Non-enzymatic formation of N-acylated amino acid conjugates in urine

J Jacobs

orcid.org 0000-0001-8786-2866

Dissertation submitted in partial fulfilment of the requirements for the degree *Masters of Science in Biochemistry* at the North-West University

Supervisor:	Prof BC Vorster
Co-supervisor:	Dr M Dercksen
Assistant supervisor:	Dr CGCE van Sittert

Graduation May 2018

20355572

PREFACE

I would like to acknowledge and thank the following people, without whom this study would not have been possible:

- My wife Marlise Jacobs and daughter Amira Jacobs for your love, support, patience and understanding.
- My mentor, Prof. L.J. Mienie, for educating me in the fields of biochemistry, analytical and organic chemistry. Prof. Mienie also initiated this study and gave full support throughout.
- My supervisors:
    - Prof. B.C. Vorster, for taking on this project, for help and support (both technical and emotional) and ensuring a high level of quality throughout.
    - Dr M. Dercksen, for your contributions and expertise in metabolism, and your high level of attention to detail.
    - Dr C.G.C.E. van Sittert, for teaching me the ins and outs of molecular modelling, and your contributions in evaluating the proposed reaction mechanisms.
- All the staff at the Potchefstroom Laboratory for Inborn Errors of Metabolism (PLIEM) – my second family – for your love, support and patience and for taking up the workload in my absence.

This project would also not have been possible without the PLIEM, for its funding, facilities and sample archive.

ABSTRACT

Comprehensive metabolic profiling is routinely utilised for the diagnosis of inborn errors of metabolism (IEMs). Observed disease-specific constituents are mostly due to induced secondary pathways, resulting in the build-up of these metabolites, which may contribute to a phenotypic presentation. There are however cases in which no enzymatic pathway is identified as a contributing factor. This is true for some N-acylated amino acid (N-AAA) conjugates, reported in isolated cases of maple syrup urine disease (MSUD). The aim of this study was to identify a potential mechanism for the formation of N-AAA conjugates, identified in urine of South African MSUD patients. Emphasis was placed on the general stereo-nonspecific nature of non-enzymatic reactions yielding racemic mixtures. A strategic approach was subsequently employed in which the enantiomeric composition of N-AAA conjugates was determined, to establish the origin of these compounds.

Several applications were utilised to identify the enantiomeric composition of the N-AAA conjugates. These included (1) a liquid-liquid extraction of N-AAA conjugates, followed by acid-hydrolysis to liberate the conjugated amino acids, (2) the separation of amino acid enantiomers by chiral derivatisation via gas chromatography-mass spectrometry (GC-MS) and (3) molecular modelling to assess the reaction mechanism for the non-enzymatic formation from 2-keto acids and ammonia. The organic acid extraction method yielded adequate amounts of N-AAA conjugates without concomitant extraction of native amino acids. Hydrolysis was complete without significant hydrolysis-induced racemisation. Amino acid enantiomers were distinguishable through GC-MS analysis with limitations noted in L-isoleucine and D-*allo*-isoleucine. After standardisation of the methods, this chiral strategy was employed to investigate an available MSUD case, which was found to contain racemic N-AAA conjugates. From the results, it was deduced that the N-AAA conjugates were indeed from non-enzymatic origin. The findings also illustrate the usefulness of a chiral strategy and molecular modelling in investigating the origin of unknown constituents in biological samples. These conjugates can now be studied as a potential disease contributing factor in MSUD and other IEMs.

Key terms:

N-acylated amino acid conjugates, Non-enzymatic reactions, Maple syrup urine disease, Inborn errors of metabolism, Racemic resolution.

OPSOMMING

Omvattende metaboliese profilering word as ’n roetine benadering vir die diagnosering van aangebore metaboliese siektes gebruik. Siekte-spesifieke metaboliete wat so waargeneem word, kan meestal aan geïnduseerde sekondêre weë toegeskryf word, wat lei tot die opbou van hierdie metaboliete en ’n bydrae lewer tot fenotipiese presentering. Daar is egter gevalle waar geen ensiemweg as ’n bydraende faktor geïdentifiseer kan word nie. Dit is ook die geval vir sekere N-geasileerde aminosuurkonjugate (N-AAA), wat tevore in sekere geïsoleerde gevalle van esdoringstroopsiekte (MSUD) waargeneem is. Die doel van hierdie studie was om ’n potensiële meganisme vir die vorming van N-AAA-konjugate (waargeneem in die uriene van Suid-Afrikaanse MSUD-pasiënte) te identifiseer. Daar is klem gelê op die verskynsel dat spontane reaksies lei tot die vorming van rasemiese mengsels. ’n Strategiese benadering is onderneem waarin die enantiomeersamestelling van die N-AAA-konjugate bepaal is om die oorsprong van die konjugate te bepaal.

Verskeie toepassings is aangewend om die enantiomeersamestelling van die N-AAA-konjugate vas te stel. Dit sluit in (1) ’n vloeistof-vloeistof ekstraksie van N-AAA-konjugate, gevolg deur suur-hidrolise om die gekonjugeerde aminosure vry te stel; (2) die skeiding van aminosuur-enantiomere deur chirale derivatisering via gaschromatografie-massaspektrometrie (GC-MS); en (3) molekuulmodellering om die reaksiemeganisme vir die nie-ensimatiese vorming vanaf 2-ketosure en ammoniak te evalueer. Die organiese-suur-metode het genoegsame hoeveelhede N-AAA-konjugate gelewer sonder gepaardgaande ekstraksie van aminosure wat oorspronklik deel van die monster was. Hidrolise was volledig sonder merkwaardige hidrolise-geïnduseerde rasemering. Onderskeid kon getref word tussen die aminosuur-enantiomere deur GC-MS analise, met beperkings wat in L-isoleusien en D-*allo*-isoleusien skeiding opgemerk is. Na afloop van metode-standaardisering is hierdie chirale strategie toegepas om ’n beskikbare MSUD geval te ondersoek, wat gelei het tot die bevinding dat die monster rasemiese N-AAA-konjugate bevat. Die resultate het gelei tot die gevolgtrekking dat die N-AAA-konjugate wel vanaf ’n nie-ensimatiese oorsprong ontstaan. Die bevindinge illustreer ook die nut van ’n chirale strategie en molekuulmodellering tydens die ondersoek van metaboliete met onbekende oorsprong in biologiese monsters. Hierdie konjugate kan dus nou verder ondersoek word as ’n potensiële bydraende faktor in MSUD en ander aangebore metaboliese siektes.

Sleutelterme:

N-asiel aminosuur konjugate, Nie-ensimaties reaksies, Esdoringstroopsiekte, Aangebore metaboliese siektes, Rasemiese skeiding.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

- ΔG : Gibbs free energy
- ΔG_{rel} : Relative Gibbs free energy
- 4HNE : 4-Hydroxy-2-nonenal
- Ac : Acetyl group
- ACY1, 2 or 3 : Aminoacylase I, II or III
- aIle : *allo*-Isoleucine
- Ala : Alanine
- AMDIS : Automated mass spectral deconvolution and identification system
- Arg : Arginine
- Asn : Asparagine
- Asp : Aspartate
- BCAA : Branched-chain amino acid
- BCKDH : Branched-chain α-keto acid dehydrogenase complex
- BD : Binding domain
- BSTFA : N,O-bis(trimethylsilyl)trifluoroacetamide
- CoA : Coenzyme A ester
- CoASH : Coenzyme A (free unbounded CoA in reduced form)
- COSMO : Conductor-like screening model
- CSM : Continuum solvation model
- Cys : Cysteine
- ddH_2O : Double-distilled water
- DFT : Density functional theory
- DIIS : Direct inversion in interactive subspace
- DNP : Double-numerical basis set with polarisation functions
- E_1 : 2-Keto acid dehydrogenase
- E_2 : Dihydrolipoyl transacylase

LIST OF ABBREVIATIONS (CONTINUED)

- E_3 : Dihydrolipoamide dehydrogenase
- α-TPP : Thiamine pyrophosphate
- E_a : Activation energy
- E_e : Electron energy
- EC number : Enzyme Commission number
- EI : Electron ionisation
- EMV : Electron multiplier voltage
- ESP : Electrostatic potential
- EZ:faast : Phenomenex EZ:faast free (physiological) amino acid analysis kit
- FAD : Flavin adenine dinucleotide
- $FADH_2$: Oxidised flavin adenine dinucleotide
- GGA : Generalised gradient approximation
- GC : Gas chromatography
- GC-MS : Gas chromatography-mass spectrometry
- Gln : Glutamine
- Glu : Glutamate
- Gly : Glycine
- Ha : Hartree
- HOMO : Highest occupied molecular orbital
- HPLC : High pressure liquid chromatography
- IEM : Inborn errors of metabolism
- Ile : Isoleucine
- LD : Lipoic acid binding domain
- Leu : Leucine
- LUMO : Lowest unoccupied molecular orbital
- MSD : Mass selective detector
- MS-MS : Tandem mass spectrometry

LIST OF ABBREVIATIONS (CONTINUED)

- MSUD : Maple syrup urine disease
- MTBSTFA : N-*tert*-butyldimethylsilyl-N-methyltrifluoroacetamide
- N.D. : Not determined/Not detected
- n.o.s. : Not otherwise specified
- N-AAA : N-Acylated amino acid
- N-AcAA : N-Acetylated amino acid
- NAD : Nicotinamide adenine dinucleotide
- NADH + H^+ : Oxidised nicotinamide adenine dinucleotide
- NAPQI : N-Ac-p-benzoquinone imine
- NIST : National institute of standards and technology
- Nva : Norvaline
- OBS : Ortmann, Bechstedt and Schmidt
- OMIM : Online Mendelian Inheritance in Man
- PES : Potential energy surface
- Phe : Phenylalanine
- R^x : R - Represents an unspecified aliphatic group attached to the rest of a molecule, unless stated otherwise. The superscript 'x' denotes the number of a group, if more than one different groups exists in the diagram.
- ROS : Reactive oxygen species
- Ser : Serine
- SCF : Self-consistent field
- SPE : Solid phase extraction
- t-BDMCS : *tert*-Butyldimethylchlorosilane
- tBDMS : *tert*-butyldimethylsilyl group
- TMCS : Trimethylchlorosilane
- TMS : Trimethylsilyl
- TS : Transition state
- U : Universal

LIST OF ABBREVIATIONS (CONTINUED)

- UDP : Uridine 5'-diphosphate
- UV : Ultraviolet
- Val : Valine

CHAPTER 1: NAMING CONVENTIONS

1.1 N-Acylated and N-acetylated amino acid conjugates

Refer to Figure 1: N-**Acetyl**ated amino acid (N-AcAA) conjugates are part of the larger N-**acyl**ated amino acid (N-AAA) conjugate group. Due to the metabolic relevance of N-AcAA conjugates, this group will be discussed separately and care should be taken to avoid confusing 'acyl' with the similar-sounding 'acetyl' (Ac).

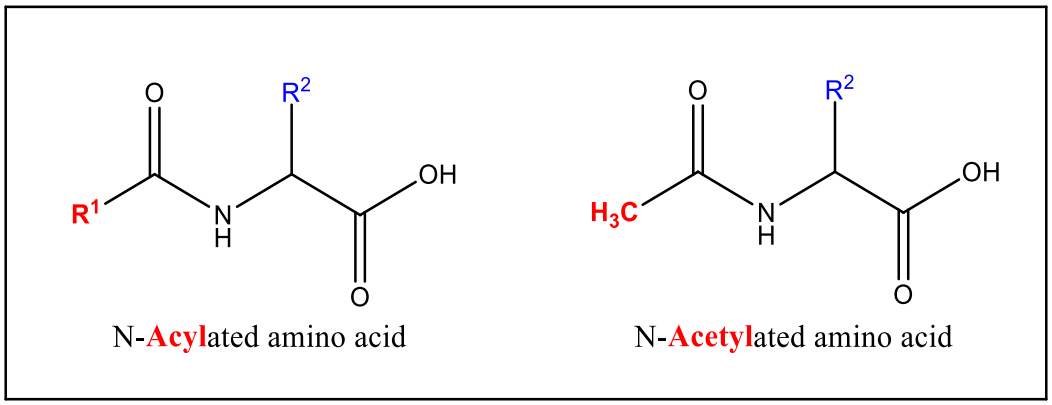

Figure 1: Difference between N-AAA (left) and N-AcAA conjugates (right) as defined for this study.

In the image above, 'R^2' represents any aliphatic side chain and 'R^1' represents any aliphatic side chain excluding methyl.

The following main N-AAA conjugates are discussed in this dissertation: N-isobutyrylvaline, N-isobutyrylleucine, N-isobutyrylisoleucine, N-isovalerylvaline, N-isovalerylleucine, N-isovaleryl-isoleucine, N-2-methylbutyrylvaline, N-2-methylbutyrylleucine and N-2-methylbutyrylisoleucine. The following main N-AcAA conjugates are discussed in this dissertation: N-acetylalanine, N-acetylvaline, N-acetylleucine, N-acetylisoleucine and N-acetylglycine.

1.2 Biotransformation and detoxification

Toxic compounds accumulate in the body as a natural result of metabolism, the ingestion of medication, preservatives in food, the environment and inborn errors of metabolism (IEMs). Biotransformation is the biochemical modification of a compound to (in most cases) decrease its harmful effect and/or to increase its solubility for excretion in urine. The term 'biotransformation' is preferred to 'detoxification' since some metabolites' toxicity may increase after the first phase of biotransformation (Liska *et al.*, 2006:122-123).

1.3 Primary, alternative, induced and secondary metabolic pathways and metabolites

These terms are used both interchangeably and with different meanings in the literature. For clarity in this dissertation, these terms are defined as follows:

- The primary metabolic pathway is the metabolic pathway currently being studied or discussed.

- An alternative metabolic pathway is a metabolic pathway that can metabolise a metabolite of the primary metabolic pathway, e.g. when considering the glycolysis pathway, the pentose phosphate pathway is the alternative metabolic pathway of glucose-6-phosphate and *vice versa*. Alternative metabolites are the intermediates and end products of an alternative metabolic pathway.

- Induced metabolic pathways are alternative metabolic pathways that are only active when a substrate of the primary pathway accumulates to a certain threshold concentration, e.g. the formation of succinylacetone from fumarylacetoacetate in tyrosinemia type I.

- Secondary metabolites are metabolites that are not directly essential for energy production or growth, i.e. the survival of the organism, e.g. pigments, toxins and antibiotics. Secondary metabolic pathways are the metabolic pathways of secondary metabolites (Wolfender *et al*., 2015:137).

1.4 Transition states, reaction intermediates and metabolic intermediates

Transition states (TSs) are sometimes erroneously referred to as intermediates. Intermediates are short-lived but well-defined chemical species that form during a reaction, whereas a TS represents the highest-energy, unstable structure (which cannot be isolated) in the reaction coordinate in-between intermediates (McMurry, 2004:156). Metabolites in a metabolic pathway, which are stable molecules, are also referred to as intermediates. For clarity in this dissertation, intermediates of a reaction mechanism are referred to as ‘reaction intermediates’ and metabolites are referred to as ‘metabolic intermediates’ or ‘metabolic substrates/products’, except where the context is obvious.

1.5 Geometrical minimisation and optimisation

In relation to the molecular modelling section of this dissertation, the term ‘geometrical minimisation’ should not be confused with ‘geometrical optimisation’ when searching for a global minimum on the energy hyper-surface, since geometrical optimisation techniques are also used to find TSs, which are not at an energy minimum.

1.6 Chiral centres

Traditionally, biological compounds with chiral centres such as amino acids and carbohydrates are classified as *levorotatory* (L) and *dextrorotatory* (D) based on their structural similarity with L- or D-glyceraldehyde when represented as a Fischer projection (Cava *et al.*, 2010:818; McMurry, 2004:944-949). Another, more systematic, approach is the *rectus* (*R*) and *sinister* (*S*) convention, which uses the Cahn-Ingold-Prelog priority rules. Based on atomic number, this unambiguous system avoids confusion. Confusion between the two systems can exists because of compounds like cysteine (Cys). Cys contains a sulphur atom on the second position of its side chain, giving the side-chain group a higher priority than the carboxyl group. Because of this, L-Cys is *R*-orientated instead of *S*-orientated like most other canonical amino acids with the same configuration. For this reason, the glyceraldehyde reference (DL-system) is favoured for amino acid, carbohydrate and enzyme names and are labelled as such in this dissertation. To avoid ambiguity, however, compounds in the molecular modelling section, diastereoisomers and chemical reagents are labelled with the *RS*-system. Please note that chiral labelling is formatted with small capital letters to avoid confusion between sulphur-atom conjugates and *sinister*-chiral centres, e.g. in the diastereoisomer S-[(1*S*,2*R*)-2-hydroxycyclohexyl]-Cys, the first 'S' denotes that the 2-hydroxycyclohexyl group is conjugated to Cys on the sulphur atom, whereas the second '*S*' denotes chirality.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

N-AAA conjugates are carboxamides consisting of an acyl group conjugated to an amino acid via a peptide bond (Schäfer & Bode, 2014:1526). Glycine N-acyltransferase (GLYAT) (Enzyme Commission number [EC]: 2.3.1.13) catalyses the conjugation of glycine (Gly) to various acyl-coenzyme A (CoA) substrates in Gly-mediated biotransformations. These biotransformation processes salvage CoAs and increases the solubility of carboxylic acids for excretion in urine (Liska *et al*., 2006:126-133). Other amino acids such as alanine (Ala), glutamine (Gln), serine (Ser) and branched-chain amino acids (BCAAs) have been described as substrates for GLYAT but at lower reaction rates than for Gly (Sweetman & Williams, cited by Dry, 1997:108; Van der Westhuizen *et al*., 2000:102-103). The biotransformation of acyl-CoAs into their respective N-acyl-Gly conjugates is especially evident in IEMs where large amounts of acyl-CoAs accumulate (Badenhorst *et al.*, 2013:346-354). Examples include the biotransformation of N-isovaleryl-CoA in isovaleric acidemia (Online Mendelian Inheritance in Man number [OMIM]: 243500), propionyl-CoA in propionic acidemia (OMIM: 606054) and various acyl-CoAs in multiple acyl-CoA dehydrogenase deficiency (OMIM: 231680) and other IEMs (Guder, 1998:26-48).

Dry (1997) discovered a variant maple syrup urine disease (MSUD) (OMIM: 248600) that presented with, amongst other things, elevated levels of N-AAA conjugates of the BCAAs. This variant case was further characterised by Deysel (2001) and Mienie *et al*. (2001). Since amino acids other than Gly have been described as substrates for GLYAT, it seems plausible that, given their elevated levels in MSUD, BCAAs could have been utilised by GLYAT in this case. Due to the higher reaction rate of Gly with acyl-CoAs, higher levels of Gly than BCAAs conjugates would then also be expected. This, however, was not the case. In addition, the formation of the respective acyl-CoA substrates required for GLYAT-mediated biotransformation is prevented by MSUD itself (Dry, 2001:138; Mienie *et al.*, 2001:51; Patrick, 1960:269). The mechanism for the formation of the N-AAA conjugates described by both Dry and Mienie thus remains to be determined.

It is important to study variant metabolic presentations of known diseases to better correlate the clinical-biochemical image, uncover new diagnostic markers, measure responsiveness to treatment, develop therapeutic interventions and increase understanding of metabolism that may be of value elsewhere. A review of MSUD and N-AAA conjugates in IEMs may shed some light on the subject.

2.2 Inborn errors of metabolism associated with N-AAA conjugates

The urinary excretion of various N-AAA conjugates is associated with IEMs as summarised in Table 1. Elevated levels of N-AAA conjugates are observed if aminoacylases (the enzymes that catalyse the hydrolysis of N-AAA conjugates) are deficient or if their capacity is exceeded, e.g. during fasting or because of an IEM (Perrier *et al*., 2005:677). Various N-AcAA conjugates are elevated in aminoacylase I (ACY1) deficiency (OMIM: 609924) (Gerlo *et al.*, 2006:191-198) and N-Ac-aspartate (Asp) is elevated in Canavan disease (Hagenfeldt *et al.*, 1967:135; Matalon *et al.*, 1988:463), as described in sections 2.3.1.1 and 2.3.1.2 later on. Most aminoacidopathies cause the conversion of a small fraction of the elevated amino acid(s) into its (their) respective N-AcAA conjugate(s), resulting in elevated urinary excretion (Jellum *et al.*, 1986:21). As an example: phenylketonuria (OMIM: 261600) is caused by a deficiency of phenylalanine-4-monooxygenase, which catalyses the conversion of phenylalanine (Phe) into tyrosine. This results in elevated levels of Phe, which are then conjugated to Ac-CoA by Phe-N-Ac-transferase. Tyrosinemia type II (OMIM: 276600) also results in elevated levels of N-Ac-Phe and -tyrosine (Jellum *et al.*, 1986:21). Other aminoacidopathies resulting in elevated levels of the respective N-AcAA conjugates include hyperlysinemia (OMIM: 238700), histidinemia (OMIM: 235800) and citrullinemia (OMIM: 215700) (Gerlo *et al.*, 2006:191-198; Jellum *et al.*, 1986:21; Salway, 2008:90-91). However, certain aminoacidopathies, such as organic acidemia, can also result in elevated levels of N-AAA conjugates. In isovaleric acidemia, N-isovaleryl-Gly is highly elevated, with lower levels of N-isovaleryl-Gln and -Ala as products of GLYAT-mediated biotransformation (Guder, 1998:34; Van der Westhuizen *et al.*, 2000:102). In propionic acidemia, N-propionyl-Gly and N-tiglyl-Gly are highly elevated (Guder, 1998:48). The N-acetylated conjugates of the BCAAs leucine (Leu), isoleucine (Ile) and valine (Val) are elevated in most cases of MSUD (Jellum *et al.* 1986:184).

A trend seems to be emerging: aminoacidopathies result in elevated levels of N-AcAA conjugates of the respective elevated amino acids, whereas organic acidemias result in elevated levels of N-acyl-Gly conjugates of the respective elevated acyl-CoAs. The only exceptions seem to be N-phenylacetyl-Gln in phenylketonuria catalysed by Gln-N-phenylacetyl-transferase (EC: 2.3.1.14) (Guder, 1998:46), N-isovaleryl-glutamate (-Glu) and -Ala in isovaleric acidemia (Du Toit *et al*., 2005:1510-1511) and the N-acylated-BCAA conjugates described in the variant MSUD cases by Deysel (2001), Dry (1997) and Mienie *et al*. (2001). Since the key sample in this study was related to MSUD, a discussion of normal BCAA metabolism and the related MSUD follows below.

Table 1: Inborn errors of metabolism resulting in elevated levels of N-AAA conjugates.

Inborn error of metabolism/pathology	OMIM #	Major/significant N-AAA conjugate(s)	Reference
Fasting	N/A	Various N-AcAA conjugates	Perrier *et al*. (2005:677)
Canavan disease	271900	N-Ac-Asp	Gerlo *et al.* (2006:191); Sass *et al.* (2006:401)
Phenylketonuria	261600	N-Ac-Phe N-Phenylacetyl-Gln	Gerlo *et al.* (2006:191); Guder (1998:46); Jellum *et al.* (1986:21)
Tyrosinemia type I, II and III	276600	N-Ac-Phe and -tyrosine	Jellum *et al.* (1986:21); Gerlo *et al.* (2006:191)
Hyperlysinemia	238700	N-Ac-Lys	Jellum *et al.* (1986:21)
Histidinemia	235800	N-Ac-histidine	Jellum *et al.* (1986:21)
Citrullinemia	215700	N-Ac-citrulline	Jellum *et al.* (1986:21)
Isovaleric acidemia	243500	N-Isovaleryl-Gly, -Glu and -Ala	Tanaka and Isselbacher (1967:2966)
Propionic acidemia	606054	N-Propionyl- and N-tiglyl-Gly	Rasmussen *et al.* (1972:665); Sweetman *et al.* (1978:198)
3-Methylcrotonyl-CoA carboxylase deficiency	210200	N-3-Methylcrotonyl-Gly	Eldjarn *et al.* (1970:521)
Beta-ketothiolase deficiency/methylacetoacetic aciduria	203750	N-Tiglyl-Gly	Gompertz *et al.* (1974:269)
Biotinidase deficiency	253260	N-3-Methylcrotonyl-Gly and N-tiglyl-Gly	Sweetman *et al.* (1977:1144)
Multiple acyl-CoA dehydrogenase deficiency	231680	N-Isovaleryl-, N-isobutyryl- and N-2-methylbutyryl-Gly	Goodman *et al.* (1980:12)

Inborn error of metabolism/pathology	OMIM #	Major/significant N-AAA conjugate(s)	Reference
Maple syrup urine disease	248600	N-Ac-Gly, -Leu, -Ile and -Val N-Lactyl-Gly, -Leu, -Ile and -Val N-2-Hydroxy-isovaleryl-Gly, -Leu, -Ile and -Val. **Variant:** N-2-Hydroxy-hexanoyl-Leu, -Ile and -Val N-2-Hydroxy-3-methylvaleryl-Leu N-Isobutyryl-Leu, -Ile and -Val N-2-Methylbutyryl-Leu, -Ile and -Val N-Isovaleryl- Leu, -Ile and -Val	Deysel (2000:18); Dry (1997:9-95); Hagenfeldt and Naglo (1987:77)
3-Hydroxy-3-methylglutaryl-CoA lyase deficiency	246450	**In severe cases:** N-3-methylcrotonyl-Gly	Gibson *et al.* (1988:76)
Medium chain acyl-CoA dehydrogenase deficiency	201450	N-Hexanoyl-, N-phenylpropionyl- and N-suberyl-Gly	Rinaldo *et al.* (1988:1308)

2.2.1 Branched chain amino acid metabolism

The catabolism of BCAAs is illustrated in Figure 2. Normal BCAA catabolism starts with the cytosolic deamination of the BCAAs Ile, Val and Leu to their corresponding 2-keto acids (2-keto-3-methylvaleric acid, 2-ketoisovaleric acid and 2-ketoisocaproic acid). This reaction is catalysed by BCAA transaminase (EC: 2.6.1.42). During this reaction, 2-ketoglutaric acid acts as an amine acceptor leading to the formation of Glu. The 2-keto acids formed are then transported into the mitochondria via the carnitine shuttle, where they are decarboxylated and oxidised by branched-chain α-keto acid dehydrogenase complex (BCKDH) (EC: 1.2.4.4). The respective CoAs (2-methylbutyryl-, isobutyryl- and isovaleryl-CoA) are subsequently formed, together with the reduction of nicotinamide adenine dinucleotide (NAD) to oxidised NAD ($NADH + H^+$). These CoAs are further catabolised via a series of enzyme-mediated reactions to Ac-CoA, succinyl-CoA and acetoacetate, respectively, and the further reduction of more NAD to $NADH + H^+$ and flavin adenine dinucleotide (FAD) to oxidised FAD ($FADH_2$). These final products of BCAA catabolism are then utilised as an energy source for production or growth, depending on the cell's current needs (Chuang *et al.*, 2012; Salway, 2004:80-88; Vockley *et al.*, 2012).

