candicate molecules for the determination of human-body odor type. Chemistry and Biodiversity *3*, 1–20.

Natsch, A., Kuhn, F., and Tiercy, J.M. (2010). Lack of evidence for HLA-linked patterns of odorous carboxylic acids released from glutamine conjugates secreted in the human axilla. Journal of Chemical Ecology *36*, 837–846.

Nayuni, N.K., Cloutman-Green, E., Hollis, M., Hartley, J., Martin, S., and Perrett, D. (2013). Critical evaluation of ninhydrin for monitoring surgical instrument decontamination. Journal of Hospital Infection *84*, 97–102.

Newstead, S. (2017). Recent advances in understanding proton coupled peptide transport via the POT family. Current Opinion in Structural Biology *45*, 17–24.

Nocek, B., Reidl, C., Starus, A., Heath, T., Bienvenue, D., Osipiuk, J., Jedrzejczak, R., Joachimiak, A., Becker, D.P., and Holz, R.C. (2018). Structural Evidence of a Major Conformational Change Triggered by Substrate Binding in DapE Enzymes: Impact on the Catalytic Mechanism. Biochemistry *57*, 574–584.

Noël, F., Piérard-Franchimont, C., Piérard, G.E., and Quatresooz, P. (2012). Sweaty skin, background and assessments. International Journal of Dermatology *51*, 647–655.

Occhipinti, E., Martelli, P.L., Spinozzi, F., Corsi, F., Formantici, C., Molteni, L., Amenitsch, H., Mariani, P., Tortora, P., and Casadio, R. (2003). 3D structure of Sulfolobus solfataricus carboxypeptidase developed by molecular modeling is confirmed by site-directed mutagenesis and small angle x-ray scattering. Biophysical Journal *85*, 1165–1175.

Olsson, M.J., Lundström, J.N., Kimball, B.A., Gordon, A.R., Karshikoff, B., Hosseini, N., Sorjonen, K., Olgart Höglund, C., Solares, C., Soop, A., et al. (2014). The Scent of Disease: Human Body Odor Contains an Early Chemosensory Cue of Sickness. Psychological Science *25*, 817–823.

Ono, E., Murota, H., Mori, Y., Yoshioka, Y., Nomura, Y., Munetsugu, T., Yokozeki, H., and Katayama, I. (2018). Sweat glucose and GLUT2 expression in atopic dermatitis: Implication for clinical manifestation and treatment. PLoS ONE *13*, e0195960.

Ouidir, T., Jarnier, F., Cosette, P., Jouenne, T., and Hardouin, J. (2015). Characterization of N-terminal protein modifications in *Pseudomonas aeruginosa* PA14. Journal of Proteomics *114*, 214–225.

Palm, G.J., and Röhm, K.H. (1995). Aminoacylase I from porcine kidney: Identification and characterization of two major protein domains. Journal of Protein Chemistry *14*, 233–240.

Parker, B.M., Taylor, I.N., Woodley, J.M., Ward, J.M., and Dalby, P.A. (2011). Directed evolution of a thermostable l-aminoacylase biocatalyst. Journal of Biotechnology *155*, 396–405.

Parkhitko, A.A., Jouandin, P., Mohr, S.E., and Perrimon, N. (2019). Methionine metabolism and methyltransferases in the regulation of aging and lifespan extension across species. Aging Cell *18*.

Patterson, M.J., Galloway, S.D.R., and Nimmo, M.A. (2000). Variations in regional sweat composition in normal human males. Experimental Physiology *85*, 869–875.

Pause, B.M., Storch, D., and Lübke, K.T. (2020). Chemosensory communication of aggression: Women's fine-tuned neural processing of male aggression signals. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Payne, J.W. (1974). Peptide transport in *Escherichia coli*: Permease specificity towards terminal amino group substituents. Journal of General Microbiology *80*, 269–276.