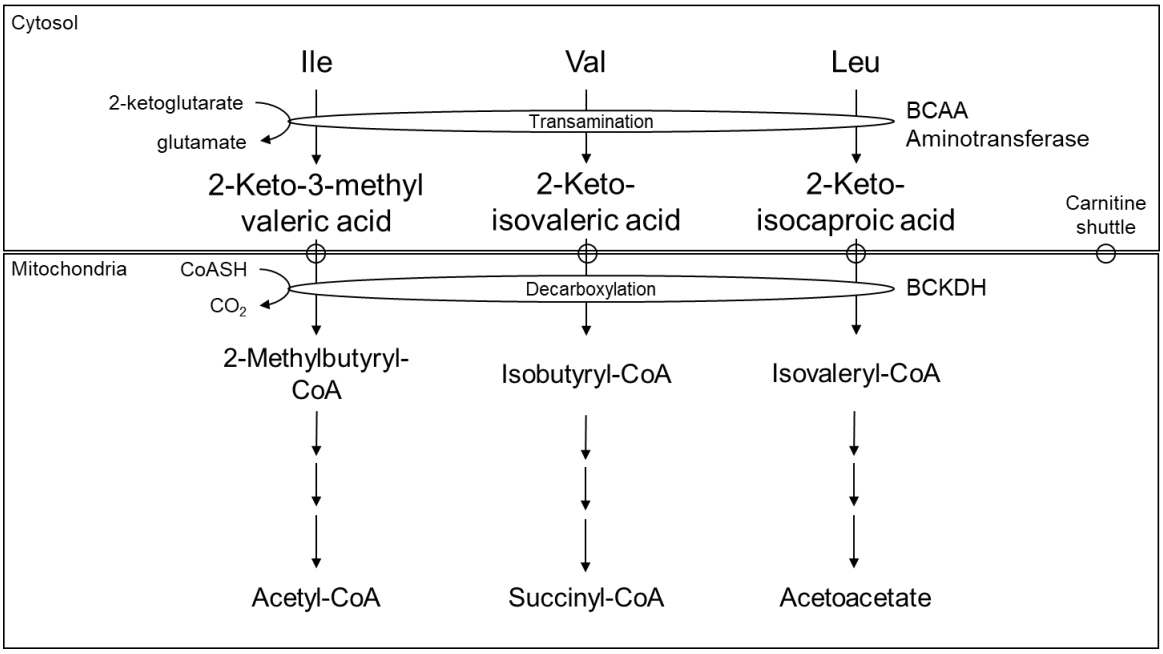

Figure 2: Simplified illustration, adapted from Salway (2004:80-88), of branched-chain amino acid (BCAA) catabolism in liver and muscle cells.

BCKDH: branched-chain α-keto acid dehydrogenase complex; CoA: coenzyme A ester; CoASH: coenzyme A (free unbounded CoA in reduced form).

The structure, function and reaction mechanism of BCKDH are illustrated in Figure 3 and Figure 4. BCKDH is a 4 million Da multi-subunit enzyme complex located on the inner membrane of mitochondria (Ævarsson *et al.*, 2000:277-278). It consists of three catalytic subunits, namely a thiamine pyrophosphate (α-TPP), dependent 2-keto acid dehydrogenase (E_1), dihydrolipoyl transacylase (E_2) and dihydrolipoamide dehydrogenase (E_3). Twenty-four E_2 subunits are arranged in an octahedral symmetry at the centre of the complex, each with a binding domain for E_1/E_3 (BD), a highly mobile lipoic acid binding domain (LD) for the covalently bound lipoic acid coenzyme and an inner-core domain (linked to BD and LD with two flexible inter-domain segments). The inner-core domain of E_2 catalyses the acyltransferase reaction. Twelve copies of E_1 and six copies of E_3 (with a bound FAD coenzyme) are non-covalently linked to the BDs. BCKDH phosphatase and kinase regulates the enzyme complex's activity (Ævarsson *et al.*, 2000:277-278).

Refer to Figure 3 and Figure 4 for the following brief description of the reaction mechanism of BCKHD: (1) E_1 binds the coenzyme α-TPP and a branched-chain aliphatic 2-keto acid. The cationic C2-carbon of the thiazolium ring of thiamine performs a nucleophilic attack on the C2-carbonyl carbon of the keto acid. The resulting conjugate then undergoes decarboxylation to form the corresponding hydroxyacyl-thiamine conjugate. (2) The hydroxyacyl-thiamine executes a nucleophilic attack on the S1-sulphur of an oxidised lipoic acid bound to the LD, releasing

α-TPP. (3) The LD is then translocated to the inner-core domain of E_2, which catalyses the transacylation of the decarboxylated ketone from the LD to the thiol of CoA. The acyl-CoA is subsequently released from the enzyme complex, whereafter the LD dissociates from E_2. (4) Finally, the LD migrates to E_3, catalysing the oxidative regeneration of lipoic acid with the FAD coenzyme, which is in turn re-oxidised by NAD.

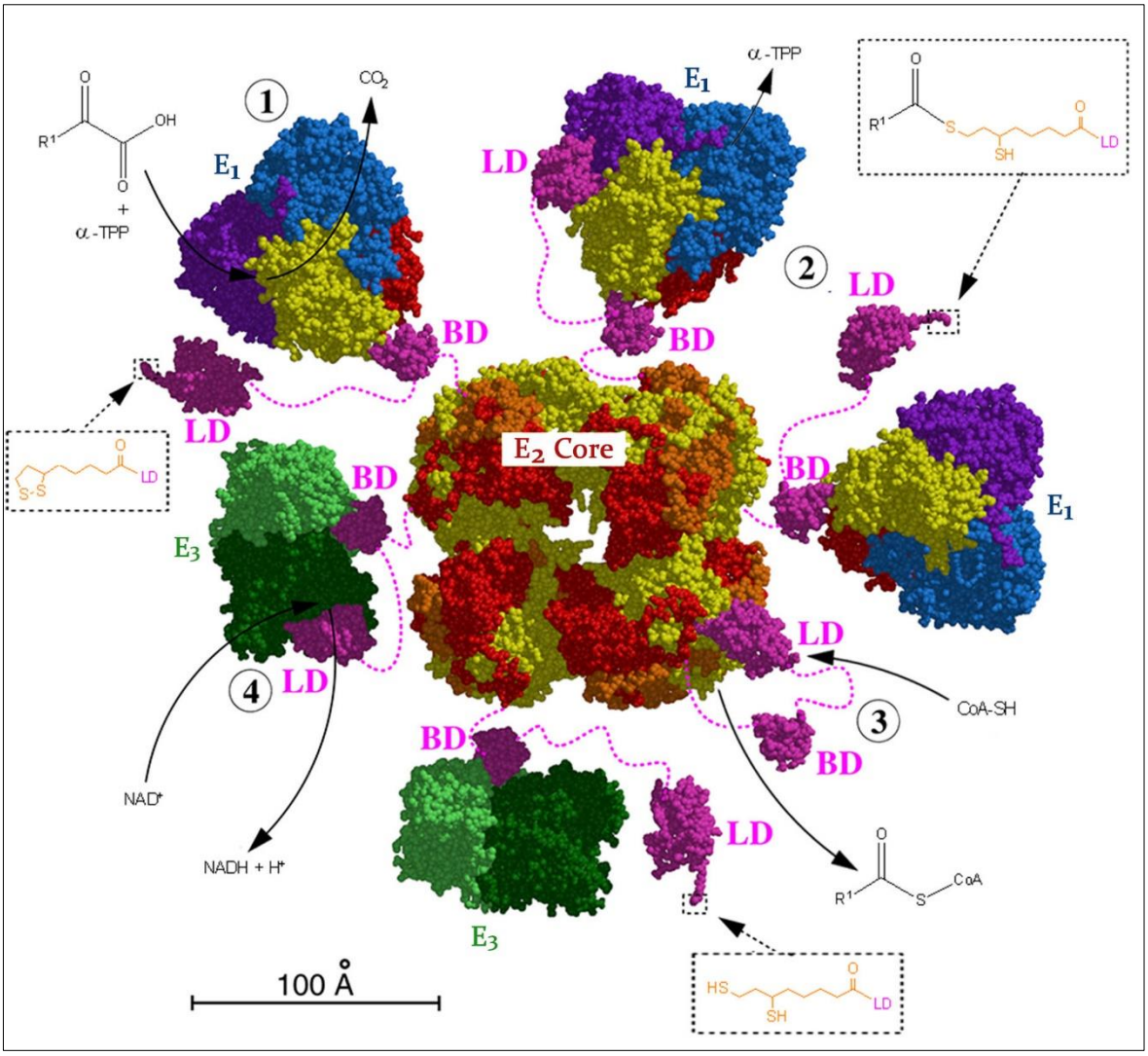

Figure 3: Components and dynamics of branched-chain α-keto acid dehydrogenase enzyme complex (BCKDH) Ævarsson *et al.* (1999:789), modified by permission of Oxford University Press.

The image above is a simplified illustration of the structure, components and functional cycle of BCKDH. A more comprehensive review and crystal structure was published by Ævarsson et al. (1999:785-792) and Ævarsson et al. (2000:277-291). R^1 correlates to either an isopropyl, 2-butyl or isobutyl side chain representing 2-keto-3-methylvaleric acid, 2-ketoisovaleric acid and 2-ketoisocaproic acid, respectively. E_1, E_2 Core and E_3 respectively denote the three catalytic subunits as described above. The purple dotted lines indicate the highly mobile and flexible inter-domain segments, hosting the BD and LD. The encircled numbers 1–4 correlate to the reaction mechanism described in the text above (numbers in brackets) in conjunction with the reaction mechanism shown in Figure 4. BD: binding domain for E_1/E_3; LD: lipoic acid binding domain; NAD: nicotinamide adenine dinucleotide; NADH + H^+: oxidised NAD; α-TPP: thiamine pyrophosphate.

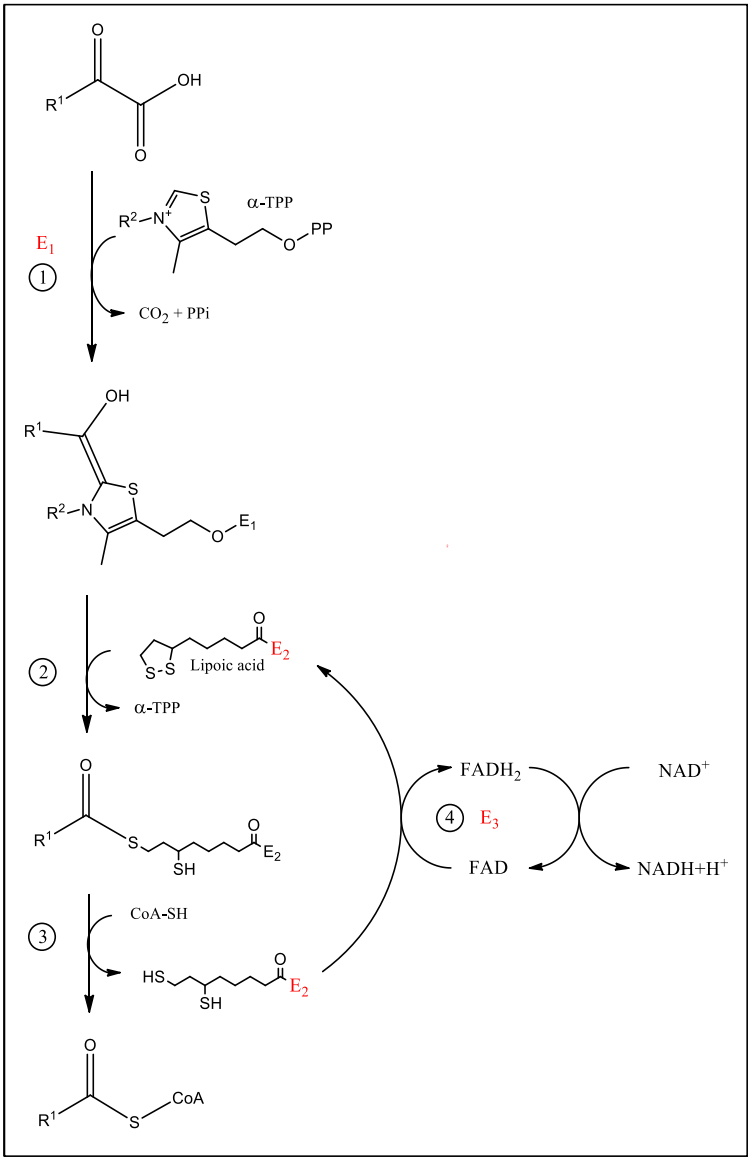

Figure 4: Simplified illustration of the reaction mechanism of branched-chain α-keto acid dehydrogenase enzyme complex (BCKDH) adapted from Deysel (2001:11) and Dry (1997:17).

In the illustration above, R^1 correlates to an isopropyl, 2-butyl or isobutyl side chain representing 2-keto-3-methylvaleric acid, 2-ketoisovaleric acid and 2-ketoisocaproic acid, respectively. R^2 represents the rest of the thiamine molecule not shown. E_1, E_2 and E_3 respectively denote the three catalytic subunits of BCKDH as described above. The encircled numbers 1–4 correlate to the reaction mechanism described in the text above (numbers in brackets) in conjunction with the components and dynamics illustrated in Figure 3. PPi: pyrophosphate; NAD: nicotinamide adenine dinucleotide; NADH + H^+: oxidised NAD; FAD: flavin adenine dinucleotide; $FADH_2$: oxidised FAD; α-TPP: thiamine pyrophosphate; CoA: coenzyme A ester; CoASH: coenzyme A (free unbounded CoA in reduced form).

2.2.2 Maple syrup urine disease

MSUD (OMIM: 248600) is an autosomal recessive disorder of BCAA metabolism due to a defective BCKDH, which was first described by Menkes *et al.* (1954:162-167). This deficiency is caused by a wide range of different possible mutations on any of the following genetic locations: 19q13.2 ($E_1\alpha$ defect) (OMIM: 608348), 6q14.1 ($E_1\beta$ defect) (OMIM: 248611), 1p21.2 (E_2 defect) (OMIM: 248610) or 7q31.1 (E_3 defect) (OMIM: 238331). Since E_3 is also a component of two other 2-keto acid dehydrogenase complexes (pyruvate and 2-ketoglutarate dehydrogenase complexes), a defect of this subunit is classified as E_3 deficiency instead of an MSUD sub-type. The various genotypes of the disorder lead to diverse phenotypes, from a mild thiamine responsive form to the more severe classical MSUD. Elevated levels of the BCAAs and especially *allo*-Ile (aIle), in all biofluids, are typical of MSUD (Dancis *et al.*, 1959:91; Snyderman *et al.*, 1964:454). In addition, the respective 2-keto acids (the substrates of BCKDH), 2-keto-3-methylvaleric acid, 2-ketoisovaleric acid and 2-ketoisocaproic acid, as well as their 2-hydroxy acid analogues, are elevated (Dancis *et al.*, 1959:91; Meister & Abendschein, 1956:171). Ammonia, N-Ac-BCAA and N-lactyl-BCAA conjugates and the conjugation of 2-hydroxyisovaleric acid with BCAAs have also been described (Deysel, 2001:3-18; Dry, 2000:10; Guder, 1998:42; Hagenfeldt & Naglo, 1987:77).

Dry (1997) conducted a study on the induced metabolism of MSUD and described several metabolites, some of which had not been described in the literature at the time, (summarised in Annexure A Table 15 together with known metabolites). These included trace levels of (1) the decarboxylated amine, isobutylamine, (2) 2-hydroxy-BCAA conjugates, N-2-hydroxy-3-methylvaleryl-BCAA and 2-hydroxy-isocaproyl-Val, (3) N-lactyl-Gly, (4) N-Ac-Gly and (5) 5-hydroxy-5-isopropyl-hydantoin, 5-hydroxy-5-methylpropyl-hydantoin and 5-hydroxy-5-isobutyl-hydantoin. According to Dry (1997:64-82), some compounds not listed above were presumably present in concentrations too low to be determined, e.g. the decarboxylated amines of 2-keto-3-methylvaleric acid and 2-keto-isocaproic acid as well as the 2-hydroxy-BCAA conjugates 2-hydroxy-isocaproyl-Ile and 2-hydroxy-isocaproyl-Val. In addition to the above-mentioned biochemical image, Dry (1997:28-31, 87-95) also described three cases of MSUD that presented with the N-AAA conjugates N-isobutyryl, N-isovaleryl and N-2-methylbutyryl conjugates of the BCAAs, which were later confirmed by Deysel (2001) and Mienie *et al*. (2001:42-44). The formation of the hydantoin and N-AAA conjugates has hitherto been unclear. This dissertation mainly focuses on the observed elevated levels of N-AAA conjugates of the BCAA.

Since acyl-CoAs are required for the formation of N-AAA conjugates (Dry, 2000:68) and the formation of the related acyl-CoAs in MSUD is improbable, a review of N-AAA conjugate metabolism may help to understand hypotheses about these findings.

2.3 Metabolic significance of N-AAA conjugates

N-AAA conjugates are intermediates of primary and alternative metabolism (as in the case of IEMs) and are implicated in the biotransformation of acyl-CoAs with Gly (Liska *et al*., 2006:126-133), degradation of both endo- and exogenous N-acetylated proteins (Perrier *et al.*, 2005:674), regulation of the urea cycle (Caldovic & Tuchman, 2003:281-285) and in various IEMs as indicated in Table 1 and described in the sections below. Most notable is the sub-group N-AcAA conjugates (refer to section 1.1 for definitions and clarification). This section (2.3) discusses the catabolism of N-AAA conjugates, focusing mainly on ACY1 and N-acylated protein degradation, followed by the anabolism of N-AAA conjugates and ending with a review of other hypothesised sources of N-AAA conjugates.

2.3.1 Catabolism of N-AAA conjugates by aminoacylases

Aminoacylases are cytosolic hydrolase enzymes, containing Zn(II) ions and having diverse substrate specificities. These enzymes act on the carbon-nitrogen bonds of linear amines in non-peptide bonds, catalysing the hydrolysis of various N-AAA conjugates, producing the corresponding amino and organic acids (Perrier *et al*., 2005:677 679), e.g. the hydrolysis of N-Ac-Gly produces the organic acid acetate and the amino acid Gly. The three main human aminoacylases that have been described, namely ACY1 (EC: 3.5.1.14), aminoacylase II (ACY2) (EC: 3.5.1.15) and aminoacylase III (ACY3) (EC: 3.5.1.114) (BRENDA, 2016a; BRENDA, 2016b; BRENDA, 2016c), are discussed below.

2.3.1.1 Aminoacylase I

ACY1, also known as N-acyl-aliphatic-L-amino acid amidohydrolase, is a dimeric enzyme comprised of two identical 45-kDa subunits, each containing a single Zn(II) ion. Substrate binding causes a conformational shift, made possible by the Zn(II) ion, bringing the catalytic site situated between the monomers around the substrate, allowing catalysis to occur (Perrier *et al*., 2005:679-681). ACY1 can also catalyse the reverse reaction, synthesising N-AAA conjugates under the correct cellular circumstances (Sass *et al.*, 2006:401). ACY1 has the highest activity in the kidneys, where it is essential for salvaging amino acids, especially during starvation and increased protein catabolism (Sass *et al*., 2006:406-407). This enzyme is also found in erythrocytes and lung, brain, heart and liver cells (Perrier *et al*., 2005:679-681). ACY1 is

overexpressed in numerous tumour cell types, especially in neoplastic epithelial cells in patients with colorectal cancer (BRENDA, 2016a).

As discussed in section 2.3.1.3, ACY1 is not only involved in the catabolism of N-AcAA conjugates but also in the hydrolysis of various other aliphatic acyl groups conjugated to amino acids (BRENDA, 2016a; Perrier *et al*., 2005:677; Van Coster *et al.*, 2005:1322). In humans, these include short and simple aliphatic acyl groups such as formyl, propionyl, 2,2-dimethylpropionyl, butyryl, isobutyryl and isovaleryl groups (the latter two having implications for this study). Longer and more complicated acyl groups have been described for other organisms, e.g. palmitoyl (*Mycobacterium* sp.) and 3-(2-furyl)acrylyl groups (*Sus scrofa*) (BRENDA, 2016a). Just as ACY1 accepts a diverse set of acyl groups conjugated to amino acids, it also accepts a diverse group of amino acids, albeit limited to amino acids conjugated to acyl groups in the L-enantiomer configuration. ACY1 has a preferred substrate specificity for all the aliphatic and neutral L-amino acids typically found in proteins (Ala, asparagine [Asn], Cys, Gly, Gln, methionine, Ser, threonine, Leu and Val [the latter two having implications for this study]) and more complicated derivatives such as the mercapturic acid conjugate S-[(1*S*,2*S*)-2-hydroxycyclohexyl]Cys. The two exceptions are the acidic amino acid conjugate N-acyl-L-Glu (but not Asp) and the aliphatic amino acid conjugate N-acyl-L-Ile, that are substrates for the enzyme as well (BRENDA, 2016a; Perrier *et al*., 2005:677; Van Coster *et al.*, 2005:1322). It is worth noting that even though the 2-methylbutyryl group (as an acyl group in an N-AAA conjugate) and Ile (as a conjugated amino acid in an N-AAA conjugate) are structurally analogous to other above-mentioned substrates, they have not been described in the literature as substrates for ACY1. N-Ac-Ile, however, has been reported to be elevated in several cases of deficient ACY1 (Sass *et al.*, 2006:401; Van Coster *et al*. 2005:1322-1326). Both N-acyl-Ile- and N-2-methylbutyryl-amino acid conjugates have implications for this study.

Refer to Figure 5: ACY1 is also integral to the metabolism of certain xenobiotics such as paracetamol and N-Ac-cysteine. N-Ac-cysteine (traded as ACC) has a mucolytic effect and is indicated for acetaminophen (traded as paracetamol) overdose (MIA, 2000). Acetaminophen readily undergoes glucuronidation and sulphation, which produce water-soluble metabolites that are excreted in the urine. Approximately 5% is oxidised to the extremely toxic metabolite N-Ac-*p*-benzoquinone imine (Mazaleuskaya *et al.*, 2015:416-417). N-Ac-*p*-benzoquinone imine conjugates with glutathione to produce the less reactive metabolite acetaminophen-glutathione. During an acetaminophen overdose, however, glutathione is quickly depleted. Due to its low abundance, Cys is the limiting factor in glutathione production. Orally administered N-Ac-Cys is readily absorbed and metabolised by ACY1 to liberate Cys for glutathione production. Thus, ACY1 also deactivates the pharmacological action (mucolytic effect) of N-Ac-Cys (Heard,

2009:285-286; Lautenberg *et al.*, 1983:980-981; Mazaleuskaya *et al.*, 2015:416-417; Perrier *et al.*, 2005:678; Uttamsingh *et al.*, 2000:625).

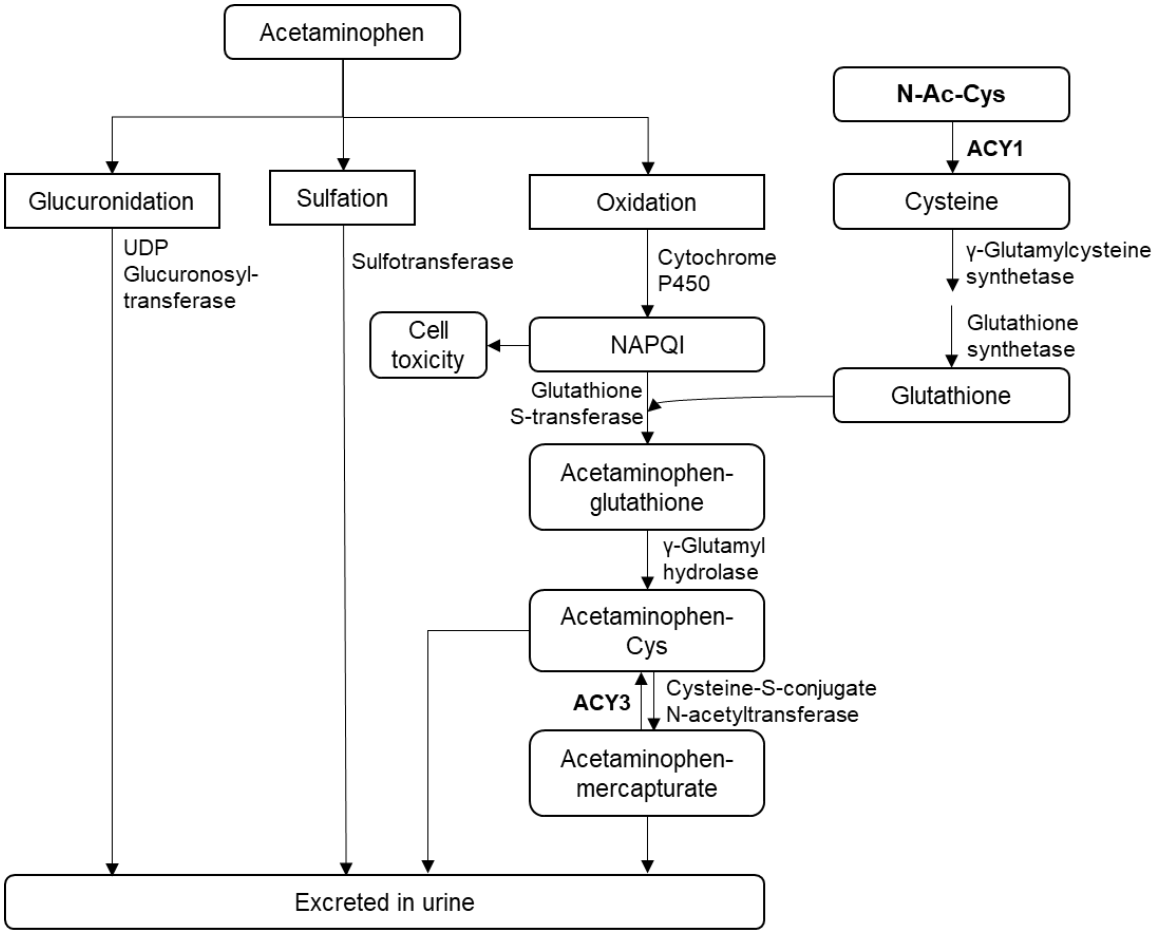

Figure 5: N-AAA conjugates and ACY1's involvement in xenobiotics. Adapted from Heard (2009:285-286) and Mazaleuskaya *et al.* (2015:416-417).

*ACY1: aminoacylase I; ACY3: aminoacylase III; NAPQI: N-Ac-*p*-benzoquinone imine; UDP: Uridine 5'-diphosphate.*

In 2005, Van Coster *et al*. (2005:1322-1326) described the first isolated case with ACY1 deficiency (OMIM: 609924). This deficiency was characterised by a marked increase in urinary excretion of N-AcAA conjugates, hypotonia, acute encephalopathy and seizures. The excreted conjugates match the substrate specificity of ACY1 as described above, with the addition of small amounts of N-Ac-L-Ile. No basic or aromatic amino acid conjugates were detected. Several cases have since been reported with a heterogeneous clinical presentation, which include the symptoms mentioned above, with a primary neurological phenotype and resembling a biotinidase deficiency (Kniffin, 2014; Sass *et al.*, 2006:401-409).

2.3.1.2 Aminoacylase II and III

ACY2, also known as aspartoacylase, is an encephalic enzyme. N-Ac-Asp is produced in a neuron's mitochondria by Asp N-Ac-transferase, whereafter it is transported to the oligodendrocytes. ACY2 catalyses the hydrolysis of N-Ac-L-Asp to supply the brain with acetate for myelin synthesis (Wijayasinghe *et al.*, 2014:4970). This enzyme has been shown to be deficient in Canavan disease (OMIM: 271900), an autosomal recessive degenerative disorder in which the defective ACY2 results in harmful accumulation of N-Ac-L-Asp, which is measurable in urine (BRENDA, 2016b; Hagenfeldt *et al.*, 1967:135). This results in dysmyelination and spongiform degeneration of white matter. Clinically, children with Canavan disease present with psychomotor deterioration, hypotonia, macrocephaly, spasticity, poor vision and a short life expectancy (Sass *et al.*, 2006:401-402; Wijayasinghe *et al.*, 2014:4970).

ACY3, also known as N-acyl-aromatic-L-amino acid amidohydrolase, catalyses the hydrolysis of N-acetylated aromatic amino acids and mercapturates (N-Ac-Cys-S-conjugates), which are not usually substrates for ACY1. ACY3 also does not catalyse N-acyl-L-aspartic acid, which is a substrate only for ACY2 (BRENDA, 2016c). Refer to Figure 6: No deficiency for ACY3 has been described, but ACY3 is implicated in the neural toxicity of 4-hydroxy-2-nonenal (4HNE) and acrolein, which might contribute to the pathogenesis and progression of both Alzheimer's and Parkinson's diseases. Both 4HNE and acrolein are conjugated to glutathione by glutathione-S-transferase, which initiates the biotransformation pathway for these compounds. The glutathione conjugates are subsequently converted to their respective Cys conjugates by gamma-glutamyl hydrolase and membrane alanyl aminopeptidase. Cys-S-conjugate N-Ac-transferase then catalyses the formation of the respective mercapturates which are subsequently excreted in the urine. ACY3 catalyses the reverse reaction of the latter, contributing to the substrate pool for beta-lyase and flavin monooxygenase, which form highly toxic compounds (LoPachin *et al.*, 2009:1499; Tsirulnikov *et al.*, 2012:303-4).

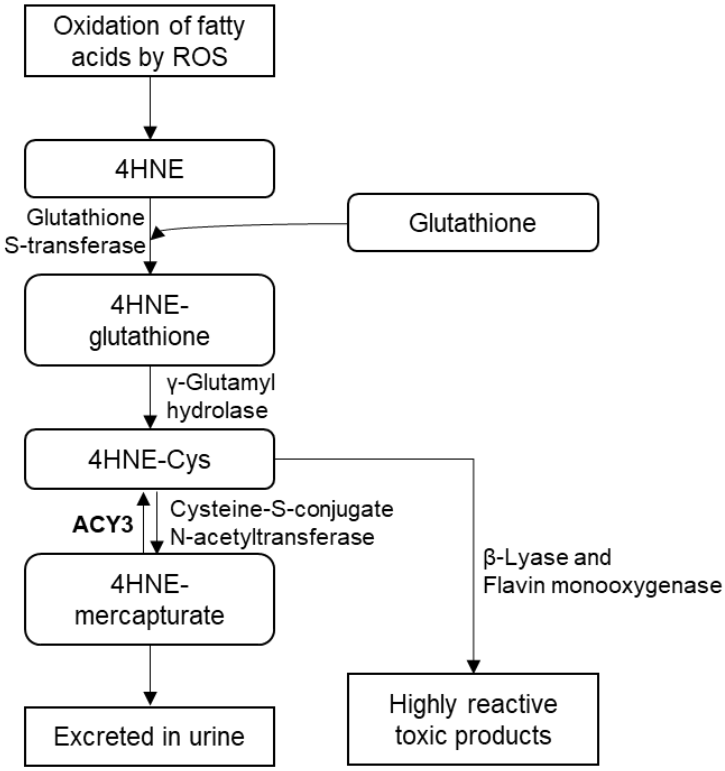

Figure 6: The role of ACY3 in the biotransformation of 4HNE. Adapted from LoPachin *et al*. (2009:1499) and Tsirulnikov *et al.* (2012:303-4).

Acrolein (not shown) follows the same reaction path as 4-hydroxy-2-nonenal (4HNE). ACY3: aminoacylase III; ROS: reactive oxygen species.

2.3.1.3 N-Acetylated proteins as a source of N-AcAA conjugates

The N-termini of most cellular proteins, especially structural proteins, are covalently acetylated. This process can take place during or after translation and is catalysed by a diverse set of peptide alpha-N-Ac-transferases with Ac-CoA as an Ac donor (Perrier *et al.*, 2005:674-677). This process is usually preceded by cleavage of the N-terminal methionine residue by methionine-aminopeptidases. N-terminus acetylation protects a protein from premature degradation and assists in localisation, regulation, stability and interactions with other proteins. Stability improvements include the prevention of unwanted non-enzymatic reactions with, amongst others, reducing sugars and cyanate ions (Perrier *et al.*, 2005:674; Van Damme *et al.*, 2011:2-3). Although the degradation process of N-acetylated proteins is still partly unclear, it is proposed by Perrier *et al.* (2005:674-677) to occur via the processes described in Figure 7 below.

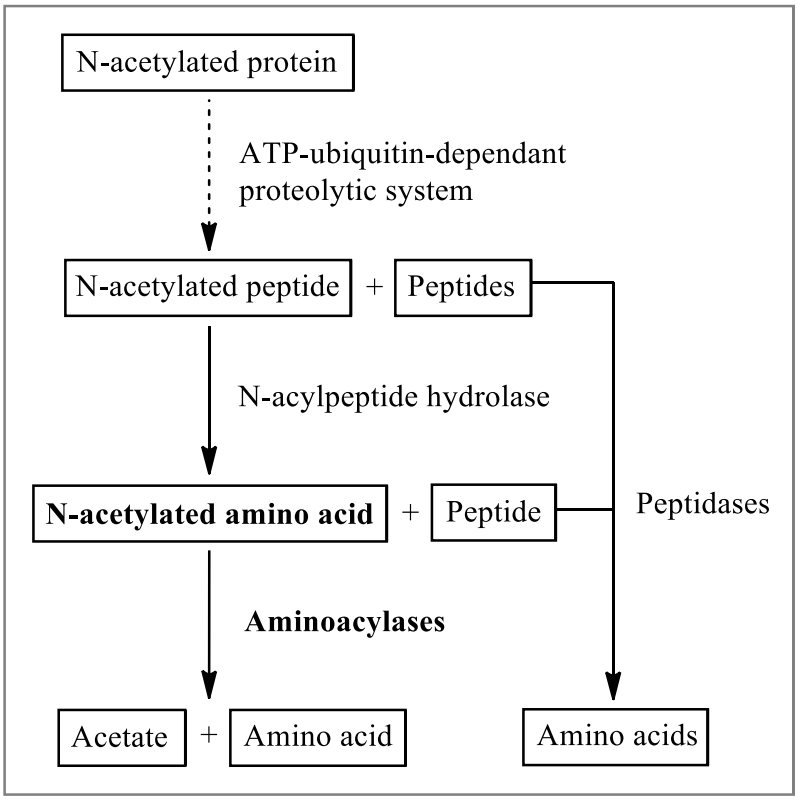

Figure 7: Overview of N-acetylated protein degradation, adapted from Perrier *et al.* (2005:674-677).

ATP: adenosine triphosphate.

According to Perrier *et al.* (2005:674-677), the first step of N-acylated protein degradation appears to be via ATP-ubiquitin-dependant proteasome system and not via lysosomal cathepsins or calcium-activated calpains as previously postulated. In this first step, an N-acetylated protein marked for degradation is hydrolysed into an N-acetylated peptide (attained from the N-terminal of the protein) and other small peptides. The N-acetylated peptide is then hydrolysed by N-acylpeptide hydrolase to liberate the N-AcAA on the N-terminus. The resulting peptide (and other peptides from the previous step) is further degraded by various peptidases, while the N-AcAA conjugate is hydrolysed by aminoacylases into a free amino acid and acetate (Perrier *et al*., 2005:673-377). N-AcAA conjugates are thus intermediates in the degradation pathway of N-acetylated proteins and the liberated amino acids are essential for synthesis of new proteins. The mechanism by which N-acetylated proteins in food are digested has not yet been fully elucidated, but it is clear that dietary N-acetylated proteins are a source of N-AcAA conjugates, albeit short-lived, because of the hydrolytic activity of aminoacylases in the epithelial cells of the gut (Perrier *et al*., 2005:677). Elevated levels of N-Ac-tyrosine have also been reported in patients receiving certain parenteral solutions (Gerlo *et al.*, 2006:191).

2.3.2 Anabolism of N-AAA conjugates

Various N-Ac-transferases (EC: 2.3.1.x), which catalyse the acyl group transfer for the formation of N-AAA conjugates, are summarised in Table 2. This also illustrates the diversity of N-AcAA conjugates in the metabolism. The most notable is GLYAT (as discussed in section 2.1), peptide alpha-N-Ac-transferase (EC: 2.3.1.88) – the latter acetylate proteins as discussed in the section above – and N-acyl-Glu in the urea cycle or arginine (Arg) biosynthesis pathway (refer to Figure 8). N-Ac-Glu is an allosteric activator of the enzyme carbamoyl phosphate synthase I, which is essential for the urea cycle function (Caldovic & Tuchman, 2003:281-285; KEGG, 2016; Kniffin, 2007; Salway 2004:74-79).

Table 2: Various N-acyltransferases involved in N-AAA conjugates' synthesis.

Enzyme	EC number	Substrates	Main product	Pathway(s)	References
N-Acetylglutamate synthase	2.3.1.1	Ac-CoA + L-Glu/L-Asp	N-Ac-Glu	Arginine biosynthesis, urea cycle	Caldovic and Tuchman (2003:281)
Cysteine-S-conjugate N-Ac-transferase	2.3.1.80	Ac-CoA + L-Cys-S-conjugate	N-Ac-L-Cys-S-conjugate/mercapturate	Glutathione-mediated biotransformation	LoPachin *et al.* (2009:1499); Tsirulnikov *et al.* (2012:303-4)
Aspartate-N-Ac-transferase	2.3.1.17	Ac-CoA + L-Asp	N-Ac-Asp	Ala, Asp and Glu metabolism Myelin biosynthesis	Wijayasinghe *et al.* (2014:4970)
Peptide alpha-N-Ac-transferase	2.3.1.88	Ac-CoA + peptide/protein	N-α-Acetylated peptide/protein	Protein/peptide co-/post-translational modifications	Perrier *et al*. (2005:679-681)
Glycine N-acyltransferase	2.3.1.13	Various acyl-CoAs (except phenylacetyl- and (indol-3-yl)Ac-CoA) + mainly Gly (also: L-Ala, L-Asp, L-Glu and L-Gln)	Various N-AAA conjugates, mostly N-Acyl-Gly	Gly-mediated biotransformation	Liska *et al*. (2006:126-133); Van der Westhuizen *et al*. (2000:102-103)
Glutamine N-acyltransferase	2.3.1.68	Indoleacetyl-/phenylacetyl-CoA + L-Gln	N-Indole-/phenylacetyl-Gln	Biotransformation of indole- and phenylacetate	Webster *et al.* (1975:3352)

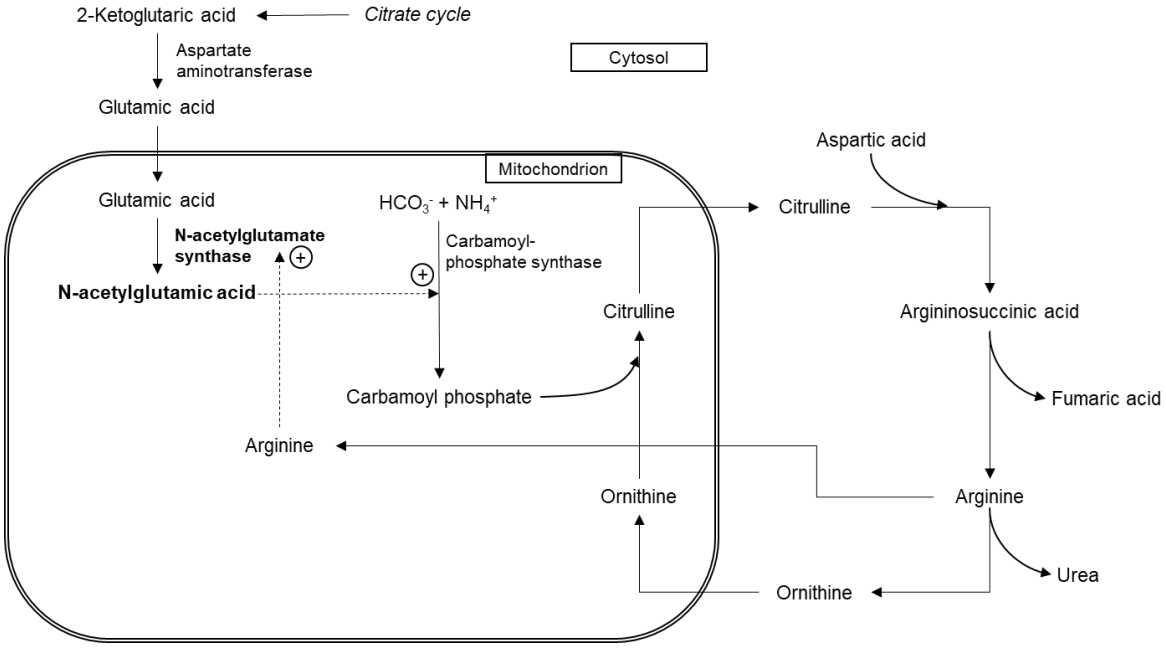

Figure 8: N-Acetylglutamate and the urea cycle. Adapted from Caldovic and Tuchman (2003:281), KEGG (2016) and Salway (2004:74-79).

The figure above is a simplified illustration of N-acyl-Glu's involvement in the urea cycle. N-Acetylglutamate is an allosteric activator of carbamoyl-phosphate synthase (indicated by the encircled positive sign), which in turn initialises the urea cycle. Arg also activates N-acetylglutamate synthase. For simplicity, the enzymes of the urea cycle itself are not shown.

As noted, the substrate of N-Ac-transferases always consists of a CoA and an amino acid/peptide. Accumulating metabolites in MSUD consist of keto acids and BCAAs and not of CoAs, yet N-AAA conjugates are abundant in specimens of patients with MSUD (Hagenfeldt & Naglo, 1987:77; Lehnert & Werle, 1988:123). While N-AAA conjugates appear not to be derived from acyl-CoAs in MSUD, the ample sources of these types of conjugates in the broad metabolism suggests other possible routes to the formation of the N-AAA conjugates (observed in the variant MSUD cases) that may or may not be related to MSUD itself.

2.3.3 Previously proposed mechanisms for the formation of N-acylated-BCAA conjugates in MSUD

Motivated by the unexpected observation of N-AAA conjugates of BCAAs and other compounds in the variant MSUD cases mentioned in section 2.1, Dry (1997:95-108) and Mienie (2001:51-52) attempted to elucidate various hypotheses as to the origin of these conjugates. The first hypothesis was that a branched-chain 2-keto acid could be decarboxylated by an oxidase enzyme, converting it to its respective aldehyde, whereafter the non-enzymatic

conjugation with BCAAs was hypothesised. However, no aldehyde intermediates could be detected (Dry 1997:95-96). Dry's (1997:97-180) second hypothesis was that if a BCKDH-E_2 deficiency was present, wherein the already decarboxylated hydroxyacyl-thiamine enzyme intermediates accumulated (catalysed by the non-deficient E_1 subunit), these intermediates could react with BCAAs to produce the observed N-AAA conjugates. Dry (1997:97-180) could establish the presence of the corresponding hydroxyacyl-thiamine conjugates and, to a certain extent, the probability of a BCKDH-E_2 deficiency in the variant MSUD cases, but no mechanism for conjugation with BCAAs was described.

Yoshioka and Uematsu (1993:783-790) described a mechanism for the formation of N-hydroxy-N-arylacetamides from 2-keto acids and nitroso aromatic compounds as illustrated in Figure 9. These authors showed that the E_1 subunit of the pyruvate dehydrogenase complex (PDH-E_1) (which is similar in structure and function to BCKDH-E_1), autonomously catalyses this reaction via the reductive acylation of a nitroso aromatic compound, instead of transferring the hydroxyacyl group to lipoic acid on the LD. Mienie *et al.* (2001:51-52) hypothesised that the BCAAs could replace the nitroso aromatic compounds in Yoshioka and Uematsu's mechanism forming the corresponding N-AAA conjugates. Mienie *et al*. proposed a new mechanism, in which the accumulating decarboxylated acyl-thiamine enzyme intermediates in a BCKDH-E_2 deficiency, described above, react with BCAAs to form the corresponding N-AAA conjugates as shown in Figure 9.

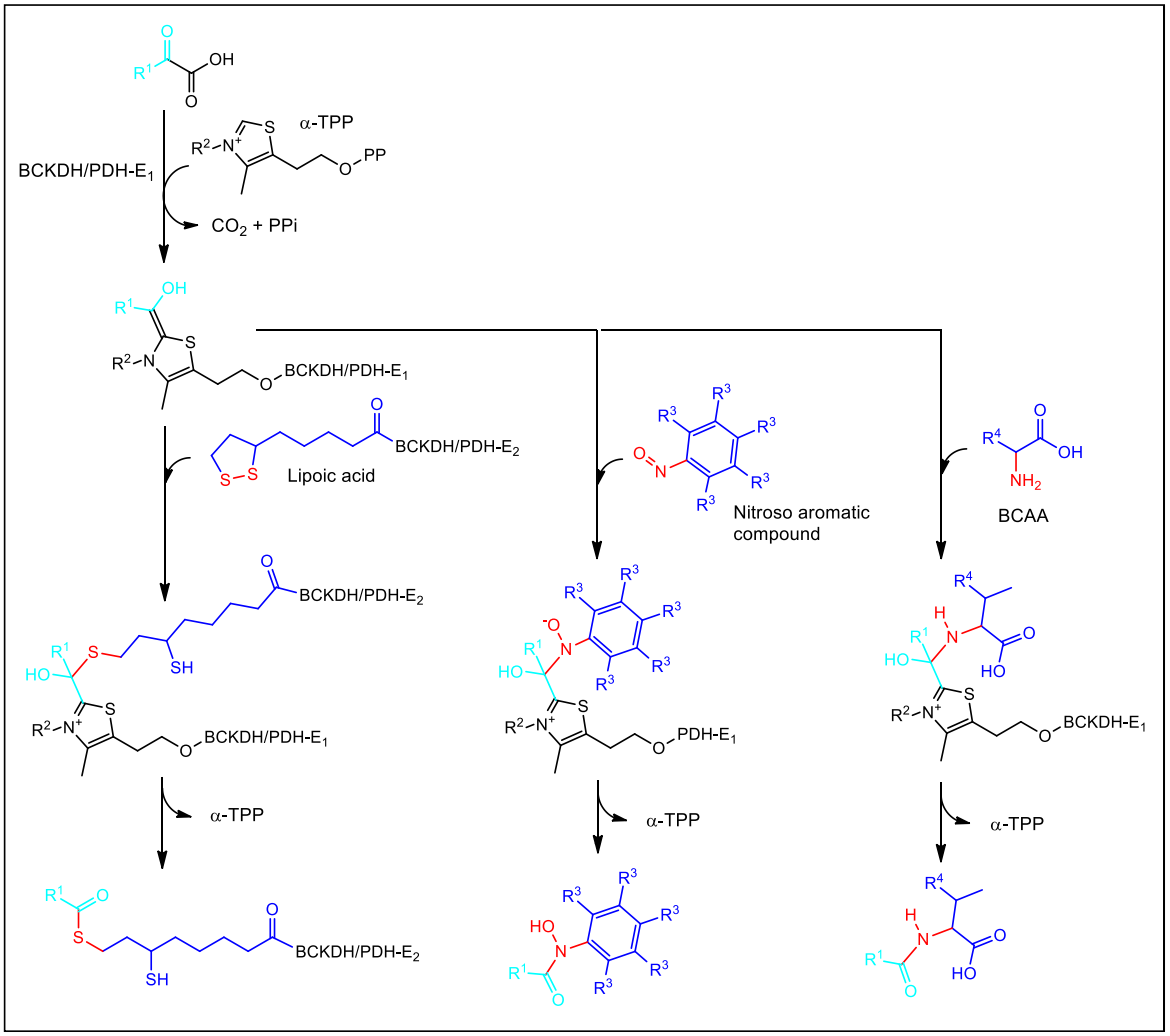

Figure 9: Alternative metabolism of 2-keto acids, adapted from Mienie *et al*. (2001:63), Dry (1997:17) and Yoshioka and Uematsu (1993:789).

In the illustration above, R^1 correlates to either a methyl, isopropyl, 2-butyl or isobutyl side chain representing pyruvate, 2-keto-3-methylvaleric acid, 2-ketoisovaleric acid and 2-ketoisocaproic acid, respectively. R^2 represents the rest of the thiamine molecule not shown. R^3 represents hydrogen or any other organic substituent and R^4 represents the same side chains as R^1 except for the methyl side chain. The primary pathway on the left shows the first similar steps of branched-chain α-keto acid dehydrogenase complex (BCKDH) and PDH as described above. The alternative pathway in the middle shows the formation of N-hydroxy-N-arylacetamides as described by Yoshioka and Uematsu (1993:789) and the pathway on the right shows the formation of N-AAA conjugates as hypothesised by Mienie et al. (2001:63). α-TPP: thiamine pyrophosphate; BCAA: branched-chain amino acid; E_1: α-TPP dependent 2-keto acid dehydrogenase; PDH: pyruvate dehydrogenase; PP: pyrophosphate.

Mienie *et al.* (2001:51-52) attempted to test this hypothesis by replicating the results of Yoshioka and Uematsu (1993:783-790), but with unspecified porcine 2-keto acid dehydrogenase (free of CoAs), 2-ketoisovaleric acid as the 2-keto acid substrate and the BCAA Val instead of a nitroso aromatic compound. Large amounts of N-isobutyryl-Val were detected in both the enzyme preparation and, unexpectedly, in the control containing no Val. A systematic elimination of reagents (and ultimately the enzyme) showed the spontaneous formation of N-isobutyryl-Val from 2-ketoisovaleric acid and ammonium sulphate (which was used to purify the enzyme). Thereafter, Mienie *et al.* (2001:52-53) incubated pyruvate and 2-ketoisovaleric acid with only ammonium sulphate, which led to the spontaneous formation of N-Ac-Ala, N-Ac-Val, N-isobutyryl-Ala and N-isobutyryl-Val. As mentioned earlier, 2-keto acids and ammonia levels are elevated in all cases of MSUD but were lowered in the three South African variants. Mienie *et al.* (2001:53) proposed that the lower levels of 2-keto acids and ammonia are due to the formation of the observed N-AAA conjugates (only observed in the three South African variants) from 2-keto acids and ammonia, depleting the indicated substrates.

Refer to Figure 10: Yanagawa *et al.* (1982) described the reaction of carbonyl compounds with ammonia in an aqueous medium to demonstrate the probable formation of amino acids in the primeval sea. These authors demonstrated that pyruvate and glyoxylic acid react with ammonium sulphate at slightly above room temperature and low pH to yield a conjugate that would produce amino acids upon hydrolysis with hydrochloric acid (HCl). With this method Yanagawa *et al.* synthesised N-oxalyl-Gly, N-oxalylsarcosine and N-Ac-Ala. Mienie *et al.* (2001:52-53) proposed the non-enzymatic formation of N-AAA conjugates from 2-keto acids and ammonia via the mechanism proposed by Yanagawa *et al.* (1982:2089).

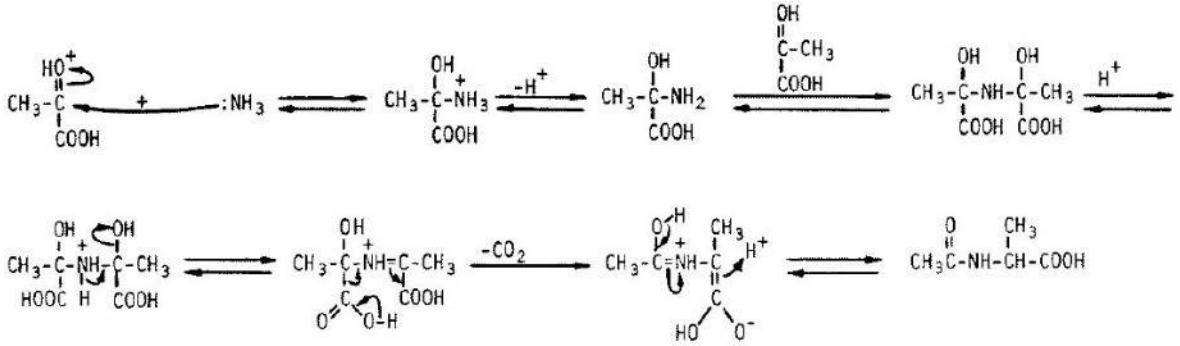

Figure 10: Reaction mechanism proposed by Yanagawa *et al*. (1982:2089) for the non-enzymatic formation of N-Ac-Ala from pyruvic acid and ammonia in an aqueous medium. Reprinted by permission of Oxford University Press.

A quotation from a more recent review by Farrell and Merkler (2008:562) states the following: 'Clearly, there is much work remaining to better define the pathways of biosynthesis and degradation for the N-acylamino acids.' In conclusion, the latter and previous work by Deysel

(2001), Dry (1997) and Mienie *et al*. (2001) emphasise the importance of providing alternative mechanisms for the synthesis of N-AAA conjugates. In order to further study the proposed mechanism for the non-enzymatic formation of N-AAA conjugates, a look into how metabolic pathways are elucidated and the role of non-enzymatic reactions in metabolism may shed some light on the subject. This may also contribute to the prospective role of non-enzymatic reactions in metabolism, health and disease.

2.4 Metabolic pathways

2.4.1 A brief history

'Both the body and its parts are in a continuous state of dissolution and nourishment, so they are inevitably undergoing permanent change' (al-Nafis cited by Wietmarschen, 2012:52). This statement made by the Arab physician Ibn al-Nafis in a theological novel (1268–1277 AD) is the earliest reference to metabolism. However, the study of metabolism and its affects dates back to the ancient Chinese, who discovered the relation between sweet-tasting urine and a specific disease – now known as diabetes (Wietmarschen, 2012:52). The first controlled metabolic experiments were conducted by the Venetian physician and inventor Santorio Santorii, who introduced the qualitative approach to the field. In 1614, Santorii described a method to measure what he called 'insensible perspiration', a prelude to the study of basal metabolic rate (Wietmarschen, 2012:52). In 1824 and again in 1828, Friedrich Wöhler partially disproved vitalism by synthesising oxalic acid and then later urea from inorganic compounds. This also showed that the principles of chemistry applied to the organic compounds in living cells, formerly believed to be driven by an unquantifiable vital force (Kinne-Saffran & Kinne, 1999:290).

In 1858, Louis Pasteur demonstrated that the chemical reaction of fermentation is catalysed by living yeast cells. However, in 1897 it was finally proven by Eduard Buchner that the fermentation process was catalysed by enzymes that could be extracted from living cells, marking the beginning of biochemistry and starting the study into the elucidation of metabolic pathways (Kinne-Saffran & Kinne, 1999:290-291). Building on the work done earlier by Albert Szent-Györgyi, the most significant contribution to the elucidation of metabolic pathways was the discovery of the citric acid cycle by Hans Krebs in 1937. Krebs also resolved the urea and glyoxylate cycles. Another vital step in the elucidation of metabolic pathways was the discovery of metabolic channelling by Paul Srere in 1987. Srere showed that sequential enzymes that were otherwise free in dilute solutions, co-precipitated (Kinne-Saffran & Kinne, 1999:291).

Armed with this knowledge and the rapidly evolving fields of analytical chemistry, molecular chemistry and genetics as well as profiling techniques, over 40 000 metabolites have been discovered to date. Of these only approximately 5 000 associated biochemical reactions have been described and about 400 pathways have been mapped (HMDB, 2016). The elucidation of metabolic pathways is thus an ongoing and continuously developing discipline.