Penn, D.J., Damjanovich, K., and Potts, W.K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. Proceedings of the National Academy of Sciences of the United States of America *99*, 11260–11264.

Picardo, M., Ottaviani, M., Camera, E., and Mastrofrancesco, A. (1992). Sebaceous gland lipids. Seminars in Dermatology *11*, 100–105.

Porter, R.H., Makin, J.W., Davis, L.B., and Christensen, K.M. (1991). An assessment of the salient olfactory environment of formula-fed infants. Physiology and Behavior *50*, 907–911.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., et al. (2018). CCP 4 i 2: The new graphical user interface to the CCP 4 program suite. Acta Crystallographica Section D: Structural Biology *74*, 68–84.

Rahman, M., Ismat, F., Jiao, L., Baldwin, J.M., Sharples, D.J., Baldwin, S.A., and Patching, S.G. (2017). Characterisation of the DAACS Family *Escherichia coli* Glutamate/Aspartate-Proton Symporter GltP Using Computational, Chemical, Biochemical and Biophysical Methods. Journal of Membrane Biology *250*, 145–162.

Rawlings, N.D., and Barrett, A.J. (2013). Introduction: Metallopeptidases and Their Clans. In

Handbook of Proteolytic Enzymes, (Elsevier Ltd), pp. 325–370.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research *46*, D624–D632.

Reitzer, L. (2003). Nitrogen Assimilation and Global Regulation in *Escherichia coli*. Annual Review of Microbiology *57*, 155–176.

Rowsell, S., Pauptit, R.A., Tucker, A.D., Melton, R.G., Blow, D.M., and Brick, P. (1997). Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. Structure *5*, 337–347.

Rudden, M., Herman, R., Rose, M., Bawdon, D., Cox, D.S., Dodson, E., Holden, M.T.G., Wilkinson, A.J., James, A.G., and Thomas, G.H. (2020). The molecular basis of thioalcohol production in human body odour. Scientific Reports *10*, 12500.

Sandilands, A., Sutherland, C., Irvine, A.D., and McLean, W.H.I. (2009). Filaggrin in the frontline: Role in skin barrier function and disease. Journal of Cell Science *122*, 1285–1294.

Satinoff, E. (1978). Neural organization and evolution of thermal regulation in mammals. Science *201*, 16–22.

Sato, K., and Sato, F. (1987). Sweat secretion by human axillary apoeccrine sweat gland in vitro. American Journal of Physiology - Regulatory Integrative and Comparative Physiology *252*.

Sato, K., Leidal, R., and Sato, F. (1987). Morphology and development of an apoeccrine sweat gland in human axillae. American Journal of Physiology - Regulatory Integrative and Comparative Physiology *252*.

Sato, K., Kang, W.H., Saga, K., and Sato, K.T. (1989). Biology of sweat glands and their disorders. I. Normal sweat gland function. Journal of the American Academy of Dermatology *20*, 537–563.

Savić, B., Tomić, S., Magnus, V., Gruden, K., Barle, K., Grenković, R., Ludwig-Müller, J., and Salopek-Sondi, B. (2009). Auxin amidohydrolases from *Brassica rapa* cleave the alanine conjugate of indolepropionic acid as a preferable substrate: a biochemical and modeling approach. Plant & Cell Physiology *50*, 1587–1599.

Schaal, B., Marlier, L., and Soussignan, R. (1995). Responsiveness to the odour of amniotic fluid in the human neonate. Neonatology *67*, 397–406.

Schaal, B., Marlier, L., and Soussignan, R. (1998). Olfactory function in the human fetus: Evidence from selective neonatal responsiveness to the odor of amniotic fluid. Behavioral Neuroscience *112*, 1438–1449.

Schaal, B., Saxton, T.K., Loos, H., Soussignan, R., and Durand, K. (2020). Olfaction scaffolds the developing human from neonate to adolescent and beyond. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Schittek, B. (2012). The multiple facets of dermcidin in cell survival and host defense. Journal of Innate Immunity *4*, 349–360.

Schittek, B., Hipfel, R., Sauer, B., Bauer, J., Kalbacher, H., Stevanovic, S., Schirle, M., Schroeder, K., Blin, N., Meier, F., et al. (2001). Dermcidin: A novel human antibiotic peptide secreted by sweat glands. Nature Immunology *2*, 1133–1137.

Scholz, C.F.P., and Kilian, M. (2016). The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus propionibacterium to the proposed novel genera

acidipropionibacterium gen. Nov., *cutibacterium* gen. nov. and *pseudopropionibacterium* gen. nov. International Journal of Systematic and Evolutionary Microbiology *66*, 4422–4432.

Shehadeh, N.H., and Kligman, A.M. (1963). The effect of topical antibacterial agents on the bacterial flora of the axilla. The Journal of Investigative Dermatology *40*, 61–71.

Shelley, W.B., and Hurley, H.J. (1953). The physiology of the human axillary apocrine sweat gland. The Journal of Investigative Dermatology *20*, 285–297.

Shelley, W.B., Hurley, H.J., and Nichols, A.C. (1953). Axillary Odor: Experimental Study of the Role of Bacteria, Apocrine Sweat, and Deodorants. A. M. A. Archives of Dermatology and Syphilology *68*, 430–446.

Sheng, S., Kraft, J.J., and Schuster, S.M. (1993). A specific quantitative colorimetric assay for L-asparagine. Analytical Biochemistry *211*, 242–249.

Shirreffs, S.M., and Maughan, R.J. (1997). Whole body sweat collection in humans: An improved method with preliminary data on electrolyte content. Journal of Applied Physiology *82*, 336–341.

Siedlecka, E.M., Kumirska, J., Ossowski, T., Glamowski, P., Gołebiowski, M., Gajdus, J., Kaczyński, Z., and Stepnowski, P. (2008). Determination of volatile fatty acids in environmental aqueous samples. Polish Journal of Environmental Studies *17*, 351–356.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology *7*, 539.

Sigmarsdottir, T.B., McGarrity, S., de Lomana, A.L.G., Kotronoulas, A., Sigurðsson, S., Yurkovich, J.T., Rolfsson, O., and Sigurjonsson, O.E. (2021). Metabolic and transcriptional changes across osteogenic differentiation of mesenchymal stromal cells. Bioengineering *8*, 208.

Simunovic, M., Zagrovic, B., and Tomić, S. (2011). Mechanism and thermodynamics of ligand binding to auxin amidohydrolase. Journal of Molecular Recognition *24*, 854–861.

Singer, A.G., and Macrides, F. (1990). Aphrodisin: Pheromone or transducer? Chemical Senses *15*, 199–203.

Smith, D.E., Clémençon, B., and Hediger, M.A. (2013). Proton-coupled oligopeptide transporter family SLC15: Physiological, pharmacological and pathological implications. Molecular Aspects of Medicine *34*, 323–336.

Smith, M.W., Tyreman, D.R., Payne, G.M., Marshall, N.J., and Payne, J.W. (1999). Substrate specificity of the periplasmic dipeptide-binding protein from *Escherichia coli*: Experimental basis for the design of peptide prodrugs. Microbiology *145*, 2891–2901.

Spielman, A.I., Zeng, X.N., Leyden, J.J., and Preti, G. (1995). Proteinaceous precursors of human axillary odor: Isolation of two novel odor-binding proteins. Experientia *51*, 40–47.

Starkenmann, C., Niclass, Y., Troccaz, M., and Clark, A.J. (2005). Identification of the precursor of (S)-3-methyl-3-sulfanylhexan-1-ol, the sulfury malodour of human axilla sweat. Chemistry and Biodiversity *2*, 705–716.

Story, S. V., Grunden, A.M., and Adams, M.W.W. (2001). Characterization of an Aminoacylase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. Journal of Bacteriology *183*, 4259–4268.