2.4.2 Elucidation of metabolic pathways

Studies of the origin, formation and clinical effects of metabolite pathways leads to a better understanding of the clinical-biochemical relationship of a disorder to ultimately improve, develop and help monitor current and novel treatment strategies and establish new biomarkers for the detection of IEMs. According to Rogers (2011:66-68), clues to the possible elucidation of a pathway can be acquired in the following ways:

- Studying an organism under stress or disease to measure the accumulation of certain metabolites (e.g. lactic acid during exercise or pyruvate accumulation due to thiamine deficiency).

- Administering a metabolic toxin to inhibit or reduce the activity of an enzyme to measure the accumulation of the enzyme's substrate.

- Administering intermediates labelled with isotopes. This technique has been used since the 1930s and demonstrates the dynamic state of the body's constituents. It was also used to prove Srere's metabolic channelling concept by showing that an addition of a 200-fold excess of a labelled intermediate does not dilute the unlabelled end product of a metabolic channel (Salway, 2004:76).

- Studying organisms with genetic defects (including IEMs) to measure the accumulating substrates of a defective enzyme, similar to the administration of metabolic toxins. In addition to the single defective enzyme's reaction, the entire organism's metabolism may be affected, resulting in accumulating intermediates, alternative metabolites, metabolite deficiencies and clinical presentation (refer to the next section).

In addition to Rogers' list, molecular modelling (discussed in section 2.6) can be used to predict and/or describe enzyme structures and catalytic site activity and biochemical reactions amongst others, the latter of which is also illustrated in this study.

2.4.3 The metabolic network

When studying perturbations in a biological system (such as IEMs), focusing on the metabolome can be more rewarding than other gene products such as mRNA and proteins. According to Fiehn (2001:156), an unobtrusive perturbation in the proteome (often overlooked) can result in a distinct change in the metabolome. The complete sequencing of the human genome and other organisms has led to considerable progress in the profiling of mRNA and proteins in recent years. However, the extensive chemical diversity and difficulty in identifying novel metabolites has meant that considerably less progress has been made in metabolite profiling (Fiehn, 2001:155).

Furthermore, the chemical composition of complex biological systems can be determined by employing different analytical strategies, although the chemical diversity of the human metabolome makes it improbable for a single approach to be sufficient for all circumstances and applications. Fiehn (2001:158-162) and Wolfender *et al*. (2015:137-138) described three typical approaches as follows:

(1) Metabolite fingerprinting is an untargeted approach wherein metabolites are not extensively identified or quantified. Little or no sample clean-up is followed by limited chromatographic separation that enables higher throughput and typically reveals the entire metabolome. High-resolution analytical techniques (e.g. nuclear magnetic resonance spectrometry, tandem mass spectrometry (MS-MS) and infrared-spectrometry) are needed to reveal metabolic deviations. Sophisticated informatics tools are subsequently employed to deconvolute raw analytical data. According to Fiehn (2001:157) 'this approach should be called metabolomics'. Metabolomics is defined by Wolfender *et al.* (2015:137) as 'a non-selective, universally applicable, comprehensive analytical approach for the identification and quantification of all metabolites in a biological system'. Data from fingerprinting can be used to study the pleiotropic effect of perturbations and is used (amongst other things) for the detection of IEMs by comparing patterns or biomarkers mined on data banks of sufficient biological replicates.

(2) Metabolite profiling is a semi-targeted approach that focuses on detection and quantification of compounds in classes. Related metabolites and pathways are studied by using a specific analytical method for each class (e.g. amino and/or organic acid analysis described in sections 4.5 and 4.6). By analysing chemically related metabolites together, sample clean-up, chromatography and detection methods can be fine-tuned to reduce matrix effects and increase sensitivity and accuracy. In this approach, the relevant data from different analytical methods are combined to yield a metabolic profile. This well-

established approach is considered by Wolfender *et al.* (2015:137) to be the precursor of metabolomics.

(3) Metabolite targeted analysis is the most widely used approach and focuses on the primary effect of a perturbation. It is aimed at accurate identification and quantification of single or closely related metabolites. Although more time-consuming and expensive, targeted analysis typically has a higher sensitivity. This approach is mainly used for screening, monitoring patients and trace compound analysis.

Generally, metabolic pathways are studied via a reductionist approach, i.e. focusing on substrates, intermediates, by-products and end products in a linear flow. Although this approach makes it easier to explain and understand the pathway being studied, care should be taken not to overlook the fact that the metabolism of an organism is a complex interlinked network (refer to Figure 11). The substrates and intermediates of the primary pathway in question may also be the substrates for various other alternative pathways, depending on a cell's needs and circumstances (Chalancon *et al.*, 2013:1263-1264; Dry, 1997:4-8; Jeong *et al.*, 2000:651-652).

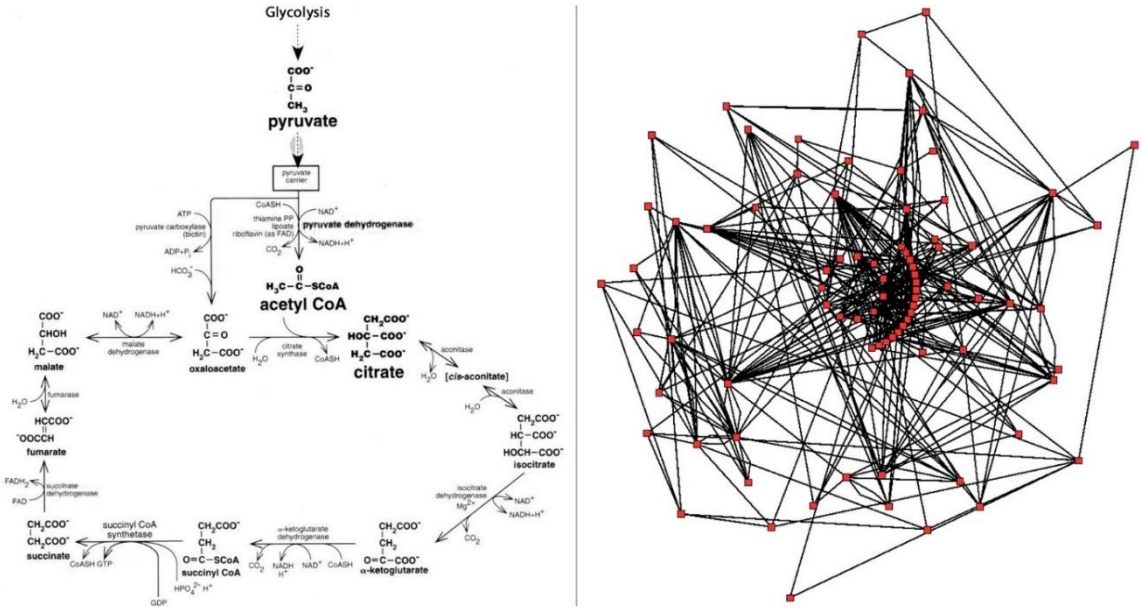

Figure 11: Comparison of a reductionist (left) and holistic view (right) of metabolism.

The image on the left, adapted from Salway (2004:11), shows a reductionist linear view of the tricarboxylic acid cycle. The image on the right, with permission from Vickers (2015), shows the network of interactions that the intermediates of the tricarboxylic acid cycle of Arabidopsis thaliana *(inner circle of red dots) have with other enzymes in alternative metabolic pathways.*

When considering IEMs, a holistic view of the metabolism is key. With a reduced or zero metabolic flux of the primary pathway, intermediates and substrates accumulate and may enter alternative pathways at higher rates than normal. This leads to the accumulation of metabolites to concentrations not normally associated with a healthy population. The depletion of products and by-products formed by the primary pathway also influences other pathways where these products serve as substrates. In addition to compounds associated with the primary pathway, depleted/elevated alternative substrates, intermediates and products can (1) correlate to a clinical-biochemical image (in the case of potentially harmful metabolites), (2) serve as diagnostic markers and (3) serve as an indicator of responsiveness to treatment (Chalancon *et al.*, 2013:1263-1264; Dry, 1997:4-8; Jeong *et al.*, 2000:651-652).

2.4.4 Involvement of non-enzymatic reactions

Enzymes are catalytic proteins and therefore decrease the activation energy (E_a) of chemical reactions by stabilising reaction intermediates. Enzymes may increase the rates of chemical reactions by up to 10^{16}-fold (Garret & Grisham, 2005:406). All enzymatic reactions can proceed uncatalysed, but at slower rates. Enzymes are essential for the regulation of pathways, albeit mostly through other cellular regulatory processes such as activators, inhibitors, co-factors, genetic regulation, compartmentalisation and degradation. Although some enzymes are less selective and can catalyse a number of substrates from the same class, most are highly selective, accepting a single substrate and a specific enantiomer (Cava *et al.*, 2011:818).

In addition to enzyme-mediated metabolism, and despite receiving little attention in recent years, non-enzymatic reactions are an integral and essential part of metabolism (Keller *et al*., 2015:153). Non-enzymatic reactions occur spontaneously but can also be catalysed by other small molecules such as acids, bases and metals. It is important to note that the same thermodynamic principles apply to both enzymatic and non-enzymatic reactions (both catalysed an uncatalysed). In all cases, the equilibrium position, total energy of the system and the change in Gibbs free energy (ΔG) remain unchanged (Keller *et al.*, 2015:153).

Non-enzymatic reactions are also involved in evolution, providing a template for the selection of enzymes. The development of an enzyme that increases the rate-limiting step of a non-enzymatic pathway could increase the growth rate and consequently improve the chances for natural selection of the organism. In some cases, enzymatic reactions can improve resistance to unfavourable conditions or alleviate the dependence on rare minerals. As an example, the isomerisation of ribulose-5-phosphate is catalysed non-enzymatically by free Fe(II), an Fe(II)-dependant enzyme in most prokaryotes and an Fe(II)-independent enzyme in most eukaryotes. Fe(II) is susceptible to oxidation to Fe(III), resulting in loss of catalytic activity in all but the

Fe(II)-independent organisms, giving eukaryotes the selective advantage (Keller *et al*., 2015:154).

According to Keller *et al.* (2015:153-154), non-enzymatic reactions can be divided into three classes:

- Indiscriminate, broad reactivity and low specificity: These reactions pertain to classes of compounds that typically react with each other, e.g. the conjugation of reducing sugars and primary amine compounds, oxidation of cellular constituents via reactive oxygen species and some alkylation, glycosylation and acetylation reactions.

- Purely non-enzymatic, highly specific and integrated: This class of non-enzymatic reactions forms an essential part of a metabolic pathway or the metabolic network, bridging gaps not catalysed by enzymes. The synthesis of vitamin D_3 can be used as an example: 7-Dehydrocholesterol is non-enzymatically (catalysed by ultraviolet [UV] radiation) converted to pre-vitamin D_3 in the skin, which in turn undergoes non-enzymatic (uncatalysed) isomerisation to vitamin D_3. Vitamin D_3 is then further metabolised by a number of enzymes.

- Parallel to enzymes: A diverse and widespread class of non-enzymatic reactions occur parallel to all six major classes of enzymatic reactions, as shown in Table 3. In some cases, non-enzymatic reactions may produce unwanted or harmful products, which is prevented by the higher reaction rates of enzyme-mediated reactions by diminishing the available substrate.

Table 3: Summary of the different classes of enzymatic reactions and corresponding parallel non-enzymatic reactions, adapted from Keller *et al.* (2015:154-155).

Class	Description	Enzyme class	Example
Redox	Transfer of electrons	Oxidoreductases	H_2O_2 degradation into H_2O and O_2. Catalysed by the enzyme catalase and non-enzymatically by metals, inorganic salts and other organic compounds.
Single replacement	Transfer of a functional group between substrates	Transferases	Biotransformation via glutathione. Glutathione-S-transferases catalyse the transfer of substrates to glutathione; this also occurs non-enzymatically with amongst others, 4-hydroxy-2-nonenal.
Hydrolysis	Breakdown of a substrate with H_2O	Hydrolases	Hydrolysis of 6-phosphogluconolactone into 6-phosphogluconate. This reaction is catalysed by 6-phosphogluconolactonase, but also occurs non-enzymatically. The increased rate of the enzymatic reaction prevents the formation of other undesired non-enzymatic end products from the same substrate.
Addition/ elimination	Non-hydrolytic addition or removal of functional groups	Lyases	Regeneration of tetrahydrobiopterin. This salvage pathway is catalysed by dihydropterin-4α-carbinolamine dehydratase and dihydropteridine reductase, but occurs non-enzymatically if the latter becomes inhibited.
Isomerisation	Intramolecular rearrangement of atomic or group positions	Isomerases	Isomerisation of dihydroxyacetone phosphate. This reaction is catalysed by triosephosphate isomerase and prevents the further non-enzymatic elimination of phosphate, which would produce the harmful metabolite methylglyoxal.
Synthesis	Formation of new compound from simpler precursors	Ligases	tRNA loading and hydrolysis. This reaction is catalysed by aminoacyl-tRNA synthetases. This may also occur non-enzymatically, especially with tryptophan, and could cause mischarging due to the lower substrate specificity of non-enzymatic reactions.

From the table above, the importance of distinguishing between enzymatic and non-enzymatic reactions is clear, but it is often neglected. The following section discusses chirality of organic molecules and how it can be used as a tool to make this distinction.

2.5 Chirality of products of enzymatic and non-enzymatic reactions

The term ‘isomers’ describes the relationship of a pair of compounds with the same chemical formula but with different structures. Refer to Figure 12 for an illustration of the relationship between the different kinds of isomers.

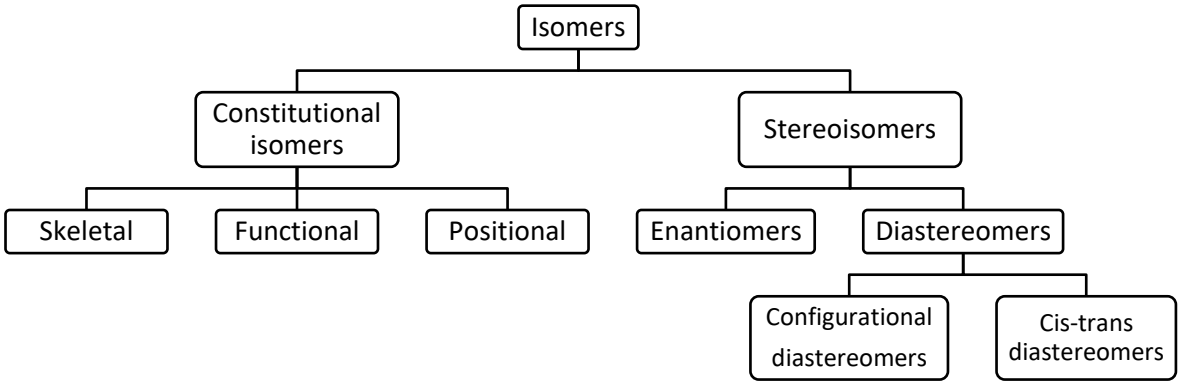

Figure 12: Summary of different kinds of isomers, adapted from McMurry (2004:294).

Constitutional isomers’ atoms are connected in different orders and can be divided into three groups: skeletal, functional and positional. Skeletal constitutional isomers have different carbon skeleton configurations (e.g. isobutane and butane), functional constitutional isomers have different functional groups constructed from the same atoms (e.g. ethyl alcohol and dimethyl ether) and positional constitutional isomers have the same functional group, but at different positions on the molecule’s backbone (e.g. isopropylamine and propylamine). With stereoisomers, the atoms/groups are connected in the same order, but with a different three-dimensional orientation or geometry. Stereoisomers can be divided into two groups: enantiomers and diastereomers (McMurry, 2004:75-77).

A molecule is stereogenic if its reflected image cannot be superimposed on the original. This is mostly caused by a chiral centre; when four different groups are covalently bound to a tetrahedral atom. In most cases, the chiral centre is a carbon atom, but this phenomenon can also be caused by silicon, sulphur, nitrogen and transition metal atoms. The latter should be in the correct oxidation state to form a tetrahedral coordination complex, such as cobalt(II), iron(II) and nickel(II). Trivalent nitrogen can act as a chiral centre with the lone pair of electrons counting as the fourth ‘group’, but its rapid umbrella-like inversion, which interconverts reflected images, complicates isolation. The same is true for trivalent phosphorus, but the much slower inversion rate allows for isolation of a stable species (McMurry, 2004:277-300).

A molecule is prochiral if it can be converted from achiral to chiral in a single reaction. In most cases, this is caused by (1) the substitution of one of two identical atoms/groups on a tetrahedral sp^3-hybridised atom, (2) the addition of a constituent to a sp^2-hybridised atom with trigonal planar molecular geometry, such as the carbon of a ketone or imine or (3) the substitution of an enantiotopic hydrogen/constituent on an achiral alkene (McMurry, 2004:301-303). A non-enzymatic reaction between two prochiral species in an achiral environment would always produce an optically inactive product, i.e. a racemic mixture (McMurry, 2004:295-297). Another structural feature that causes inherent stereogenicity, in the absence of a chiral centre, is dissymmetry, i.e. lacking an improper axis of rotation, planes of symmetry or an inversion centre; this includes planar chirality, axial chirality and atropisomers. Triphenylphosphine does not have a chiral centre, but is chiral due to its dissymmetrical propeller-like structure. A pair of chiral molecules can be defined as enantiomers (formally optical isomers) if they have the same chemical and structural formula and are non-superimposable mirror images of each other (Cava *et al.*, 2011:817).

The chemical and physical properties of enantiomers are identical in a symmetrical environment, but their three-dimensional conformation causes some unique interactions. Opposing enantiomers rotate the plane of plane polarised light (an asymmetrical medium) with the same amount but in opposite directions. A sample containing equal amounts of both enantiomers is termed racemic and is optically inactive due to a zero net-rotation of plane polarised light (McMurry 2004:280-281). In some cases, a chiral compound does not rotate plane polarised light due to its electronic properties and is termed cryptochiral. Opposing enantiomers also differ in some processes that are affected by the accessibility of their atoms/groups in the same way that a left hand would not fit in a right-handed glove (e.g. the active sites of most enzymes, inclusion complexation in cyclodextrin stationary phase chromatography and interactions with other chiral molecules) (Cava *et al.*, 2011:817). Enantiomers are distinguished from each other either based on their optical activity (+/-), relationship with glyceraldehyde (D/L) or configuration (*R*/*S*) (Cava *et al.*, 2011:817; McMurry 2004:280-285). Refer to section 1.6 for the naming conventions used to distinguish stereoisomers in this dissertation.

Multiple chiral centres in a molecule lead to multiple stereoisomers. Each chiral centre increases the maximum number of possible conformations by a factor of two. A pair of stereoisomers with more than one chiral centre can be defined as diastereomers if they have the same chemical and structural formulae and at least one similar and one dissimilar chiral centre. Contrary to enantiomers, diastereomers have different chemical properties and are not mirror images of each other (Wang & Chen, 2013:1663). Refer to Figure 13 for an illustration of the relationship

between enantiomers and diastereomers and Figure 14 for the phenomenon of *meso*-compounds.

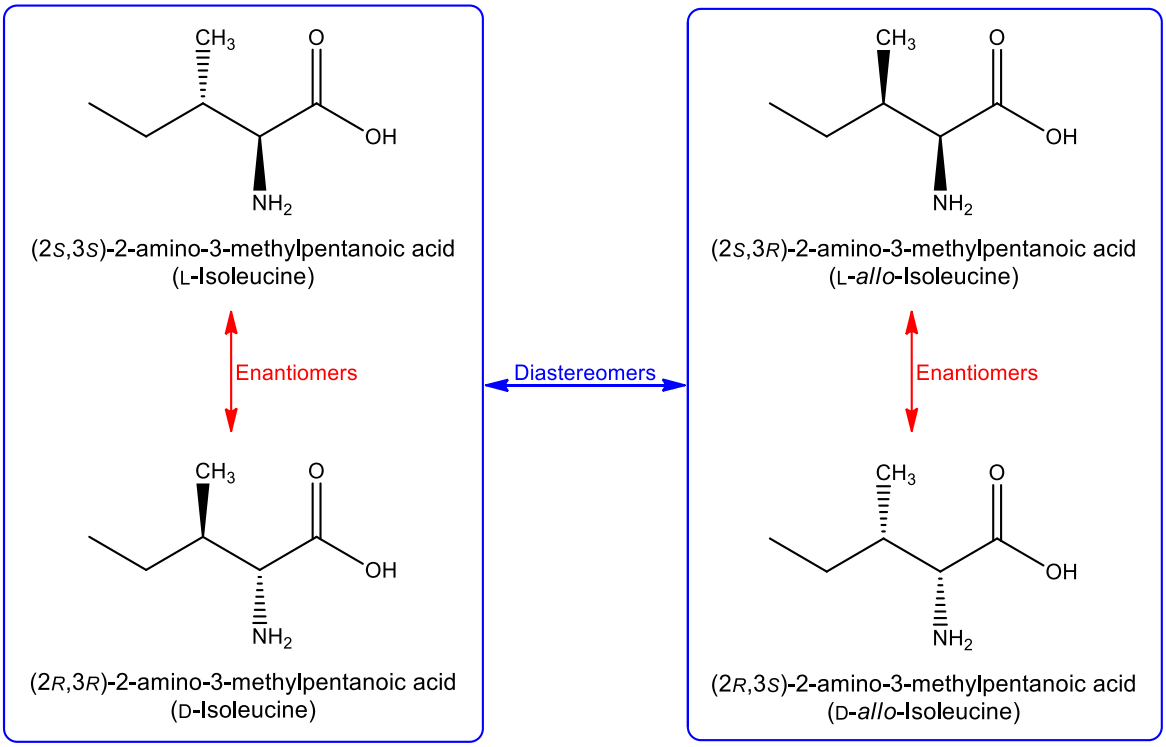

Figure 13: Illustration of the relationship between enantiomers and diastereomers.

Please note that in the illustration above, the double-headed arrows are not resonance arrows but instead indicate relationship. L-Ile and D-Ile differ on both chiral centres and are thus enantiomers. L-Ile has one similar chiral centre and one dissimilar when compared to both L-aIle and D-aIle and is thus a diastereomer of both these compounds. The same is true when comparing D-Ile with L- and D-aIle and vice versa *(McMurry, 2004:287). The full relationship of the stereoisomers of Ile is given in Table 4.*

Table 4: Relationship between the four stereoisomers of Ile, adapted from McMurry (2004:287).

Stereoisomer	Enantiomeric with	Diastereomeric with
L-Ile (2S,3S)	D-Ile (2R,3R)	L-aIle (2S,3R) and D-aIle (2R,3S)
D-Ile (2R,3R)	L-Ile (2S,3S)	L-aIle (2S,3R) and D-aIle (2R,3S)
L-aIle (2S,3R)	D-aIle (2R,3S)	L-Ile (2S,3S) and D-Ile (2R,3R)
D-aIle (2R,3S)	L-aIle (2S,3R)	L-Ile (2S,3S) and D-Ile (2R,3R)

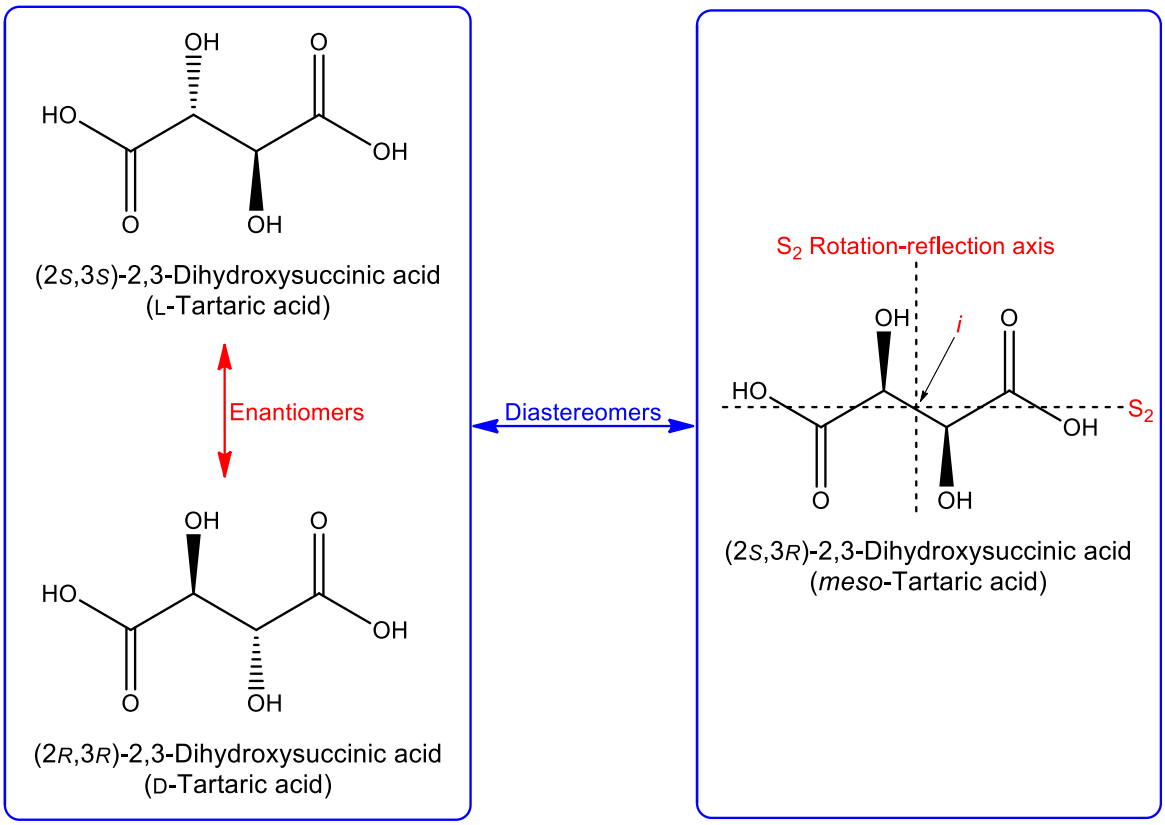

Figure 14: Illustration of the phenomenon of *meso*-compounds.

In the image above, i *is an inversion centre and S_2 is a two-fold improper rotation axis.*

In some instances, multiple chiral centres can result in a conformation with symmetry elements. *meso*-Tartaric acid (2S,3S) has two symmetry elements: an inversion centre (*i*) and a two-fold improper rotation axis (S_2). Inversion centres are present when the signs of a molecule's Cartesian coordinates can be reversed without the molecule's appearance changing. In other words, an identical atom exists diametrically opposite the inversion centre at equal distances from it. A two-fold improper rotation axis, or an S_2-rotation-reflection axis, is an axis around which a molecule can be rotated by 180°, followed by a mirror-reflection perpendicular to the axis. This reflection should then be superimposable on the original. Molecules with symmetry elements are achiral and optically inactive, but since they still possess chirality centres, they are termed *meso*-compounds. For example, in the case of *meso*-tartaric acid, the 2S,3R- and 2R,3S-conformation is the same molecule. When comparing *meso*-tartaric acid with DL-tartaric acid, however, the one similar and one dissimilar chiral centre still qualifies it as a diastereomer of the set (McMurry, 2004:288-291).

In some instances, a set of molecules may differ only by the positions of substituents around a restriction to free rotation, such as a double bond or ring: one molecule with the substituents on the same side (*cis*) and the other with the substituents on the opposite sides (*trans*) of the

restriction. The atoms of the different molecules in the set are still connected in the same order, but differ spatially around the restriction and are thus stereoisomers. They are not mirror images of each other and are thus not enantiomers. In most cases, *cis-trans* isomers differ in physical and chemical properties, retaining symmetry elements but lacking chirality centres. Therefore, they are termed optically inactive achiral diastereomers. If one side of the restriction consists of identical substituents, the *cis-trans* isomerism is not possible. The *cis-trans* notation is only used if the substituents in question are identical (on opposite sides of the restriction). For tri-, tetra- and dissimilar substituted alkenes, the unambiguous E/Z-notation is used instead (McMurry, 2004:93).