Stowers, L., Cameron, P., and Keller, J.A. (2013). Ominous odors: Olfactory control of instinctive fear and aggression in mice. Current Opinion in Neurobiology *23*, 339–345.

Strauss, J.S., and Kligman, A.M. (1956). The bacteria responsible for apocrine odor. The Journal of Investigative Dermatology *27*, 67–71.

Sun, S.-W., Lin, Y.-C., Weng, Y.-M., and Chen, M.-J. (2006). Efficiency improvements on ninhydrin method for amino acid quantification. Journal of Food Composition and Analysis *19*, 112–117.

Tamames, B., Sousa, S.F., Tamames, J., Fernandes, P.A., and Ramos, M.J. (2007). Analysis of zinc-ligand bond lengths in metalloproteins: Trends and patterns. Proteins: Structure, Function and Genetics *69*, 466–475.

Tanimoto, K., Higashi, N., Nishioka, M., Ishikawa, K., and Taya, M. (2008). Characterization of thermostable aminoacylase from hyperthermophilic archaeon *Pyrococcus horikoshii*. FEBS Journal *275*, 1140–1149.

Taylor, D., Daulby, A., Grimshaw, S., James, G., Mercer, J., and Vaziri, S. (2003). Characterization of the microflora of the human axilla. International Journal of Cosmetic Science *25*, 137–145.

Tewari, Y.B. (1990). Thermodynamics of industrially-important, enzyme-catalyzed reactions. Applied Biochemistry and Biotechnology: Part A: Enzyme Engineering and Biotechnology *23*, 187–203.

Thody, A.J., and Shuster, S. (1989). Control and function of sebaceous glands. Physiological Reviews *69*, 383–416.

Tidjani Alou, M., Khelaifia, S., Michelle, C., Andrieu, C., Armstrong, N., Bittar, F., Sokhna, C., Diallo, A., Fournier, P.E., Raoult, D., et al. (2016). *Anaerococcus rubiinfantis* sp. nov., isolated from the gut microbiota of a Senegalese infant with severe acute malnutrition. Anaerobe *40*, 85–94.

Tomić, A., Brkić, H., Matić, A., and Tomić, S. (2021). Unravelling the inhibitory zinc ion binding site and the metal exchange mechanism in human DPP III. Physical Chemistry Chemical Physics *23*, 13267–13275.

Tonzetich, J., and Richter, V.J. (1964). Evaluation of volatile odoriferous components of saliva. Archives of Oral Biology *9*, 39–45.

Toogood, H.S., Hollingsworth, E.J., Brown, R.C., Taylor, I.N., Taylor, S.J.C., McCague, R., and Littlechild, J.A. (2002). A thermostable L-aminoacylase from *Thermococcus litoralis*: Cloning, overexpression, characterization, and applications in biotransformations. Extremophiles *6*, 111–122.

Troccaz, M., Starkenmann, C., Niclass, Y., Van De Waal, M., and Clark, A.J. (2004). 3-methyl-3-sulfanylhexan-1-ol as a major descriptor for the human axilla-sweat odour profile. Chemistry and Biodiversity *1*, 1022–1035.

Troccaz, M., Borchard, G., Vuilleumier, C., Raviot-Derrien, S., Niclass, Y., Beccucci, S., and Starkenmann, C. (2009). Gender-specific differences between the concentrations of nonvolatile (R)/(S)-3-Methyl-3-Sulfanylhexan-1-Ol and (R)/(S)-3-Hydroxy-3-Methyl-Hexanoic acid odor precursors in axillary secretions. Chemical Senses *34*, 203–210.

Troccaz, M., Gaïa, N., Beccucci, S., Schrenzel, J., Cayeux, I., Starkenmann, C., and Lazarevic, V. (2015). Mapping axillary microbiota responsible for body odours using a culture-independent approach. Microbiome *3*, 3.

Vaglio, S., Minicozzi, P., Kessler, S.E., Walker, D., and Setchell, J.M. (2021). Olfactory signals and fertility in olive baboons. Scientific Reports *11*, 8506.

Valdez, C.E., Smith, Q.A., Nechay, M.R., and Alexandrova, A.N. (2014). Mysteries of metals in metalloenzymes. Accounts of Chemical Research *47*, 3110–3117.

Veloo, A.C.M., de Vries, E.D., Jean-Pierre, H., and van Winkelhoff, A.J. (2016). *Anaerococcus nagyae*

sp. nov., isolated from human clinical specimens. Anaerobe *38*, 111–115.

Walling, L.R., Kouse, A.B., Shabalina, S.A., Zhang, H., and Storz, G. (2022). A 3 UTR-derived small RNA connecting nitrogen and carbon metabolism in enteric bacteria. Nucleic Acids Research *50*, 10093–10109.

Warneford, J. (1916). The ninhydrin reaction with amines and amides. Journal of Biological Chemistry *25*, 337–350.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.

Wedekind, C., Seebeck, T., Bettens, F., and Paepke, A.J. (1995). MHC-dependent mate preferences in humans. Proceedings of the Royal Society B: Biological Sciences *260*, 245–249.

Wildeboer-Veloo, A.C.M., Harmsen, H.J.M., Welling, G.W., and Degener, J.E. (2007). Development of 16S rRNA-based probes for the identification of Gram-positive anaerobic cocci isolated from human clinical specimens. Clinical Microbiology and Infection *13*, 985–992.

Williams, J., and Ringsdorf, A. (2020). Human odour thresholds are tuned to atmospheric chemical lifetimes. Philosophical Transactions of the Royal Society B: Biological Sciences *375*.

Yamazaki, K., Boyse, E.A., Miké, V., Thaler, H.T., Mathieson, B.J., Abbott, J., Boyse, J., Zayas, Z.A., and Thomas, L. (1976). Control of mating preferences in mice by genes in the major histocompatibility complex. Journal of Experimental Medicine *144*, 1324–1335.

Yokoyama, Y., Aragaki, M., Sato, H., and Tsuchiya, M. (1991). Determination of sweat constituents by liquid ionization mass spectrometry. Analytica Chimica Acta *246*, 405–411.

Yoshiura, K.I., Kinoshita, A., Ishida, T., Ninokata, A., Ishikawa, T., Kaname, T., Bannai, M., Tokunaga, K., Sonoda, S., Komaki, R., et al. (2006). A SNP in the ABCC11 gene is the determinant of human earwax type. Nature Genetics *38*, 324–330.

Zeden, M.S., Kviatkovski, I., Schuster, C.F., Thomas, V.C., Fey, P.D., and Gründling, A. (2020). Identification of the main glutamine and glutamate transporters in *Staphylococcus aureus* and their impact on c-di-AMP production. Molecular Microbiology *113*, 1085–1100.

Zeng, C., Spielman, A.I., Vowels, B.R., Leyden, J.J., Biemann, K., and Preti, G. (1996). A human axillary odorant is carried by apolipoprotein D. Proceedings of the National Academy of Sciences of the United States of America *93*, 6626–6630.

Zeng, X., Leyden, J.J., Lawley, H.J., Sawano, K., Nohara, I., and Preti, G. (1991). Analysis of characteristic odors from human male axillae. Journal of Chemical Ecology *17*, 1469–1492.

Zhang, Y., Fu, Y., Zhou, S., Kang, L., and Li, C. (2013). A straightforward ninhydrin-based method for collagenase activity and inhibitor screening of collagenase using spectrophotometry. Analytical Biochemistry *437*, 46–48.

Zouboulis, C.C., Baron, J.M., Böhm, M., Kippenberger, S., Kurzen, H., Reichrath, J., and Thielitz, A. (2008). Frontiers in sebaceous gland biology and pathology. Experimental Dermatology *17*, 542–551.