2.5.1 Molecular symmetry in nature

An interesting phenomenon is the difference in smells and tastes of enantiomers, caused by the specificity of the receptors in sensory organs. Where L-amino acids are generally tasteless, D-amino acids are sweet-tasting, *R*-carvone has a spearmint odour, whereas *S*-carvone has the anise-like odour of the spicy caraway fruit and *R*-limonene has the odour of oranges, whereas *S*-limonene has the odour of lemons (McMurry, 2004:300). In mammals, carbohydrates are predominantly D-orientated and amino acids mostly L-orientated (Podlech, 2001:44). Significant levels of D-amino acids (up to millimolar concentrations) are present in bacterial systems. Functions include synthesis of peptidoglycan layers, inhibiting growth, signalling molecules, biofilm disassembly control, spore germination regulation and a source of nutrients in some bacteria (Cava *et al.*, 2011:828). In addition, D-amino acids are incorporated into proteins of some animals (refer to Table 5), providing resistance to proteases and contributing to bioactivity. This is achieved either by post-translational isomerisation by racemases in some eukaryotes or by peptide synthesis by non-ribosomal peptide synthetases in prokaryotes (Cava *et al.*, 2011:828).

Table 5: Some peptides containing D-amino acids, adapted from Cava *et al.* (2011:819).

Peptide	Organism and source	Function
Dermorphin	Skin secretion of *Phyllomedusa sauvagii* (Argentinian tree frog)	Potent analgesic
Bombinins	Skin secretion of Bombinatoridae (fire-bellied toads)	Antimicrobial and haemolytic
Paecilodepsipeptide A	*Paecilomyces cinnamomeus* (fungus)	Anti-parasitic activity against *Plasmodium falciparum* (malaria)
ω-Agatoxin-IVB	Venom sac of *Agelenopsis aperta* (funnel-web spider)	Blocks voltage-sensitive calcium channels
Crustacean hyperglycaemic hormone	Decapod crustaceans	Neurohormone governing hyperglycaemia
Gramicidin S	*Bacillus brevis* (bacteria)	Antimicrobial, disrupts lipid bilayer
Lipopeptidolactones	*Pseudomonas* spp. (bacteria)	Biosurfactant

The most significant D-amino acids present in mammals include: D-Asp (a neurogenesis regulator) as well as D-Ser and D-Ala, which act as co-agonists of N-methyl-D-Asp-type Glu receptors in the brain. D-Ser is the most abundant D-amino acid in human urine (0.0021–10.1 µmol/mmol creatinine) where it also alters gene expression of uropathogenic *Escherichia coli*. (Luykx *et al.*, 2013:2020). Ser also has the higest D/L-ratio of human urine amino acids at 42%, with D-Arg (27%), D-Ala (18%) and D-Asn (17%) also showing high ratios (Wouter *et al.*, 2011:7134). Most other amino acids in human urine have been reported at below a 10% D/L-ratio for the D-enantiomer (Armstrong *et al.*, 1993:375). The D-BCAAs D-Val (0.88%), D-Leu (2.7%) and D-Ile (0.04%), however, are at much lower ratios in urine (Wouter *et al.*, 2011:7134).

Since two prochiral substrates (such as 2-keto acids) in an achiral environment would always produce a racemic mixture (McMurry, 2004:295-297) and most enzymes require and produce a specific enantiomer (Cava *et al.*, 2010:818), it could be established whether a reaction proceeded enzymatically or non-enzymatically by determining the enantiomeric composition of a sample. For this, a method to separate enantiomers is required. The N-AAA conjugates described in the variant MSUD cases above (refer to section 2.3.3) can be investigated in such a manner. By employing this chiral strategy, the main concern is the natural presence of D-amino acids in urine, which in the case of BCAA seems insignificant.

2.5.2 Separation of enantiomers

Asymmetric synthesis and resolution of racemic mixtures is of particular importance in the food, agrochemical, cosmetic (perfumes) and, particularly, the pharmaceutical industries. With most chiral drugs exhibiting a selective efficiency for a particular stereoisomer, racemic drugs are generally less efficient and potentially harmful (Wang & Chen, 2013:1663; McMurry, 2004:304-305). The separation of enantiomers in a sample is challenging due to the fact that enantiomers exhibit identical physical and chemical properties in an achiral environment. In the absence of a chiral environment, the resolution of enantiomers is impossible (Ward & Ward, 2014:20). Resolving enantiomers therefore relies on the introduction of an enantiomerically pure chiral agent to the otherwise racemic or achiral environment; this includes an enantiomerically pure reagent, catalyst, solvent, adsorption surface, stationary phases, acids/bases, buffers and/or seeding crystals. A number of general racemic resolution methods are discussed below:

(1) Kinetic resolution: Allowing a racemic sample to react with an enantiomerically pure reagent, or in the presence of an enantiomerically pure catalyst (which includes enzymes and microorganisms), results in an enantioenriched sample of the less reactive enantiomer. This results either from the different chemical properties (e.g. E_a) of the diastereomeric TS/substrate-catalyst complex or due to the conformation of a substrate allowing or disallowing fit into an active site. The unreacted enantiomer substrate and altered product of the opposite enantiomer can then be separated by conventional methods. This route would, however, be ineffective if the E_a difference is small, the reagents/catalysts are expensive and/or the separation of starting racemates and end products is problematic. The industrial application described in section 2.5.3 is an example of such a method (Sato & Tosa, 1993:3-14).

(2) Fractional crystallisation: In 1848, Louis Pasteur proved that chirality has a molecular basis. His observation that sodium ammonium tartrate crystals had a chiral macrostructure allowed him to manually sort dissimilar crystals into two piles. A solution of the separate piles rotated the plane of plane polarised light in different directions (McMurry, 2004:281). In 1882, Pasteur showed that seeding a supersaturated solution of racemic tartaric acid with the different enantiomeric crystals allowed the formation of the respective crystals on opposite sides of the reactor. Although seemingly impractical, this classical method of enantiomeric resolution remains an active area of development in both the laboratory and industry (Wang & Chen, 2013:1663).

(3) Chiral reagents: In 1853, this method was also introduced by Louis Pasteur. A set of enantiomers (containing dissimilar chiral centres) is allowed to react with an

enantiomerically pure reagent, which then creates an additional, but similar, chiral centre on both enantiomers. This chiral auxiliary fulfils the requirements for a set of diastereomers, which can then be separated by conventional methods, including gas chromatography-mass spectrometry (GC-MS), high-pressure liquid chromatography (HPLC) and nuclear magnetic resonance spectrometry. The most widely used method in the industry is the formation and separation of crystalline dissociable diastereomeric salts (Wang & Chen, 2013:1663). For chiral derivatisation to be effective, a number of requirements must be met: (1) the derivatisation reagent must be enantiopure to limit the formation of an additional set of diastereomers, which would be enantiomers of the first, nullifying the effect, (2) the reaction must continue to completion under the reaction conditions to prevent kinetic resolution as described above, (3) the reaction conditions should not racemise the analytes or reagents and (4) the resulting diastereomeric products should possess desirable qualities required by the final resolution step of the method (e.g. a chromophore for UV detection, adequate volatility for GC-MS analysis or an adequate solubility difference for differential crystallisation) (Wang & Chen, 2013:1663). A chiral reagent was used in this study to separate amino acids enantiomers lysed from the N-AAA conjugates.

(4) Chiral chromatography: This method can be subdivided into three approaches: (1) Pertaining to HPLC, a chiral mobile phase or modifiers may be used, though this proves to be expensive in most cases, (2) a column containing a chiral stationary phase, forming temporary diastereomeric adsorption complexes, or (3) a column containing inclusion complexes that exhibit differential retention for enantiomers based on accessibility and conformation (refer to Figure 15 and subtext) may be used. The use of a chiral or inclusion complex stationary phase is a powerful and versatile method. However, the selectivity of the phase depends on a number of factors that are hard to predict (Lambrecht, 2011). On a small scale, selecting the best column is an expensive and tedious venture, but once a method is standardised, chiral column chromatography has a lower running cost and is less laborious than chiral derivatisation. For this reason, chiral column chromatography has superseded the use of chiral derivatisation in most small-scale routine laboratory applications, but chiral derivatisation remains more practical at a larger industrial scale or for single-use investigative research such as this study (Ward & Ward, 2014:20-23).

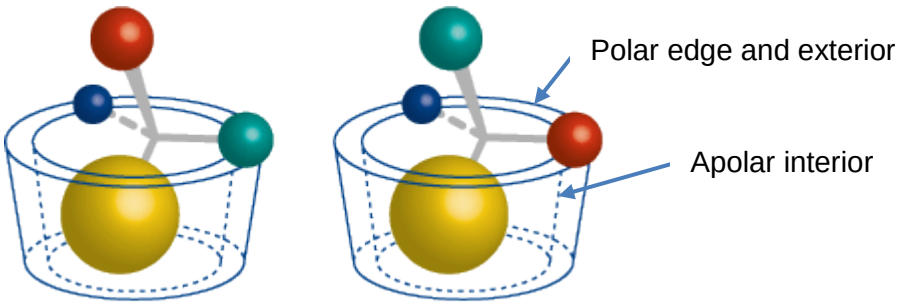

Figure 15: Cyclodextrin inclusion complexes of enantiomers, adapted by permission of Lambrecht (2011).

Cyclodextrins are cyclic oligosaccharides. Six to eight α-D-glucopyranoside units are linked through α-1,4-linkages, forming a toroid with a larger and smaller opening (cone-like shape) with an apolar interior and a polar edge and exterior. The primary and secondary hydroxyl groups of the α-D-glucopyranoside subunits are accessible in the inside and outside edges, respectively. In the image above, the yellow groups of the enantiomers form an inclusion complex with cyclodextrin, temporarily retaining the compound. Because of the spatial arrangement differences, the compound on the left further interacts with cyclodextrin via the teal and blue atoms, while the compound on the right interacts with the red and blue atoms. The difference in chemical properties between the teal (left) and red atoms (right) leads to separation of the enantiomers (Lambrecht, 2011; Sigma-Aldrich, 2016).

2.5.3 Chemical synthesis and industrial application

The routine method for synthesising N-AAA conjugates in the laboratory is the acylation of an amino acid with a carboxylic acid. This is achieved in the presence of a catalyst, usually an aqueous base, and a coupling agent such as the chloride salt of the organic acid (Schäfer & Bode, 2014:1526). The resulting N-AAA conjugates are a racemic mixture. Since aminoacylases are L-amino acid-specific, they have been used since the 1950s to produce enantiomerically pure L-amino acids from N-acylated-DL-amino acid conjugates. Aminoacylases are immobilised on a diethylaminoethyl cellulose crossed-linked dextran gel column and the reaction mixture is then passed through. The liberated L-amino acids can subsequently be easily purified and the remaining N-acylated-D-amino acid conjugates can either be racemised and reinjected into the column flow or chemically hydrolysed without racemisation (with HCl) to procure the enantiomerically pure D-amino acids (Sato & Tosa, 1993:3-14).

In addition to the enantiomer resolving methods described above, a sample (especially biological fluids) first needs to be pre-treated to remove interfering matrix substituents to a point where adequate detection of the enantiomer of interest is possible (Wilson & Walker, 2005:498).

The following section discusses the techniques used in this study to prepare a urine sample to investigate the enantiomeric composition of N-AAA conjugates via GC-MS.

2.6 Analytical techniques and methods

2.6.1 Sample preparation

Commonly, urine requires little or no pre-treatment, depending on the method to follow. Preservatives such as toluene, sodium carbonate, thymol and certain acids are sometimes added to prevent bacterial growth during transport and storage, but are mostly avoided to maximise the number of tests that can be performed. Some methods (e.g. direct infusion MS-MS) require crude urine, in which case centrifugation – to remove insoluble material – and isotope dilution may be the only sample preparation required (Wolfender *et al*., 2015:144). Other pre-treatment strategies include pH adjustment (to increase/decrease solubility of a given compound class) (Wilson & Walker, 2005:499), addition of a buffer, enzymatic digestion of unwanted compounds (Van Slyke *et al.*, 1943:251) and chemical stabilisation (e.g. oximation of keto acids to prevent keto-enol tautomerisation) (Wolfender *et al*., 2015:142). The use of pH adjustment, together with heat, can also be used for the hydrolysis of conjugates (Kaiser & Benner, 2005:318; Sato & Tosa, 1993:3–14).

Modifications to the matrix of a sample may be required before chromatography and detection to enhance the effectiveness of these processes. The aim of sample clean-up, therefore, is to reduce the matrix effect by extracting the metabolite class of interest from the background. In most cases, this is achieved as a compromise between yield, purity, selectivity and unwanted chemical modifications (Wilson & Walker, 2005:498-499). Sample clean-up strategies include, but are not limited to, solid phase extraction (SPE), liquid-liquid extraction, evaporation, desalting and protein precipitation. In this study, SPE was used to remove interfering substances for amino acid analysis, whereas liquid-liquid extraction was used for organic acid analysis. These methods are described as follows:

- SPE relies on the principles of chromatography. A column containing a porous stationary phase is placed upright onto an SPE-tank. A native sample or sample diluted in a mobile phase is added to the column, which then passes through the column with the aid of gravity or a vacuum. Typically, the analytes of interest are strongly retained, whereafter the column is washed with clean solvents. The compounds of interest are eluted with an eluting agent that alleviates the intermolecular bonding forces between the analytes and the column (Wilson & Walker, 2005:499).

- Solvent extraction (also known as liquid-liquid extraction) is the most common sample clean-up technique and depends on the principles of partition coefficients. Weak electrolytes, like organic acids, can exist in both ionised and unionised forms in an aqueous environment. By manipulating the sample's pH, the equilibrium can be pushed towards a favourable state, e.g. the addition of an acid produces unionised organic acids, which favour a more apolar solvent. These unionised compounds can be extracted by mixing the sample with a low-boiling-point water-immiscible solvent such as ethyl acetate. This often leads to the concurrent extraction of some water, which can be removed with the addition of an anhydrous salt (Wilson & Walker, 2005:499).

2.6.2 Derivatisation and chromatography

Derivatisation is the chemical modification of a compound to produce a new compound with a similar structure, but with new favourable chemical properties for the method to follow. Generally, in organic chemistry this is achieved as a substitution reaction (two reactants exchange parts to yield two new products) (Schummer *et al*., 2009:1482). According to Schummer *et al*. (2009:1482-1483) and Wilson and Walker (2005:500), derivatisation can be useful on various levels:

- Optimisation of chromatography by increasing/decreasing polarity or allowing separation of racemic mixtures by converting enantiomers into diastereoisomers.

- Decreasing the compound's evaporation temperature and increasing thermal stability. This allows the evaporation of compounds in the temperature range achievable on a gas chromatography (GC) system.

- Improving detectability (e.g. enabling neutral loss scan on an MS-MS), improving fragmentation pattern in a mass selective detector (MSD) and introducing a detector-orientated marker on analytes (e.g. the addition of a chromophore/fluorophore, which enables detection on a UV-visible detector).

Chromatography is a technique used for the separation of analytes in a mixture, achieved by passing the mixture over or through a stationary phase with the aid of a mobile phase. The different analytes in the mixture have varying solubilities in the mobile phase and varying affinities for the stationary phase. Analytes are temporarily adsorbed onto the stationary phase (some stronger than others) through intermolecular bonding forces, which delays their movement and facilitates separation (Wilson & Walker, 2005:485-486).

Various types of chromatography, including thin layer chromatography, liquid chromatography and GC, can be utilised to separate analytes in a complex mixture (Wilson & Walker, 2005:487). Coupled with an MSD, GC is a high-resolution method, able to analyse complex mixtures with high sensitivity and was applied in this study. GC is one of the most widely used techniques despite the need for derivatisation in most cases (Wolfender *et al.*, 2015:142). The gas (mobile phase) carries the analytes through a column with a bonded stationary phase and is therefore referred to as a carrier gas. The column is placed in an oven that is either kept isothermal or gradually heats up. Analytes are separated based predominantly on their evaporation temperatures and are then carried through the column where additional separation occurs based on their affinity for the stationary phase. Analytes also separate into bands, which are visualised by converting the signal from a detector (such as an MSD or flame ionisation detector) into a visual representation called a chromatogram with the aid of computer software (Wilson & Walker, 2005:541-545). Analytes with formal charges (such as choline) or strong dipole moments cannot be analysed on a GC system. In addition, analytes should be volatile and thermally stable at or below the operating temperatures of a GC. For this reason, suitable analytes are typically restricted to relatively low molecular weights of approximately 800 Da. Derivatisation can be applied to assure that analytes comply with these requirements (Wolfender *et al*., 2015:142).

2.6.3 Detection and data processing

Detection of metabolites in a complex mixture has developed greatly since the 1960s: from identifying chromatography peaks solely on retention times to combining with sophisticated detection techniques such as MS, time-of-flight mass spectrometry and infrared-spectrometry with software-assisted identification and quantification (Fiehn, 2001:158; Walker, 2005:545). The main function of a detector is the accurate identification and quantification of compounds in the relevant matrix while maintaining high sensitivity, dynamic range and reproducibility. Detectors can be divided into two main groups, namely destructive and non-destructive. GC systems are usually coupled with destructive detectors such as an MSD, flame ionisation detector or photoionisation detector. These detectors typically have a higher sensitivity and dynamic range. Liquid chromatography systems are usually coupled with non-destructive detectors such as UV/visible, diode array or fluorescence detectors. Eluent from these non-destructive detectors can be collected in fractions or discarded as waste (Stashenko & Martínez, 2014:17).

The essential features of an MSD can be summarised as follows: (1) production of ion fragments through electron impact or chemical ionisation, (2) acceleration of ions by means of an electric field, (3) separation of ions based on their mass-to-charge ratio in a mass selector

and (4) detection of ions and measurement of their abundances (Wilson & Walker, 2005:405). The detector can be utilised in single-ion monitoring or scan mode in which multiple analytes are assessed within one analytical run. Deconvolution software such as the Automated Mass Spectral Deconvolution and Identification System (AMDIS) from the United States National Institute of Standards and Technology (NIST) or ChemStation from Agilent Technologies, are used to facilitate data processing, which includes peak identification through deconvolution (Wilson & Walker, 2005:445-446). In the case of destructive detectors, where identifiable unique patterns are generated, patterns are compared with those in libraries of data to determine the identity of a particular compound. The software also uses peak integration to calculate concentrations of the specific analyte, which can be compared to population averages, known as reference ranges (Wolfender *et al*., 2015:142). In addition to the analytical techniques and methods described, theoretical methods such as molecular modelling can also be used to investigate N-AAA conjugates.

2.7 Molecular modelling

Advances in computer technology and understanding of molecular and quantum mechanics (the mathematical description of subatomic particles) have made the use of high-performance computing and advanced computer software for molecular modelling possible (Leach, 2001:8). Molecular modelling is a computational chemistry technique that uses powerful software to provide insights into, amongst other things, molecular characteristics of substrates, reaction intermediates and products. In addition, the feasibility of a proposed reaction mechanism can be evaluated by determining the E_a required for chemical reactions to take place (Leach, 2001:1; Das *et al.*, 2014:6-14). For example, for substrates to react, the following must be true: There should not be substantial steric hindrance, the energy gap between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of substrates should be small enough, there should be a similarity between the symmetry of the reacting reagent's molecular orbitals and the bonding orbitals should be correctly orientated (Jensen, 1999:347-351; McMurry, 2004:350). However, results generated from computational chemistry techniques such as molecular modelling should, if possible, be verified with laboratory experimental data and *vice versa* (Fourches *et al.*, 2010:1189).

Molecular modelling aids in the study of biochemical reactions, enzyme structures and active sites and cell membrane characteristics, amongst others, especially where the component under investigation is unstable and short-lived. In the case of non-enzymatic reactions, the E_a of the reaction can be compared to the typical (experimentally determined) E_a of enzymes to ascertain the feasibility of an *in vivo* reaction. Refer to Table 6 for an example of typical enzymes' E_a.

Table 6: Examples of typical enzyme activation energy (E_a) requirements.

Enzyme	E_a (kJ/mol)	Reference
Human glucokinase (activation step)	65.5	Matschinsky *et al.* (2010:239)
Human glucose-6-phosphate dehydrogenase	34.1*	Aksoy *et al.* (2001:290)
Human lecithin cholesterol acyltransferase	74.9–82.0*	Zorich *et al.* (1985:8833)
Mammalian alkaline phosphatase	39.0–91.0	Craig *et al.* (1996:5251)
Mammalian ACY1	Approx. 31	Su *et al*. (2007:580)
Mammalian glucose 6-phosphatase	41.8–54.8*	Kelmer-Bracht *et al.* (2003:53)
Mammalian pyruvate carboxylase	38.1*	Scrutton (1974:7060)
Gluconobacter diazotrophicus pyruvate decarboxylase	46.0	Van Zyl *et al.* (2014:4)
Jack bean (*Canavalia ensiformis*) urease	35.3	Fidaleo and Lavecchia (2003:315)
Bacteriophage T4 DNA ligase (end-joining reaction)	68.2*	Cherepanov and De Vries (2003:4320)

* Some values were reported in kcal/mol and were converted to kJ/mol by multiplication with 4.184. ACY1: aminoacylase I.

Ideally, for the purpose of this study, the data in Table 6 should have contained enzymes from human origin only, catalysing reactions similar to the proposed reaction mechanism (C-N bond-forming enzymes and decarboxylation enzymes). However, this data proved difficult to obtain. Experimentally, the most commonly used method to determine the E_a of a reaction is the Arrhenius equation. Reactants together with any required constituents (such as a catalyst) are incubated at different temperatures and the rate constant is determined. The natural logarithm of the rate constant is subsequently plotted against the inverse of the temperature and the E_a is derived from the slope (slope = $-E_a/R$, where R is the gas constant). It is worth noting that in the case of enzymatic reactions, this method is imperfect for the following reason: Incubating enzymes at different temperatures may result in poor enzyme performance, resulting in a non-linear curve on the Arrhenius plot (Londesborough, 1980:211).

However, before any properties of a proposed reaction mechanism can be determined, the various structures need to be geometrically optimised and verified as stable minimum-energy structures or TS structures.

2.7.1 Geometric optimisation and potential energy surfaces

Optimal bond lengths and angles for most small molecules are known and widely available in the literature. With larger molecules, however, dipole moments, Coulomb repulsion, Van der Waals interactions and strain energy, amongst others, may cause a small shift in the standard

bond lengths and angles within a molecule in comparison with smaller molecules. This results in deviations in atoms' coordinates and the electronic field in the entire molecule, affecting its electron energy (E_e), e.g. the C-C bond length in butane is 153 pm, but the C-C bond length in a diamond is 154 pm (Bartell, 1959:3497). Since all the bond lengths and angles for larger molecules are not available, molecular modelling software is required to calculate the geometrically minimal configuration of a molecule (and its relative $_{Ee}$) before any other calculations can be done.

Initially, a structure called the guess value is drawn in the software. One approach (the approach used in this study) is then to use density functional theory (DFT) which relates electron density in the structure to the E_e of the structure. The software therefore first calculates the starting relative E_e of the roughly optimised structure and then makes small changes in the atoms' coordinates. If the positions of a molecule's nuclei change, the bond angles, bond lengths, torsion angles and proximity of atoms/groups change, which lead to changes in the electron density. This change in electron density leads to a change in E_e. If these changes in E_e are plotted against each set of changes in the structure (the so-called reaction coordinates, i.e. bond length, etc.), a curve as shown in Figure 16 is obtained. The process is repeated until the first derivative of the function of E_e to the atoms' coordinates approaches zero (refer to Equation 1) (Leach, 2001:5,253-258).

Equation 1: First derivative approaches zero of the function for energy to atom coordinates.

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\frac{\delta E}{\delta x}=0$$

In Equation 1, E *denotes energy,* x *denotes Cartesian coordinates of a set of atoms and* h *represents the change in the atoms' coordinates, which can be a combination of several simultaneous changes in the molecule's three-dimensional conformation (not to be confused with bonding configuration/arrangement).*

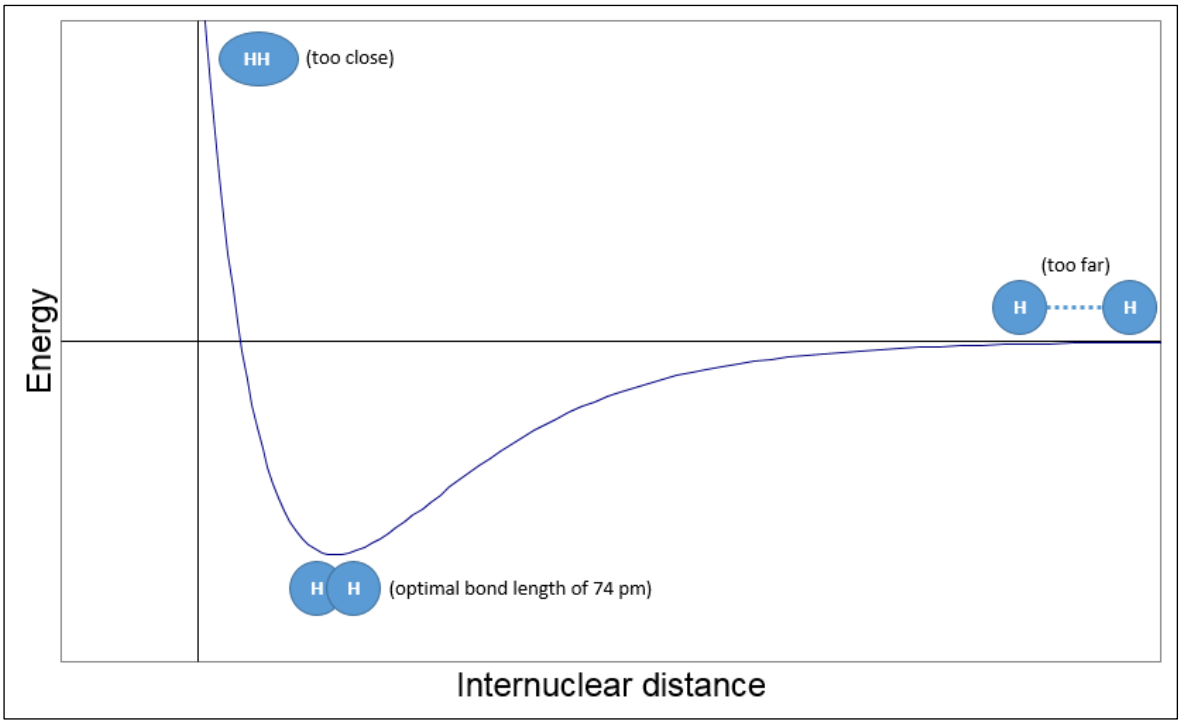

Figure 16: Optimal bond length for diatomic hydrogen, simplified illustration adapted from McMurry (2004:11).

As a simple example, Figure 16 is a plot of energy versus the bond length of diatomic hydrogen. If the nuclei are too close together, the positively charged nuclei will repel each other, which increases the energy of the molecule. At the optimal bond length of 74 pm, the valence electrons are shared, resulting in a stable molecule with the lowest energy value. If the bond length increases further, the atoms would not be able to share the bonding electrons, resulting in a less stable system (McMurry, 2004:10-11). At a bond length of 74 pm, diatomic hydrogen is in a geometrically minimal configuration and the first derivative (slope) of a function of its energy to bond length is zero. This means that further changes in the bond length will result in an increase in energy. In the example above, the energy change was the result of a change in a single parameter, which provides a two-dimensional graph. In larger molecules with multiple bonds, bond angles and torsion angles, changes in energy can be considered as movement on a multidimensional surface or potential energy surface (PES) (Leach, 2001:5), as shown in Figure 17.

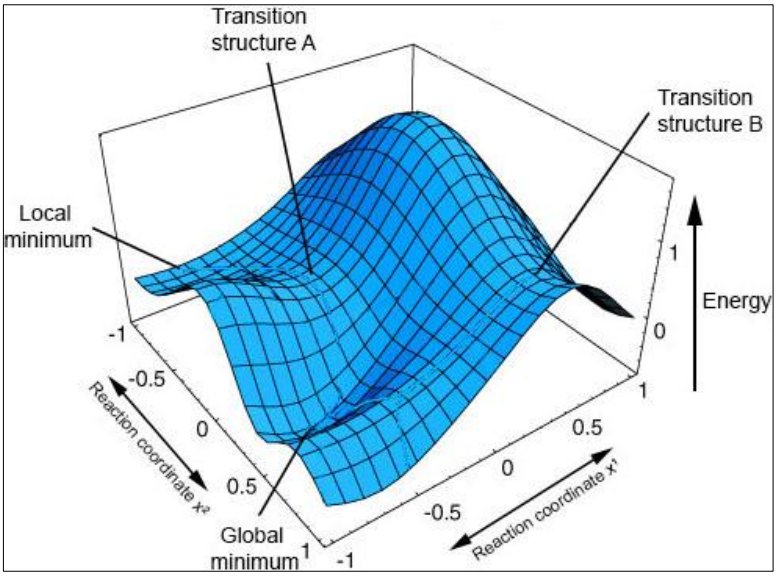

Figure 17: Potential energy surface in larger molecules with multiple bonds, bond angles and torsion angles, simplified illustration adapted from Schlegel (2003:1514) by permission of Oxford University Press.

If the first derivative of energy over a coordinate ($\delta E/\delta x$) equals zero, a turning point in the energy curve is observed. If the second derivative of energy over the above-mentioned coordinate ($\delta^2 E/\delta x^2$) is positive, it is at a minimum. However, if the second derivative is negative, the structure is considered a transition (Leach, 2001:4).

One of the limitations of geometrical optimisation (refer to section 1.5 for nomenclature) is that it may produce either a local or a global minimum (ground state conformation), depending on the starting structure, e.g. refer to Figure 17 and Figure 18: Cyclohexane has two stable structural isomers, chair and twist-boat (Bauld, 2001). Each geometrical configuration lies within a minimum and is thus stable (the first derivative for both is zero and any change in coordinates will result in an increase in E_e for both conformations). The chair conformation has the lowest relative E_e of the two structures and therefore occurs in a global minimum, whereas the twist boat occurs in a local minimum. If the starting structure of a geometrical minimisation for cyclohexane resembled a structure in between a chair and a half-chair conformation, the resulting optimised structure would be a chair conformation (global minimum). However, if the starting structure resembled a structure in between a half-chair and twist-boat conformation, the resulting optimised structure would be a twist boat (local minimum) (refer to red arrows in Figure 18) (Bauld, 2001).

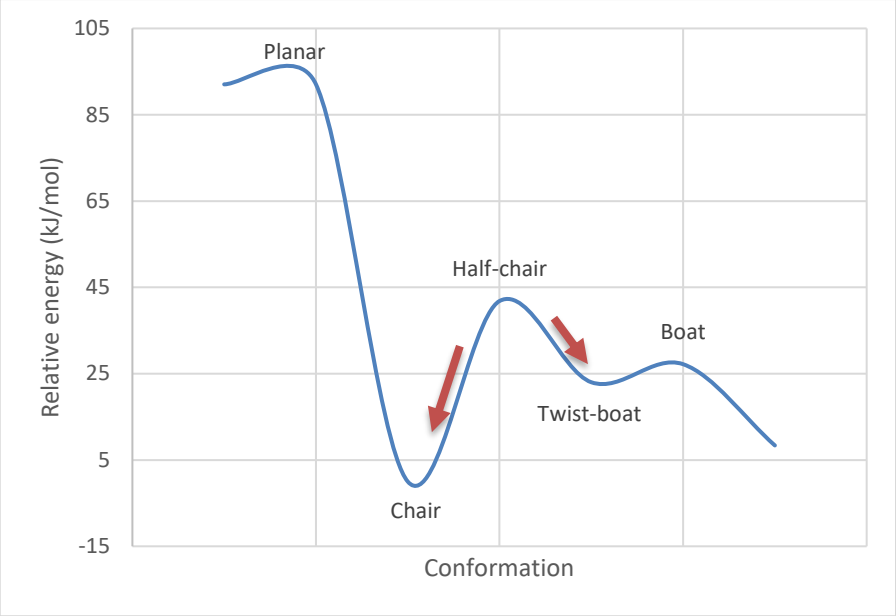

Figure 18: Conformational energy diagram for cyclohexane, adapted from Bauld (2001).

A search for the global minima of a complex compound is possible but extremely time-consuming, since complex compounds may have multiple minima. Starting a TS search from a local minimum (instead of a global minimum) would result in a lower calculated E_a and a larger or smaller ΔG for exo- or endothermic reactions, respectively. Although this is not ideal, the E_e of a molecule is still significantly lower in a local minimum position than it would be without geometric minimisation and should not equate to significantly erroneous results. In some cases, however, a local/global minimum does not lie in the correct reaction path on the PES, which results in a failure to acquire a TS. In such cases, the structure's conformation should be revised to search for a better minimum (Jensen, 1999:339-340). Atoms are not motionless but vibrate constantly and deviate from equilibrium positions. As part of the geometric minimisation/optimisation calculations, vibrational energy can also be calculated.

2.7.2 Infrared spectrometry and imaginary vibrations

The change in energy associated with the above-mentioned vibrations falls in the region of infrared light. By calculating vibrations, the previously calculated E_e can be converted to the more meaningful ΔG value (Leah, 2001:273-274). By computing frequencies, it is possible to identify modes of negative curvature (imaginary vibrations) or positive curvature (real vibrations). Imaginary vibrations obtained from vibrational analysis are by convention given as real negative numbers. A minimised structure should not have any imaginary vibrations, whereas an optimised TS should have only one imaginary vibration (Leah, 2001:279-282). Vibrational analysis can therefore be used as a tool to indicate whether the desired minimised

structure or optimised TS has been obtained. Refer to Figure 17: A TS can be defined in terms of molecular modelling as follows: When considering the function for E_e against the reaction coordinate, the TS has a zero value for the first derivative and a negative value for the second derivative. When observing a PES, the TS lies on the highest point in a valley, known as the saddle point. Energy increases with any change in reaction coordinates, except for one direction, which is the reaction path (Jensen, 1999:296-298).

In conclusion, molecular modelling can be used to study chemical properties of molecules and chemical systems by simulating the behaviour of molecules. This can in turn can be applied to study biochemical reaction mechanisms and determine the E_a required for chemical reactions to take place. The latter is useful when investigating non-enzymatic reactions, which can be compared to typical E_a values of enzymes.

CHAPTER 3: AIMS AND OBJECTIVES

3.1 Background

During routine analysis of biofluids in an analytical laboratory, it is not uncommon to come across a sample with unknown constituents in significant concentrations. This is especially true for untargeted or semi-untargeted applications, including analysis of amino acids and organic acids, where the MSD is used in scan mode. These unknown constituents may be artefacts of exogenous origin (e.g. dirty laboratory apparatus, impure reagents and unwanted reactions with reagents), metabolites associated with diet and medication intake, bacterial contamination/degradation of sample (microbiome) or unidentified metabolites not in the compound library. These unknowns may increase doubt as to whether the metabolites are associated with a disorder/condition or whether all relevant information has been extracted from the data. In some cases, a constituent may be positively identified, but its origin may be unclear and the functionality thereof unknown. Determining the origin of the constituent may provide insights into the patient's disorder or condition, help elucidate unknown metabolic pathways, and/or assist in troubleshooting a troublesome method (in the case of exogenous artefacts).

As a practical example, the N-AAA conjugates (N-isobutyryl, N-isovaleryl and N-2-methylbutyryl conjugates of Val, Leu and Ile) previously identified and quantified in the urine of three cases of MSUD (Dry, 1997:28-31,87-95; Mienie *et al*., 2001:42-44) were investigated. The above-mentioned nine N-AAA conjugates were a group of many alternative metabolites observed in these three cases. An extensive literature search did not reveal other reported MSUD cases exhibiting this unique metabolic profile. The mechanism of formation and the clinical-biochemical effect of the observed N-AAA conjugates were deemed unclear. The original studies in which these conjugates were observed (Dry, 1997:86-87,95-109; Mienie *et al*., 2001:51-52) tested, described and proposed various hypotheses to elucidate the formation of N-AAA conjugates in the above-mentioned samples.

The most apparent origin of these N-AAA conjugates would have been the biotransformation of the respective N-acylated-CoAs with BCAAs by the GLYAT enzyme. The observation of these N-AAA conjugates in MSUD cases was unexpected, since Gly is a preferred substrate of GLYAT and N-acyl-Gly conjugates were not observed (Sweetman & Williams, cited by Dry, 1997:108; Van der Westhuizen *et al*., 2000:102-103). The biotransformation hypothesis was further complicated by the fact that the formation of the most apparent substrates of GLYAT in this case (N-isobutyryl-, N-isovaleryl-, and N-2-methylbutyryl-CoA) was prevented/depleted by the enzyme defect in MSUD (Dry, 2001:138; Mienie *et al.*, 2001:51). Confirming an alternative

mechanism of N-AAA formation as proposed by Dry (1997:95-108) and Mienie *et al.* (2001:51-52) subsequently served as motivation for this study. The most probable hypothesis included the non-enzymatic N-AAA conjugate formation via the reaction between the elevated 2-keto acids and ammonia described in section 2.3.3.

Since non-enzymatic reactions yield both L- and D-enantiomers and enzymatic reactions use and yield only specific enantiomers (Cava *et al.*, 2011:818), determining the chirality of unknown compounds may provide insights into their origin and potential functionality in disease and health. In addition, molecular modelling may be used as an assistive technique to verify proposed mechanisms as to how these compounds may form.

3.2 Aims

The aims of this study were as follows:

(1) To develop a method for investigating the chirality of amino acid conjugates.

(2) To set up a molecular modelling procedure to verify the feasibility of proposed reaction mechanisms for non-enzymatic organic reactions.

3.3 Objectives

The following objectives were formulated:

(1) In relation to aim (1), the first objective was to demonstrate the chiral strategy by developing a method for investigating the chirality of the N-AAA conjugates reported in the urine of sample 11 (one of the three MSUD cases mentioned above) by:

   (a) Setting up a method for the extraction of N-AAA conjugates from urine, without simultaneous extraction of native BCAAs, and the subsequent hydrolysis to liberate free BCAAs from the extracted N-AAA conjugates.

   (b) Developing a qualitative method for the separation and identification of enantiomers of racemic BCAAs, by chiral derivatisation with *R*-2-butanol and N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide (MTBSTFA) with 1 % tert-butyldimethylchlorosilane (t-BDMCS) and subsequent GC-MS analysis.

   (c) Developing a method to investigate the feasibility of the proposed substrates for non-enzymatic formation of N-AAA conjugates, by attempting synthesis in an applicable *in vitro* environment.

(2) In relation to aim (2), the second objective was defined as the following:

(a) Proposing a refined reaction mechanism based on the mechanism proposed in the literature for the non-enzymatic formation of N-AAA conjugates from 2-keto acids and ammonia.

(b) Investigating the feasibility of this proposed reaction mechanism in terms of E_a requirements and assessing whether the energy requirements were achievable in a physiological and/or laboratory environment (e.g. during sample handling).

3.4 Scope, substantiation and challenges of the study

3.4.1 Qualitative approach

The aim of the methods developed and set up in this study was to investigate the chirality of the observed N-AAA conjugates and to test the feasibility of the formation of these conjugates from the proposed substrates. Since accurately quantifying the amino acid enantiomers and synthesised N-AAA conjugates was not part of the aims and objectives, developing and validating quantitative methods were not pursued.

For the separation of racemic amino acids, two approaches are possible: (1) using a chiral stationary phase or (2) derivatisation with chiral reagents. Since a chiral stationary phase method is more expensive to develop, chiral derivatisation was pursued. The use of *S*-2-butanol as an O-esterification chiral derivatisation reagent, to separate racemic amino acids during GC, has been described by Bremer *et al*. (1981:483). Bremer *et al*. (1981:483) also used pentafluoropropionic anhydride or trifluoroacetic anhydride for N-acylation. A modified method with *R*-2-buthanol (instead of *S*-2-butanol) and MTBSTFA with 1% T-BDMCS (instead of fluoroacyl anhydrides) derivatisation was selected for this study due to availability and the latter's ability to react with amines to completion and improve the separation of enantiomers via GC-MS (Schummer *et al*., 2009:1482).

3.4.2 Sample availability

Unfortunately, previous studies exhausted all collected material from all but one of the samples described in section 3.1. Only 3 mL urine remained from this sample in which the N-AAA conjugates were most abundant. This sample will be referred to by its anonymised identity as "sample 11". Since the samples were anonymised, recollection was not an option. No other MSUD urine samples available to the study exhibited the same metabolic profile. In the light of this, only sample 11 could be used to demonstrate the method. The experimental findings were based on an isolated observation in an amino-acid-specific disease group. Care should

therefore be taken not to generalise findings presented in this dissertation. Future MSUD patients exhibiting the same metabolic profile, and other defects with elevated levels of N-AAA conjugates, may potentially be subjected to the methods of this study to further support or contradict any conclusions. The next chapter gives a comprehensive description of the materials and methods used to persue the aims and objectives of this study.

CHAPTER 4: METHODS

4.1 Introduction on methods of investigation

In short, the analytical strategy of this investigation was to first isolate N-AAA conjugates from a urine sample by means of an organic acid extraction protocol, without the concomitant extraction of native amino acids. Secondly, the N-AAA conjugates were hydrolysed to liberate BCAAs without the concomitant racemisation of the amino acids. Finally, BCAA enantiomers were separated via GC-MS to establish the enantiomeric composition of the starting N-AAA conjugates in the urine sample. In addition, the non-enzymatic formation of N-AAA conjugates was attempted *in vitro* and molecular modelling was used to determine the validity of the proposed reaction mechanism to gather insights on the requirement of a catalyst/enzyme.

Sample selection proved difficult due to the low incidence of metabolic disorders, in particular those that produce N-AAA conjugates. As discussed in section 3.4.2, only one sample (sample 11) with significant concentrations of N-AAA conjugates, could be used for this study. Since only a low sample volume of 3 mL was available and a minimum of 1 mL was required for the planned analysis, proper planning and rehearsal was needed to prevent unnecessary wastage of the limited material. For this purpose, 10 anonymised samples, left over from a previous project, were acquired and used as urine controls. In addition, N-AAA conjugate standards would have been purchased to standardise the various methods. Unfortunately, the exact N-AAA conjugates observed in sample 11 could not be acquired and a selection of the structural analogues, N-AcAA conjugate standards, was purchased instead.

Some of the methods used in this study are routinely used in-house (NWU Centre for Human Metabolomics, 2017) and was adopted as is. This included creatinine determination as well as amino acid and organic acid analysis. Other methods had to be standardised, which included the non-enzymatic synthesis of N-AAA conjugates from ammonia and 2-keto acids, hydrolysis of N-AAA conjugates, preparation of *R*-2-butanolic HCl (enantiopure and water-free) and a method to separate racemic amino acids via GC-MS.

4.2 Experimental outline

The experimental phase of this study was subdivided into the following five phases:

- Phase 1: Standardisation of new methods.
- Phase 2: Optimisation and rehearsal of the analytical strategy.
- Phase 3: Assessment of the non-enzymatic reaction between 2-keto acids and ammonia.

- Phase 4: Determining the enantiomeric composition of the N-AAA conjugates observed in sample 11.
- Phase 5: Molecular modelling to investigate the non-enzymatic formation of N-AAA conjugates.

An outline of the five phases is provided in sections 4.2.1 to 4.2.5 with detailed descriptions following in sections 4.3 to 4.11.

4.2.1 Phase 1: Standardisation of new methods

First, the three methods not routinely used in-house were standardised, namely:

(1) In relation to objective (1)(a), a method was developed to hydrolyse N-AAA conjugates without significant racemisation of the liberated BCAAs. Refer to section 4.7.

(2) In relation to objective (1)(b), a qualitative GC-MS method for the separation of racemic BCAAs was standardised. Refer to section 4.8.

(3) In relation to objective (1)(c), a duplication of the results obtained by Mienie *et al*. (2001:51-52) and Yanagawa *et al*. (1982:2089) was attempted: the non-enzymatic *in vitro* synthesis of N-AAA conjugates from 2-keto acids and ammonia. Refer to sections 2.3.3 and 4.9.

4.2.2 Phase 2: Optimisation and rehearsal of the analytical strategy

A summary of the approach that was followed for this phase is presented in Table 7. As outlined at the start of this chapter, the analytical strategy used in this study required the sequential use of various methods. Whether these methods could be used sequentially (e.g. to determine whether constituents left over from a previous method would interfere with the next method) therefore needed to be assessed. In addition, anonymised urine controls were needed to standardise and rehearse the analytical strategy. Therefore, the feasibility of using the urine controls for this purpose also needed to be ascertained:

(1) To prevent unexpected results when spiking the urine controls with N-AcAA conjugates, it needed to be ascertained whether any N-AcAA conjugates were natively present in the urine controls. For this, the standard in-house organic acid analysis was applied (refer to section 4.6).

(2) To prevent the observation of a pseudo-enantiomeric excess of BCAAs, it needed to be determined whether any other compounds in the controls would liberate BCAA enantiomers when applying the standardised hydrolysis method. This was achieved as

follows: An organic acid extraction was performed on the control samples, followed by the hydrolysis protocol on the dried organic acid extract. After drying the hydrolysis end product, the samples were derivatised with a chiral reagent and analysed via GC-MS.

(3) In addition to step 2 above, it was necessary to confirm that no free BCAAs were extracted by the organic acid extraction protocol. For this, the amino acid analysis was performed on reconstituted organic acid extracts of the urine controls. This was further tested by spiking the controls with significant amounts of BCAAs prior to organic acid extraction, as would be the case with MSUD.

(4) To enable the qualitative approach, it was necessary to determine whether any constituents of the urine controls would overlap with the peaks of BCAA enantiomers on the GC chromatograms. This was achieved by spiking the controls with N-Ac-BCAA conjugates followed by organic acid extraction, hydrolysis of the conjugates, conversion of the liberated BCAAs to diastereomers with a chiral derivatisation reagent and subsequent separation and identification of the diastereomeric derivatives via GC-MS. Results were compared to the non-spiked samples in step 2 above.

Table 7: Phase 2: Optimisation and rehearsal of the analytical strategy.

Determining whether N-AcAA conjugates are present in the urine controls	**Determining whether any compounds in urine controls would liberate BCAAs upon hydrolysis**	**Confirming that no free BCAAs are extracted during organic acid extraction**	**Determining whether constituents of urine controls would overlap with DL peaks of BCAA enantiomers on the gas chromatograms**
Not spiked with N-AcAA conjugates.	Not spiked with N-AcAA conjugates.	One set spiked with BCAAs, one set not spiked.	Spike with N-AcAA conjugates.
Organic acid extraction (4.6.1).	Organic acid extraction (4.6.1).	Organic acid extraction (4.6.1).	Organic acid extraction (4.6.1).
No hydrolysis.	Hydrolysis of the N-AcAA conjugates (4.7.2).	No hydrolysis.	Hydrolysis of the N-AcAA conjugates (4.7.2).
Standard organic acid derivatisation and GC-MS analysis (4.6.2 & 4.6.3).	Chiral derivatisation and separation of diastereomeric derivatives via GC-MS (4.8).	Standard amino acid analysis (4.5).	Chiral derivatisation and separation of diastereomeric derivatives via GC-MS (4.8).
Determine the presence of N-AcAA conjugates.	Determine whether BCAA enantiomers are present.	Determine whether BCAAs are present in significant concentrations.	Assess chromatograms for overlapping peaks.

Corresponding sections are indicated in brackets. GC-MS: gas chromatography-mass spectrometry; BCAA: branched chain amino acid; N-AcAA: N-acetylated amino acid.

4.2.3 Phase 3: Assessment of the non-enzymatic reaction between 2-keto acids and ammonia

A summary of the approach that was planned for this phase is presented in Table 8. During the standardisation of the method for the non-enzymatic synthesis of N-AAA conjugates (refer to phase 1, section 4.2.1) the standard organic acid method and GC-MS analysis were performed to determine yield. If a significant yield was to be obtained, the enantiomeric composition of the synthesised conjugates would be determined. This would be achieved by performing an organic acid extraction followed by hydrolysis and subsequent chiral derivatisation and GC-MS analysis. This would have shown whether the non-enzymatic reaction between 2-keto acids and ammonia would indeed yield racemic N-AAA conjugates.

Table 8: Phase 3: Assessment of the non-enzymatic reaction between 2-keto acids and ammonia.

Determining yield of non-enzymatic formation of N-AAA conjugates	**Determining enantiomeric composition of synthesised N-AAA conjugates**
Synthesis of various N-AAA conjugates from a selection of 2-ketoacids by the standardised method (4.9).	Synthesis of various N-AAA conjugates from a selection of 2-ketoacids by the standardised method (4.9).
Organic acid extraction (4.6.1).	Organic acid extraction (4.6.1).
No hydrolysis.	Hydrolysis of N-AcAA conjugates (4.7).
Standard organic acid derivatisation and GC-MS analysis (4.6.2 & 4.6.3).	Chiral derivatisation and separation of diastereomeric derivatives via GC-MS (4.8).
Determine whether adequate amounts of N-AAA conjugates were obtained.	Determine whether a racemic or enantiomeric excess was obtained.

Corresponding sections are indicated in brackets. GC-MS: gas chromatography-mass spectrometry; N-AcAA: N-acetylated amino acid; N-AAA: N-acylated amino acid.

4.2.4 Phase 4: Determining the enantiomeric composition of the N-AAA conjugates observed in sample 11

A summary of the approach that was followed for this phase is presented in Table 9. After standardising the non-standard methods and the optimisation and rehearsal of the analytical strategy, the enantiomeric composition of the N-AAA conjugates observed in sample 11 was determined. To confirm the presence of N-AAA conjugates, the standard organic acid method and GC-MS analysis were first performed. This was followed by an organic acid extraction, hydrolysis and subsequent chiral derivatisation and GC-MS analysis on a fresh aliquot.

Table 9: Phase 4: Determination of the enantiomeric composition of N-AAA conjugates observed in sample 11.

Confirming the presence of N-AAA conjugates	Determining the enantiomeric composition of N-AAA conjugates
Organic acid extraction (4.6).	Organic acid extraction (4.6.1).
No hydrolysis.	Hydrolysis of the N-AAA conjugates (4.7).
Standard organic acid derivatisation and GC-MS analysis (4.6.2 & 4.6.3).	Chiral derivatisation and separation of diastereomeric derivatives via GC-MS (4.8).
Determine whether adequate amounts of N-AAA conjugates were present.	Determine whether racemic or an enantiomeric excess was obtained.

Corresponding sections are indicated in brackets GC-MS: gas chromatography-mass spectrometry; N-AAA: N-acylated amino acid.

4.2.5 Phase 5: Molecular modelling to investigate the non-enzymatic formation of N-AAA conjugates

This section outlines the molecular modelling approach. For a detailed description of the method used, refer to section 4.11. Refer to Figure 19: The first step of the molecular modelling procedure was to interpret the ambiguous reaction mechanism proposed by Yanagawa *et al*. (1982:2089). Based on the chemical principles of this mechanism, a new mechanism was derived that would be applicable for the 2-keto acid substrates and N-AAA conjugate products investigated in this study (refer to section 2.3.3 and Figure 10). Next, the substrates, reaction intermediates and products derived from the new mechanism were drawn using the Visualizer module in the BIOVIA Materials Studio software (v16.1.0.31 Dassault Systèmes, Vélizy-Villacoublay, France). To simplify calculations while optimising the modelling approach, pyruvate as a 2-keto acid (focusing on the *R*-configuration) was selected, resulting in N-Ac-*R*-Ala as the end product. The drawn structures were roughly optimised using the built-in function of the software and subjected to geometric optimisation (refer to section 4.11 for details). Thereafter the optimised structures were assessed by visual inspection and vibrational analysis. If the optimised structure did not resemble the intended structure or if the structure constituted one or more imaginary vibrations, the structure's configuration was adjusted and the geometric optimisation procedure repeated.

Observations made during visual inspections and vibrational analyses were reapplied to the newly proposed reaction mechanism and the process was repeated. After a satisfactory geometrically optimised structure was obtained, an energy calculation was performed to determine the E_e of the optimised structure (refer to section 4.11.3). This value, together with

the correction factor obtained from the vibrational analysis, was used to determine the total ΔG of the molecule. Along with the energy calculation, the HOMO, LUMO and electrostatic potential (ESP) map were calculated to assist in the interpretation of the mechanism.

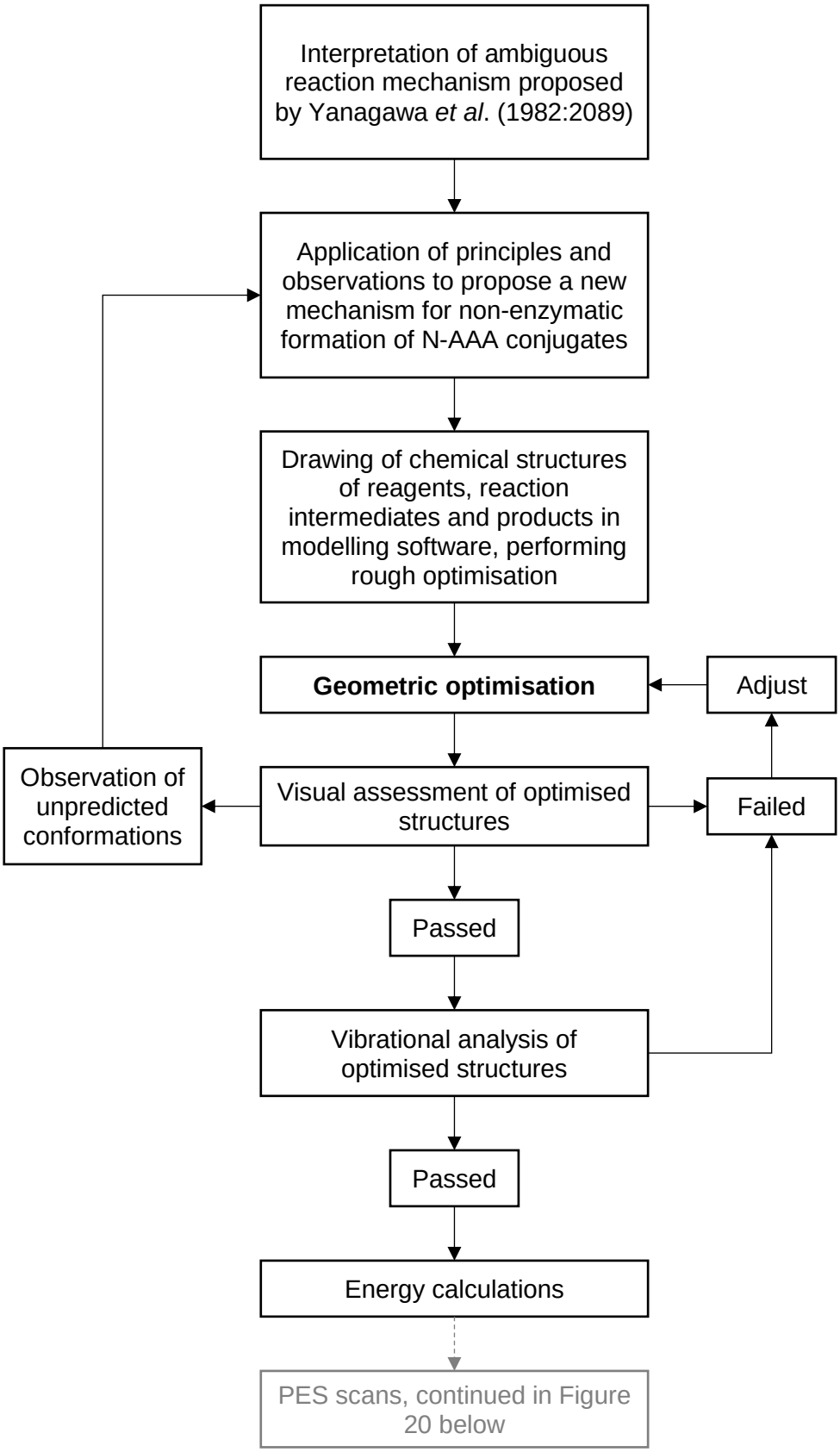

Figure 19: Phase 5-A: Procedure for geometric optimisation.

Refer to Figure 20: After all the reaction intermediates in the proposed reaction mechanism were geometrically optimised, TS searches were done on all the reaction steps by performing PES scans (Refer to section 4.11.4). During a PES scan procedure, incremental changes were made to the reaction coordinate (bond forming or breaking) with a geometric optimisation between each incremental change. These changes (conformational changes after each optimisation) were evaluated in the Visualizer module of the Materials Studio software. This proved useful for gathering information on how a reaction step may proceed and led to new insights that were applied to the proposed reaction mechanism. As a practical example, with each incremental lengthening in the N-H bond in intermediate #E (refer to Figure 25 in section 4.11), the geometric optimisation would bring the hydrogen closer to the hydroxyl group on the adjacent carbon. This resulted in a water molecule leaving, instead of just a hydrogen atom. A PES scan was deemed successful if the final molecule made sense in terms of the reaction mechanism and was stable (e.g. did not result in the decomposition of the starting molecule into three or more molecules).

After a successful PES scan, the resulting E_e of each optimised molecule in the PES scan was plotted against the reaction coordinate to obtain an energy profile. This energy profile was used to determine whether a TS was present in the reaction coordinate. As mentioned in section 2.7.1, if the first derivative of the energy profile equalled zero, a turning point in the energy curve was observed. If the second derivative of the energy profile was negative, a TS was present at this turning point. If no turning point was apparent, a bond was presumed to form/break without a TS. The TS(s), if present, was further geometrically optimised and the E_e was determined. Finally, all the calculated energy values were plotted against the reaction coordinate to obtain an energy profile of the total reaction mechanism. This profile was compared to the enzymatic reactions given in Table 6, to ascertain the feasibility of an *in vivo* reaction.

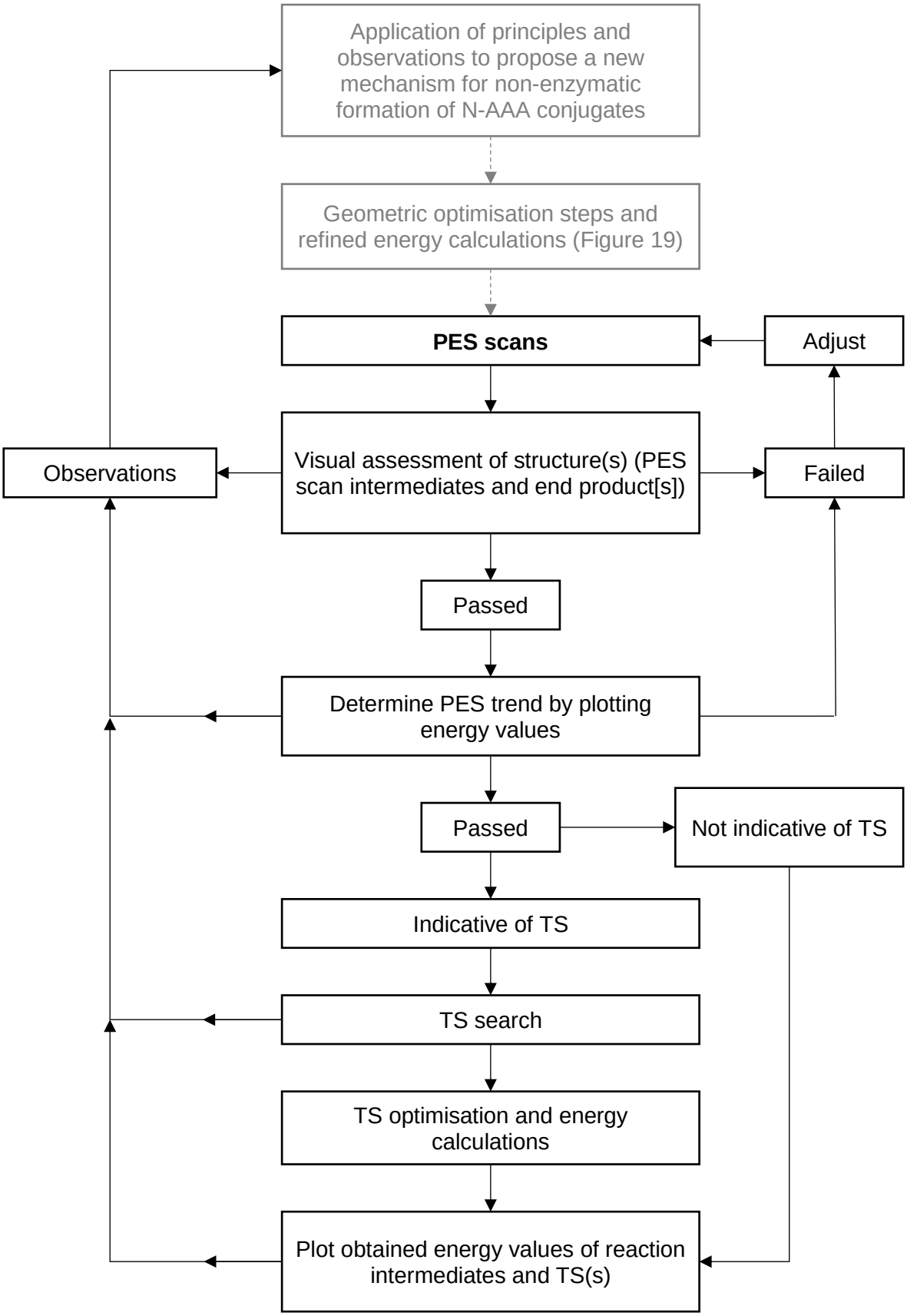

Figure 20: Phase 5-B: Procedure for potential energy surface (PES) scans.

N-AAA: N-acylamino acid; PES: potential energy surface; TS: transition state.

4.3 Chemicals and reagent preparation

The enzymatic creatinine determination kit and other reagents and consumables required for the Thermo Fisher Scientific (Waltham, Massachusetts, USA) Indiko clinical chemistry analyser type 863 were purchased from Thermo Fisher Scientific. The kit for amino acid analysis, 'Phenomenex EZ:faast for free (physiological) amino acid analysis by GC-MS' (EZ:faast), was purchased from Phenomenex (Torrance, California, USA).

Amino acid standards (DL-Ala, L-Ala, L-aIle, D-aIle, L-Leu, D-Leu, L-Val, D-Val, DL-norvaline [Nva], L-Ile and DL-Ile), internal standards (3-phenylbutyric acid and 4-methyl-DL-tryptophan), derivatisation reagents (N,O-bis(trimethylsilyl)trifluoroacetamide [BSTFA], trimethylchlorosilane [TMCS], MTBSTFA with 1% t-BDMCS and *R*-2-butanol), 2-ketoacids standards (2-keto-3-methylvaleric acid, 2-ketobutyric acid, 2-ketocaproic acid, 2-ketoisocaproic acid, 2-ketoisovaleric acid, 2-ketovaleric acid, phenylpyruvate, 4-ketophenylpyruvate and pyruvate), urea, ammonium acetate, magnesium sulphate, ammonium carbonate, ammonium bicarbonate, ammonium sulphate, monopotassium phosphate, dipotassium phosphate and α-TPP were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). N-AAA conjugate standards (N-isobutyryl-, N-isovaleryl- and N-2-methylbutyryl conjugates of Leu, Val and Ile) could not be purchased due to unavailability. Instead, the structural analogues N-AcAA conjugate standards (N-Ac-L-Ala, N-Ac-D-Ala, N-Ac-L-Val, N-Ac-D-Val, N-Ac-L-Leu, N-Ac-D-Leu, N-Ac-L-Ile and N-Ac-D-aIle) were acquired from Sigma-Aldrich.

HCl (37%), potassium hydroxide (KOH), sulphuric acid, glacial acetic acid, sodium hydroxide (NaOH), isooctane, pyridine, acetone, ethyl acetate, diethyl ether, *n*-hexane, anhydrous sodium sulphate and ammonium hydroxide (25%) were purchased from Merck Millipore (Darmstadt, Germany). Amino acid isotopes (universal [U]-^{13}C algal amino acid mixture [refer to Annexure B-Table 17], L-Gln-U^{13}C-U^{15}N, L-Asn-U^{13}C-U^{15}N and DL-cystine-2,2',3,3'-d_4) were purchased from Cambridge Isotope Laboratories (Andover, Massachusetts, USA). Nitrogen gas, for solvent evaporation, was generated with an NP60 nitrogen generator from Nitralife (Shrewsbury, Sandton, South Africa), coupled with a Jemaco AN SPX refrigerated air dryer from Artic Driers International (Benoni, South Africa). Instrument-grade helium (baseline 5.0) for use as a carrier gas in GC operations was purchased from African Oxygen Limited (Johannesburg, South Africa).

4.3.1 Amino acid internal standard mixture preparation

Some of the isotopes and standards contained in the amino acid internal standard mixture were not relevant to the study but are used during in-house routine analysis (NWU Centre for Human Metabolomics, 2017a). The mixture was therefore used as is. The preparation of the isotope mixture consisted of preparing a separate 1 mmol/L stock solution of each of the standards including 4-methyl-DL-tryptophan, L-Gln-U^{13}C-U^{15}N, L-Asn-U^{13}C-U^{15}N and DL-cystine-2,2',3,3'-d_4 by dissolving in double-distilled water (ddH_2O). Thereafter 25.66 mg U-^{13}C algal amino acid mixture was dissolved in 1 mL ddH_2O, aiding solubility with the addition of one drop of 5 N NaOH. A 1 mL aliquot of the algal amino acid mixture was transferred to a 50 mL volumetric flask and a 5 mL aliquot of each of the stock solutions of 4-methyl-DL-tryptophan, L-Gln-U^{13}C-U^{15}N, L-Asn-U^{13}C-U^{15}N and DL-cystine-2,2',3,3'-d_4 was added. The mixture was mixed and ddH_2O added up to the 50 mL mark of the volumetric flask.

4.3.2 Preparation of *R*-2-butanolic HCl

Moisture-free enantiopure *R*-2-butanolic HCl stock solution was prepared by slowly bubbling moisture-free HCl gas through *R*-2-butanol on ice for 30 min, with the setup as illustrated in Figure 21. Concentrated HCl(aq) (37%) was slowly added (one drop every 2 s) by means of an addition funnel to concentrated sulphuric acid in a glass flat-bottomed three-necked flask. The mixture was continuously mixed with a magnetic stirrer. The liberated HCl(g) was bubbled through additional sulphuric acid in a separate round-bottomed flask, connected with chemical-resistant tubing (Tygon formula 2075 laboratory tubing purchased from Sigma-Aldrich) and a vacuum adapter to further remove any moisture that may have evaporated from the exothermic mixture in the flat-bottomed three-necked flask. This was achieved by the positive pressure generated in the setup by the HCl(g). The HCl(g) was then bubbled through *R*-2-butanol with a glass pasture pipette connected with tubing, while the mixture was kept cold on ice. The HCl(g) dissolved in the *R*-2-butanol to produce moisture-free *R*-2-butanolic HCl.

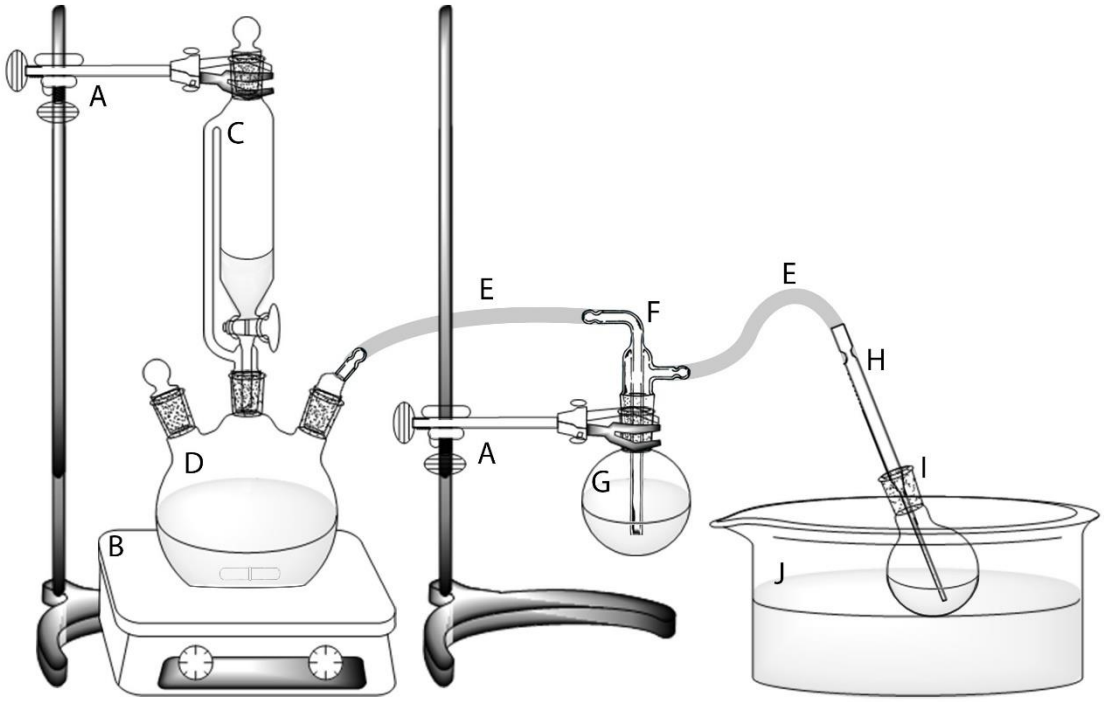

Figure 21: Setup used to prepare R-2-butanolic HCl.

(A) Burette stands and clamps, (B) magnetic stirrer and Teflon-coated stirring bar, (C) 100 mL glass addition funnel with glass stopper containing 37% HCl(aq), (D) 250 mL glass flat-bottomed three-necked flask with glass stopper and glass hose adapter, containing concentrated sulphuric acid, (E) chemical-resistant tubing, (F) modified glass long-stem vacuum adaptor, with the stem submerged in sulphuric acid, (G) 50 mL glass round-bottomed flask containing sulphuric acid, (H) glass Pasteur pipette, (I) 50 mL glass round-bottomed flask containing R-2-butanol, (J) large glass beaker containing ice.

The prepared ca. 6 mol/L *R*-2-butanolic HCl, *R*-2-butanol stock and a glass pipette were left to equalise in a walk-in-freezer at -20°C for 30 min. This was necessary since pipetting a 6 mol/L 2-butanolic HCl preparation at room temperature proved impossible due to vapour pressure. The prepared ca. 6 mol/L *R*-2-butanolic HCl was diluted to approximately 3 mol/L with *R*-2-butanol stock. Since the enantiopure *R*-2-butanol is relatively expensive and hard to acquire, the prepared *R*-2-butanolic HCl was not titrated to determine its concentration. Instead, the cheaper racemic *RS*-2-butanolic HCl was prepared in the same manner and titrated with 1 N NaOH until neutral pH was achieved as indicated by a universal pH indicator. *RS*-2-Butanol was found to be saturated with HCl after 30 min, at a concentration of approximately 6 mol/L. *R*-2-Butanolic HCl can also be prepared by the addition of Ac-chloride to *R*-2-butanol on ice. This will however, lead to residual Ac-chloride acetylating amino acids. Bubbling HCl(g) through *R*-2-butanol resulted in a purer preparation.

4.4 Creatinine analysis

Creatinine values were determined for all urine samples prior to organic and amino acid analysis. Creatinine was determined in-house via an enzymatic method on a Thermo Fisher Scientific Indiko Clinical Chemistry Analyser, type 863 (NWU Centre for Human Metabolomics, 2017b). A 500 µL aliquot urine sample was transferred to a sample cup and inserted into the analyser together with the reagent kit. Thereafter the assay was automatically performed by the analyser and reported in mmol/L.

4.4.1 Principle of the method

The creatinine analysis kit consisted of two reagent mixtures. Due to the proprietary nature of the kit, exact details on some of the constituents were not available. Reagent A consisted of 3-[[tris(hydroxymethyl)methyl]amino]propanesulfonic acid buffer at pH 8.1 and a concentration of 30 mmol/L. Also in the solution was creatinase, sarcosine oxidase and ascorbate oxidase with at least 333, 133 and 33 µkat/L activities respectively (enzymes isolated from unspecified microorganisms) and 2,4,6-triiodo-3-hydroxybenzoic acid, as well as unspecified detergents and preservatives. Reagent B also consisted of 3-[[tris(hydroxymethyl)methyl]amino]propanesulfonic acid buffer, but at pH 8.0 and a concentration of 50 mmol/L. Also in reagent B was creatininase from unspecified microorganisms with at least 500 µkat/L activity, peroxidase from horseradish (*Armoracia rusticana*) with at least 16.7 µkat/L activity, 4-aminophenazone at 2.0 mmol/L, potassium hexacyanoferrate(II) trihydrate at 163 µmol/L and an unspecified detergent and preservative.

The creatininase enzyme catalyses the hydrolysis of the creatinine in the sample to creatine. Thereafter, creatinase catalyses the hydrolysis of creatine to produce urea and sarcosine. Sarcosine oxidase catalyses the oxidative demethylation of sarcosine with oxygen as an electron acceptor, yielding Gly, formaldehyde and hydrogen peroxide. Finally, peroxidase catalyses the reduction of 4-aminophenazone and 2,4,6-triiodo-3-hydroxybenzoic acid in the reagent, with the hydrogen peroxide produced from previous steps as the electron donor. This final step yields water, hydrogen iodine and a quinineimine chromogen. The colour intensity of the chromogen is directly proportional to the concentration of creatinine and is measured photometrically at 540 nm. Interfering intermediates were removed by a pre-incubation step with reagent A.

4.4.2 Procedure

Creatinine was determined automatically by the analyser as follows: 120 µL of reagent A was pipetted to a micro cuvette together with 3 µL of a urine sample and diluted with 45 µL ddH_2O.

The mixture was agitated with a drill-like mixer and incubated at 37°C for 5 min. After incubation, an end-point blank measurement was taken at 540 nm and confirmed with a side-wavelength of 700 nm. This was followed by the addition of 60 µL of reagent B, mixing and incubation at 37°C for 5 min. Thereafter, an end-point measurement was taken at 540 nm and confirmed with a side-wavelength of 700 nm. If the calculated creatinine concentration was found to be below 0.2 mmol/L, the analyser repeated the procedure without dilution and corrected the final concentration by adjusting the dilution factor. If the creatinine value was found to be above 40 mmol/L, the analyser repeated the procedure with an increased dilution of 240 µL ddH_2O. This enabled an absorbance measurement in the linear range of the calibration curve, i.e. between 0.01 and 0.11 absorbance units. The test limits were set as a minimum of 0.01 and a maximum of 200.00 mmol/L. This method had an acceptable working range between 0.20 and 20.00 mmol/L.

4.5 Amino acid analysis (quantitative)

Please note that the quantitative amino acid analysis discussed below was not the method used to separate racemic amino acids (refer to section 4.8), which was only qualitative. It was instead used to accurately quantify amino acids with the aid of stable isotopes and not to separate enantiomers. The quantitative amino acid analysis method used was a modified method of the EZ:faast kit (validated in-house for diagnostic use) (NWU Centre for Human Metabolomics, 2017a). The four main in-house modifications to the EZ:faast method were:

(1) The addition of a mixture of amino acid isotopes and additional internal standards (refer to section 4.3.1) to improve the quantification of the method.

(2) The addition of NaOH or HCl to adjust the pH of the sample as required.

(3) The omission of the final step of the kit-procedure (drying of sample and reconstitution in reagent 6 – re-dissolution solvent).

(4) The selective modification of the GC parameters specified by the kit.

This modified method is the standard routine method used in-house. Due to the proprietary nature of the kit, the exact composition of most of the reagents and the composition and type of the SPE sorbent tips are unknown and referred to as ‘not otherwise specified’ (n.o.s.).

4.5.1 Sample preparation

First, the elution reagent was prepared fresh each day by combining three parts of reagent 3A (0.33 mol/L NaOH) with two parts of reagent 3B (80% n-propanol and 20% 3-picoline), vortex-

mixed and set aside. If the sample was dried by the final step of a previous method, it was first reconstituted in 1 mL ddH_2O with one drop of 0.1 mmol/L NaOH and vortexed. Subsequently, 100 µL of the sample was pipetted into a small reaction vial and 100 µL of the isotope mixture and 100 µL of reagent 1 (internal standard solution with 0.2 mmol/L Nva, 10% n-propanol and 20 mmol/L HCl) were added and vortex-mixed. The pH was determined with a universal indicator paper and adjusted (if necessary) to a pH of between 1.5 and 6.0 with 1 mol/L NaOH or 1 mol/L HCl as necessary. The sample was passed through the SPE sorbent tip provided with the kit with the provided syringes. Then, 200 µL of reagent 2 (washing solution, with 30% n-propanol, n.o.s.) was added to the same vial and passed through the resin with the syringe. The syringe was disconnected from the sorbent tip and the waste discarded. Thereafter, 200 µL of the elution reagent was added to the vial, passed through the sorbent tip with a clean syringe and ejected back into the vial, eluting the amino acids from the sorbent tip.

Fifty µL of reagent 4 (chloroform, propyl chloroformate, n.o.s.) was added to the sample and vortex-mixed for 8 s with the vortex mixer set to pulse mode and 80% of maximum speed. Subsequently, the sample was left to incubate at room temperature for 8 min. The sample was then vortex-mixed again for 8 s and left to incubate at room temperature for another 8 min, after which 100 µL of reagent 5 (80% isooctane and 20% chloroform) was added to the sample, vortex-mixed for 5 s and left again to incubate at room temperature for 8 min. Finally, the top organic layer was transferred to a GC vial with insert, cap and septa.

4.5.2 GC-MS conditions

The sample was injected on a Hewlett Packard (Palo Alto, California, USA) 6890 series GC system with an Agilent Technologies (Santa Clara, California, USA) 5973N MSD fitted with an electron ionisation (EI) ion source, an Agilent Technologies 7683 autosampler and a 7683B injector. The GC system was fitted with a Phenomenex GC FocusLiner liner for HP (split/splitless, w/wool, single taper, 4 mm internal diameter x 78.5 mm length x 6.3 mm outer diameter [part number AG0-4680]), a Phenomenex Zebron EZ-AAA amino acid GC column (10 m x 0.25 mm x 0.25 µm [part number CG0-7169]) and an Agilent Technologies gas clean filter system (part number CP17973).

Helium was used as a carrier gas at a constant flow rate of 1.3 mL/s. A 2 µL sample was injected in splitless mode with a viscosity delay of 1 s and the injector was set to 250°C. The oven was set to an initial temperature of 60°C for 1 min and subsequently increased to 110°C at 50°C/min, immediately followed by an increase to 185°C at 20°C/min, then to 235°C at 25°C/min, 320°C at 30°C/min and, finally, kept isothermal at 320°C for 1 min. The MSD's transfer line heater was set to 280°C, the source to 230°C and the quadrupole to 150°C.

Solvent delay was 3 min and the electron multiplier voltage (EMV) was set to a gain factor of 1.00. The MSD was set to scan from 40 to 500 m/z with a threshold of 150.

4.6 Organic acid analysis

The organic acid analysis method used was the standard routine method used in-house, which is a modified method from that described by Blau *et al.* (2008:140) (NWU Centre for Human Metabolomics, 2017c).

4.6.1 Organic acid extraction

The amount of urine used was based on the creatinine content (refer to section 4.4 for creatinine determination, Equation 2 for conversion from mmol/L to mg/dL and Table 10 for calculation of the sample amount).

Equation 2: Conversion of creatinine from mmol/L to mg/dL.

$$Creatinine\ value\ (mg/dL) = creatinine\ value\ (mmol/L)\ \times\ 11.312$$

Table 10: Calculation of sample amount for organic acid analysis.

Creatinine value (mg/dL)	Amount of sample used (mL)
Above 100.0	0.5
5.0–99.9	1.0
2.0–4.9	2.0
Below 2.0	3.0
Not determined (e.g. standards)	1.0

The calculated amount of sample was aliquoted to a test tube with screw cap and septa. The amount of internal standard (3.197 mmol/L 3-phenyl butyric acid) added to the sample was also calculated based on the individual sample's creatinine value (refer to Equation 3) and the amount of sample used.

Equation 3: Calculation of the amount of internal standard used for organic acid analysis.

$$Internal\ standard\ volume\ (\mu l) =\ creatinine\ value\ (mg/dL)\ \times sample\ volume\ (mL) \times 5$$

The calculated amount of internal standard was added to the sample. For samples where no creatinine value was available, 100 µL of internal standard was used. The sample's pH was

adjusted to less than 2 with 6 drops of 5 N HCl, after which it was vortexed and the pH checked with a universal indicator paper. More HCl was added if necessary, the sample was vortex-mixed and the pH re-measured. Thereafter, organic acids were extracted by adding 6 mL ethyl acetate to the sample. The test tube was capped, checked for leaks and mixed on a rotary mixer for 30 min at approximately 24 revolutions per minute, whereafter the sample was centrifuged for 3 min at a relative centrifugal force of 760 × g. The top organic layer was transferred to a clean test tube and the extraction steps were repeated with 3 mL diethyl ether. After centrifugation, the top organic layer was aspirated and added to the ethyl acetate layer previously extracted. Next, excessive water was removed from the ethyl acetate and diethyl ether sample mixture with one spatula (approximately 1 g) of sodium sulphate and the sample was vortexed and centrifuged for 3 min at 760 × g. The supernatant was subsequently carefully poured into a clean test tube and dried under a steady stream of nitrogen at 40°C for approximately 45 min.

4.6.2 Derivatisation of extract

The amount of derivatisation reagents used was also calculated based on the creatinine value and sample volume (refer to Equation 4).

Equation 4: Calculation of the amount of derivatisation reagents used for organic acid analysis.

$$BSTFA\ volume\ (\mu L) = creatinine\ value\ (mg/dL) \times sample\ volume\ (mL) \times 2.0$$
$$TMCS\ volume\ (\mu L) = creatinine\ value\ (mg/dL)\ \times sample\ volume\ (mL) \times 0.4$$
$$Pyridine\ volume\ (\mu L) = creatinine\ value\ (mg/dL) \times sample\ volume\ (mL)\ \times 0.4$$

For samples where no creatinine values were available, 100 µL BSTFA, 20 µL TMCS and 20 µL pyridine was used. The calculated amount of derivatisation reagents was added to the dried sample. The samples were then vortexed and incubated at 75°C for 45 min, whereafter it was transferred to a GC vial with insert, cap and septa.

4.6.3 GC-MS conditions

The samples were injected on an Agilent Technologies 7890A GC system with a 5975C VL MSD with a triple-axis detector fitted with an EI ion source and a 7693 automatic liquid sampler. The GC system was fitted with an Agilent Technologies GC split liner (single taper, deactivated glass wool, low pressure drop, 4 mm internal diameter, 870 µL [part number 5183-4701]), an Agilent Technologies J&W GC column (DB-1MS ultra inert, 30 m x 0.25 mm x 0.25 µm capillary column [part number 122-0132UI]) and an Agilent Technologies gas clean filter system (part number CP17973).

Helium was used as a carrier gas at a constant column flow rate of 0.945 mL/s. One µL sample was injected with a viscosity delay of 2 s in split mode with a split ratio of 1:12, a split flow of 11.345 mL/min and the injector set to 280°C. The oven was set to an initial temperature of 50°C for 1 min and then increased to 60°C at 20°C/min, immediately followed by an increase to 120°C at 5°C/min and, finally, to 280°C at 7°C/min, where it was kept isothermal for 4 min. The post-run temperature was 300°C for 1 min. The MSD's transfer line heater was set to 280°C, the source to 230°C and the quadrupole to 150°C. Solvent delay was set to 7 min, EMV to a gain factor of 2.00 and the MSD to scan from 50 to 600 m/z with a threshold of 15.

4.7 Standardisation of hydrolysis of N-AAA conjugates

Refer to Figure 22: Enantiopure N-Ac-L-Ala (not N-Ac-Gly) was selected to standardise the hydrolysis method, since it is the simplest N-AAA conjugate with a chiral centre. Several acids (HCl, glacial acetic acid, sulphuric acid and formic acid) and bases (NaOH, ammonium hydroxide and KOH) were tested at various concentrations (0.1, 0.5, 2.0, 5.0 and 10 mol/L), temperatures (room temperature, 60°C and 110°C) and incubation times (no incubation, 10 min, 1 h and overnight) on N-Ac-L-Ala. An amino acid analysis (refer to section 4.5) was performed on all samples prior to hydrolysis to serve as a baseline and to confirm that no amino acids were present in the preparation. An organic acid analysis (refer to section 4.6) was also performed prior to hydrolysis to confirm the presence of N-Ac-L-Ala conjugates in the preparation. After hydrolysis, amino acid and organic acid analyses were repeated to measure the effectiveness of the hydrolysis method. After successful hydrolysis, the amino acid analysis results showed sufficient amounts of liberated amino acids and organic acid analysis showed a very low concentration or zero amounts of N-AcAA conjugates; a provisional method was subsequently selected. The hydrolysate was then treated with a chiral derivatisation reagent (as described in section 4.7.2) and the diastereomeric derivatives separated via GC-MS to determine the extent of hydrolysis-induced racemisation.

The provisional method was tested on individual N-AcAA conjugates and then on a mixture of the full range of N-AcAA conjugates available to the study (refer to section 4.3). Thereafter, the method was tested on the 10 urine controls. An organic acid analysis was performed on the controls prior to hydrolysis to confirm that no N-AAA conjugates were present that could interfere with the results. The controls were divided into two aliquots, one set of which was spiked with a mixture of N-AcAA, while the other was not. The spiked set was used to determine whether any other compounds in urine interfered with the efficiency of hydrolysis and the non-spiked set was used to determine whether any other compounds present in the controls would liberate amino acids during hydrolysis. Refer to section 4.2.2 and Table 7.

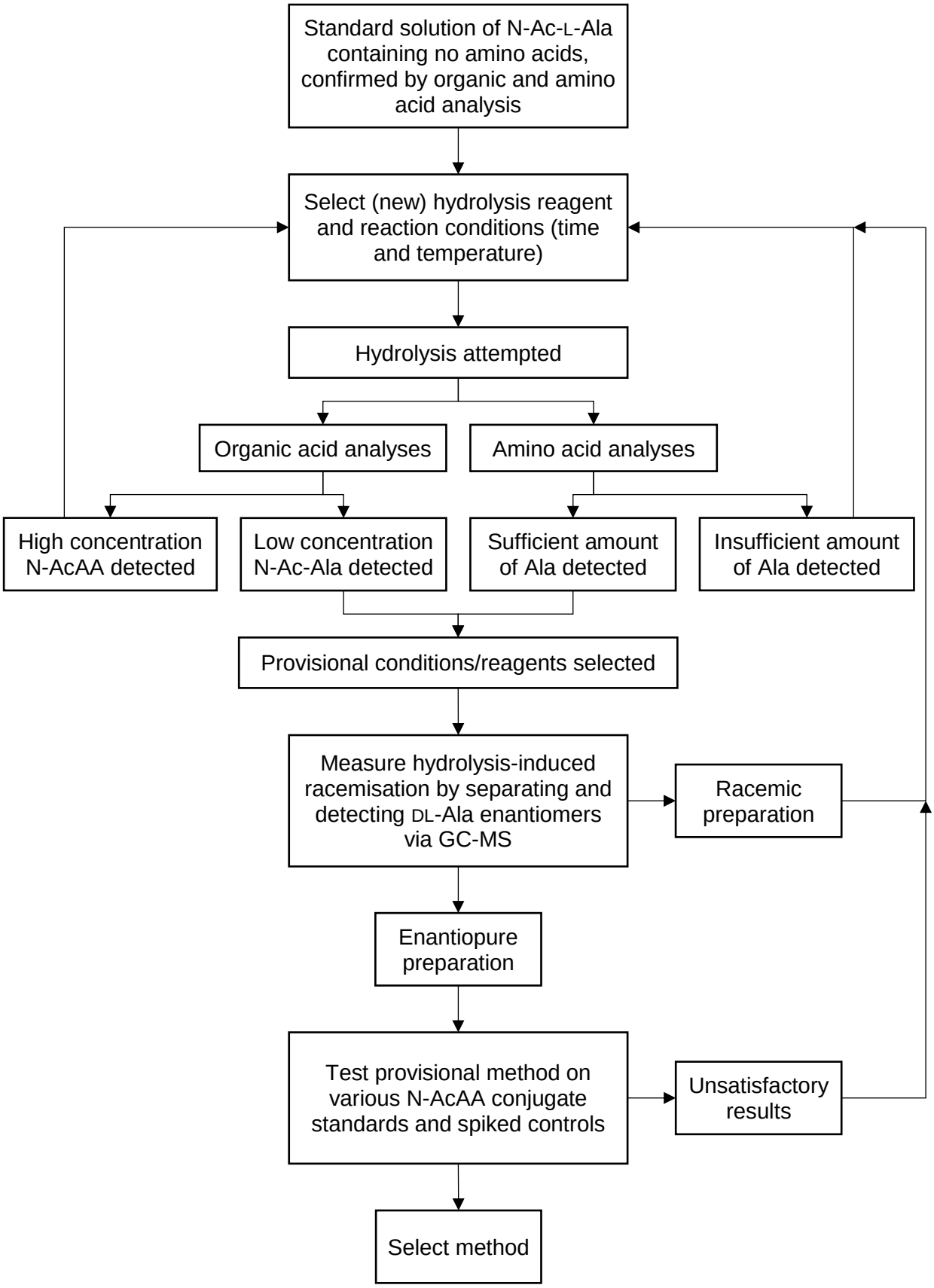

Figure 22: Procedure for standardising hydrolysis of N-AAA conjugates.

4.7.1 Provisional hydrolysis procedure

The stock solutions of the N-AcAA conjugate standards were prepared separately by dissolving 100 mg standard in 10 mL ddH_2O. A working solution was prepared by performing a 10-fold dilution with ddH_2O. A working mixture of the conjugates was prepared by adding 1 mL of the respective stock solutions to a 10 mL volumetric flask and adding ddH_2O up to the 10 mL mark. The spiked urine control samples were prepared as follows: 100 µL working mixture, containing 1 mg/mL of each of the N-AcAA conjugate standards, was transferred to a vial and dried under a steady stream of nitrogen at 40°C. The residue was reconstituted in 1 mL urine from a urine control. Thereafter an organic acid extraction was performed, followed by hydrolysis on the dried extract.

For standards, a 100 µL sample containing 1 mg of a N-AcAA conjugate standard (or 1 mg each for a mixture of standards) was transferred to a small glass test tube and dried under a steady stream of nitrogen at 40°C. Thereafter, 2 mL of the selected hydrolysis reagent was added to the vial and vortexed. The vial was capped and left to incubate at the selected incubation temperature and time. After incubation, the test tube was centrifuged briefly to collect the condensation on the sides of the test tube and in the cap, whereafter is was dried under a steady stream of nitrogen at 40°C for approximately 4–6 h. The dried hydrolysate was reconstituted in 1 mL ddH_2O, a 100 µL aliquot was removed and set aside for amino acid analysis and the rest was used for organic acid analysis.

4.7.2 Final selected method for hydrolysis of N-AAA conjugates

The final standardised method for the hydrolysis of N-AAA conjugates was as follows: The dried organic acid extract of a sample or standard containing approximately 100 µg N-AAA conjugates was reconstituted in 2 mL 2.0 N HCl and incubated for 1 h at 110°C. The test tube was briefly centrifuged and the hydrolysate dried under a steady stream of nitrogen at 40°C for approximately 4–6 h.

4.8 Separation of racemic amino acids on GC-MS (qualitative)

GC was selected as the analytical instrument of choice because of its inherent high resolution, relatively short analytical time, dynamic range and ability to be coupled with an MSD (Bertrand *et al.*, 2008:131). Derivatisation with a chiral reagent instead of a chiral column was pursued for reasons described in section 3.4.1. The method selected was a modified method based on that described by Bremer *et al*. (1981:483). Since S-3-methyl-2-butanol and pentafluoropropionic anhydride used by Bremer were not available at the time when this study was initiated, it was substituted with the enantiopure *R*-2-butanol and MTBSTFA with 1% t-BDMCS, respectively.

These reagents were selected based on reasons described in section 3.4.1. To increase reactivity, *R*-2-butanolic HCl was prepared by bubbling HCl(g) through *R*-2-butanol as described in section 4.3.2. The enantiopure *R*-2-butanolic HCl added an additional chiral centre to the racemic amino acid analyte, creating diastereomers, which enables separation via GC, while MTBSTA aided separation, increased stability and reduced evaporation temperatures to enable analysis via GC as described in section 2.6.2.

Enantiopure DL-Val, D-Val and L-Val were selected to standardise the separation of racemic amino acids. Different concentrations of *R*-2-butanolic HCl (1.5, 3.0 and 6.0 mol/L) were tested at different incubation temperatures (70 and 110°C) and incubation times (30 min, 90 min, 2 h and overnight). Samples were subsequently derivatised with MTBSTFA with 1% t-BDMCS and analysed via GC-MS. Thereafter, the method was tested on several racemic amino acids (refer to section 4.3), hydrolysate of N-AcAA conjugate standards and finally on hydrolysate of the urine controls (with and without spiking with N-AcAA conjugate standards – refer to section 4.2.2).

4.8.1 Sample preparation

After standardisation, the following procedure was selected and subsequently used on all other samples: A sample containing approximately 100 µg amino acid was aliquoted to a test tube with a screw cap and septa. If the sample contained HCl (from hydrolysis), it was first evaporated to dryness under a steady stream of nitrogen at 40°C for approximately 4–6 h, otherwise the sample was lyophilised overnight. Subsequently, 150 µL (3 mol/L) *R*-2-butanolic HCl was added to the dried sample and the test tube was capped, vortexed and incubated at 110°C for 2 h. The sample was then centrifuged briefly to collect condensation from the sides and inside the cap. The sample was evaporated to dryness under a steady stream of nitrogen at 40°C for approximately 30 min. Next, 100 µL MTBSTFA with 1% t-BDMCS and 50 µL pyridine were added to the sample, the test tube capped and the sample was vortexed and incubated at 75°C for 30 min. The sample was then transferred to a GC vial with insert, cap and septa and analysed via GC-MS.

4.8.2 GC-MS conditions

The sample was injected on an Agilent Technologies 7890A GC system with 5977A MSD (EI ion source) as detector and a 7693 automatic liquid sampler. The GC system setup further consisted of an Agilent Technologies GC split liner (single taper, deactivated glass wool, low pressure drop, 4 mm internal diameter, 870 µL [part number 5183-4701]), an Agilent

Technologies J&W GC Column CP-Sil 19CB (30 m x 0.32 mm x 0.25 µm capillary column [part number CP8842]) and an Agilent Technologies gas clean filter system (part number CP17973).

Helium was used as a carrier gas at a constant column flow rate of 1.4 mL/s. One µL of the sample was injected with a viscosity delay of 1 s in split mode with a split ratio of 1:12 and split flow of 16.8 mL/min. The injector was set to 280°C and the oven was set to an initial temperature of 70°C for 3 min and was then increased to 120°C at 5°C/min. The oven temperature was further set to increase at a near isothermal rate of 2°C/min to 140°C, until all D/L amino acids of interest eluted. The oven temperature was finally set to increase at 30°C/min to 280°C and was kept isothermal for 2.5 min. Post-run was at 300°C for 1 min. The MSD's transfer line heater was set to 280°C, the source to 230°C and the quadrupole to 150°C. Solvent delay was 8 min and EMV was set to a gain factor of 1.00. The MSD was set to scan from 70 to 600 m/z with a threshold of 50 and a scan speed of 1.562 Da/s.

4.9 Non-enzymatic *in vitro* synthesis of N-AAA conjugates

In relation to objective (1)(c), a duplication of the results obtained by Mienie *et al*. (2001:51-52) and Yanagawa *et al*. (1982:2089) was attempted. In addition, reaction conditions were selected that would simulate realistic reaction conditions expected *in vivo* and/or during sample handling. For this the following criteria were pursued:

1. The reaction conditions should be comparable to physiological or pathological conditions, including any part of inter- or extracellular physiology, and body fluids (e.g. urine or the intermembrane space of mitochondria).

2. If (1) cannot be achieved, the conditions should at least match that of a urine sample during transport, storage, handling and/or during organic acid analysis sample preparation, taking into account that in worst-case scenarios a sample may be exposed to at/above room temperature for as long as a weekend.

3. The N-AAA conjugates should be prepared from 2-keto acids and ammonia (as per the hypothesis described by Mienie *et al.* [2001:51-52]) and not from N-acyl-chloride and amino acids or with any other methods.

An exact duplication of the results reported by Mienie *et al.* (2001:51-52) was not achievable, possibly due to the fact that not all reaction conditions and reagent concentrations were reported. Thus, a modified method based on that reported by Mienie *et al.* (2001:51-52) was standardised. Several 2-keto acids (refer to section 4.3) were individually incubated with various sources of ammonia, including ammonium hydroxide, urea, ammonium acetate, ammonium

carbonate, ammonium bicarbonate and ammonium sulphate. Different pH values (<2.0, 3.5, 7.0 and >13.0), temperatures (room temperature, 37°C and 110°C), incubation times (no incubation, 1, 24 and approximately 65 h) and sample matrixes (a normal urine control and ddH_2O) were attempted.

A single 2-keto acid (reacting with another molecule of the same 2-keto acid) would result in a single product, serving as both the formyl- and carboxy-donor. By using the proposed net reaction given in Figure 23, the predicted N-AAA conjugate end product from a single 2-keto acid substrate was derived as shown in Table 11. Refer to Figure 24: Two different 2-keto acids in the reaction mixture could result in four possible products; each could possibly donate a formyl- and a carboxy group, which could also react with itself, e.g. a reaction between pyruvate and 2-ketoisovaleric acid could result in the formation of N-Ac-Ala, N-isobutyryl-Val, N-Ac-Val and N-isobutyryl-Ala. Three substrates could result in nine possible end products, four in sixteen, and so forth. Therefore, it is clear that the possible number of end product combinations can be calculated as: n^2, where n equals the number of 2-keto acids in the reaction, i.e. the number of substrates squared. This means that with the nine 2-keto acids selected for this study, 81 end products could be possible. To complicate matters further, each N-AAA conjugate end product could produce two peaks, with different mass spectra (due to incomplete derivatisation), after derivatisation with BSTFA and separation via GC-MS. Mass-spectra were only available for 27 of the 162 possible compounds (refer to Annexure C-Table 18 for a summary of sources). For this reason, the non-enzymatic reactions were limited to a single 2-keto acid substrate per reaction.

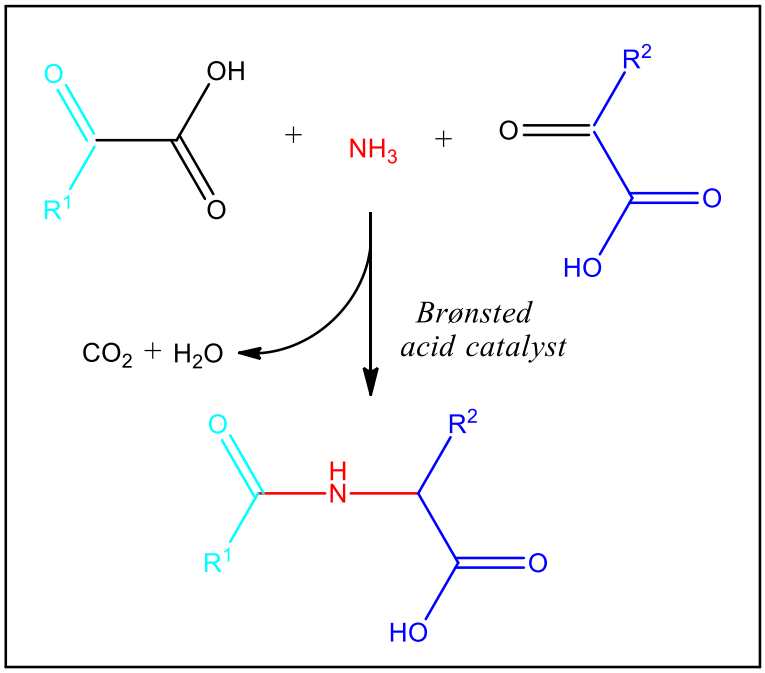

Figure 23: Net reaction for the formation of N-AAA conjugates from 2-keto acids and ammonia.

The net reaction shown above was used during the standardisation of the non-enzymatic in vitro *synthesis of N-AAA conjugates to predict which N-AAA conjugates would form from a set of different combinations of 2-keto acid substrates. R^1 and/or R^2 may refer to a methyl, isopropyl, 2-butyl or isobutyl side chain.*

Table 11: Expected N-AAA conjugate products from a single substrate reaction.

2-Keto acid	Expected result
2-Keto-3-methylvaleric acid	N-2-Methylbutyryl-Ile
2-Ketobutyric acid	N-Propionyl-2-aminobutyric acid
2-Ketocaproic acid	N-Pentanoyl-*nor*-Leu
2-Ketoisocaproic acid	N-Isovaleryl-Leu
2-Ketoisovaleric acid	N-Isobutyryl-Val
2-Ketovaleric acid	N-Butyryl-Nva
Phenylpyruvate	N-(Phenylacetyl)-Phe
p-Hydroxyphenylpyruvate	N-(*p*-hydroxyphenylacetyl)-(*p*-hydroxy-Phe)
Pyruvate	N-Ac-Ala

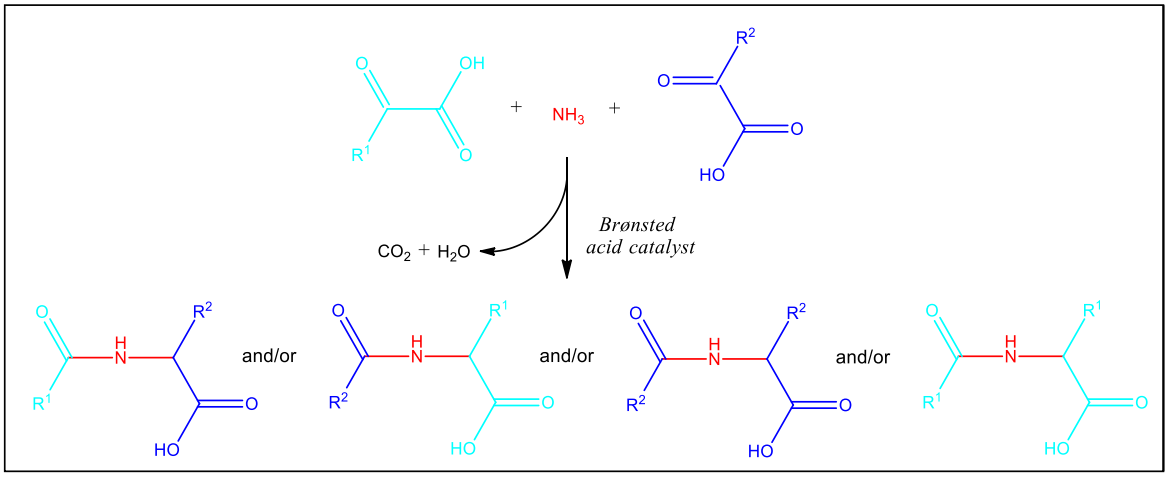

Figure 24: Possible products resulting from the reaction between two 2-keto acids and ammonia.

In the figure above, R^1 and/or R^2 may refer to a methyl, isopropyl, 2-butyl or isobutyl side chain.

4.10 Data mining

AMDIS version 2.71 from NIST was used for the deconvolution, identification and quantification of all data acquired from GC-MS analysis in this study. Identification of compounds was aided by an in-house library. This library was updated with spectra obtained from a previous study (Dry, 1997:51-52,70,77-79,89-94), data acquired from analysis of pure standards and the NIST Mass Spectral Search Program for the NIST/EPA/NIH mass spectral library, version 2.0f. The Agilent Technologies MSD Enhanced ChemStation F.01.00.1903 was also used for identification and quantification of GC-MS data. Furthermore, ChemStation was used for overlaying chromatograms for visual comparison.

4.11 Molecular modelling

In this study, molecular modelling was used to determine the E_a required for the non-enzymatic formation of N-AAA conjugates and to assess whether the energy requirements were achievable in a physiological and/or laboratory environment (during normal sample handling and organic acid analysis). The resulting E_a values were also compared to typical enzymatic E_a values (refer to Table 6). During the course of the modelling process, the newly proposed mechanism (refer to Figure 25) was assessed to determine the feasibility of the reaction.

4.11.1 Reaction mechanism

The reaction mechanism proposed by Yanagawa *et al*. (1982:2089) (refer to Figure 10) was ambiguous and incomplete for various reasons: Firstly, the mechanism by which the first 2-keto acid substrate is protonated to form an oxonium ion was not indicated. Furthermore, after the nucleophilic addition of ammonia, it is not shown how the amine cation is deprotonated to form the corresponding amine. The second oxonium ion (protonated from a 2-keto acid) is shown to be protonated on the carbonyl oxygen but was not notated with a formal charge. The mechanism for the subsequent nucleophilic addition between the amine and the second oxonium ion was not indicated. In addition, the number of protons from the amine and second oxonium ion does not compare with the total number of protons on the secondary amine cation conjugate. Yanagawa does, however, suggest the subsequent protonation of the conjugate, but this seems unnecessary given the latter. In the steps thereafter, it is unclear whether the curved arrows used to show electron movement indicate to/from atoms or bonds. The hydroxyl group on the conjugate is shown to leave the molecule without previously being protonated; this seems unlikely since a hydroxyl is not a good leaving group. In addition, the flow of electrons is incomplete for the reaction intermediate undergoing decarboxylation. It is improbable for the electrons resulting from the decarboxylation to be stabilised between the carbons of the opposite carboxyl group; instead it would be expected to be stabilised by the more electronegative nitrogen or oxygen atoms via the formation of a secondary ketamine or an alkoxide anion, respectively.

Despite these shortcomings, the reaction intermediates proposed by Yanagawa's reaction mechanism, together with known principles of organic chemical reactivity, enabled a preliminary reaction mechanism to be derived as shown in Figure 25. This preliminary mechanism served as a starting point for molecular modelling and was constantly updated as new insights were acquired from observations made during the modelling procedures. Furthermore, the preliminary reaction mechanism was generalised (replacing $-CH_3$ with R^1 and R^2) to accommodate other 2-keto acid substrates and N-AAA conjugated products. Thereafter the reaction equation was balanced in terms of molarity and charge (refer to Equation 5).

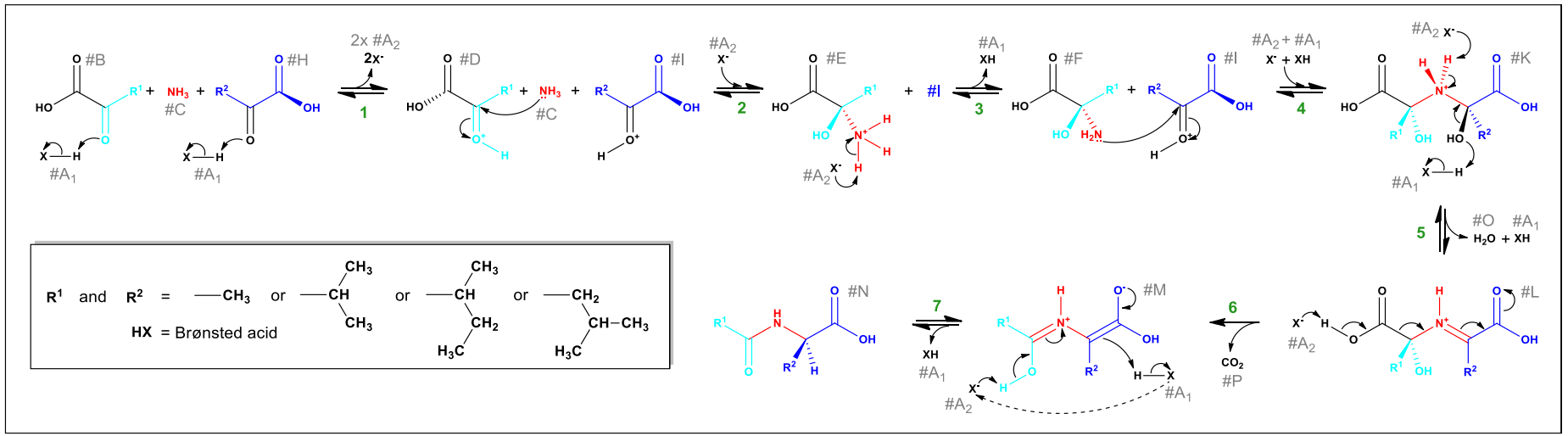

Figure 25: Preliminary reaction mechanism used to initiate the molecular modelling process.

In Figure 25, R^1 and/or R^2 may refer to a methyl, isopropyl, 2-butyl or isobutyl side chain. For example, replacing R^1 with a methyl group and R^2 with an isopropyl group produces the proposed reaction mechanism for the non-enzymatic formation of N-Ac-Val from pyruvic acid and 2-ketoisovaleric acid. The red, light-blue and dark-blue coloured groups correlate to the net reaction shown in Figure 23. Reaction steps are numbered in green, whereas reaction intermediates are labelled with grey letters preceded by a hash symbol. Missing letters, such as #G and #J, were reserved for possible expansion of the proposed mechanism. The symbol 'HX' is used to indicate a Brønsted acid. The main adjustments made to Yanagawa's reaction mechanism (at this stage) included (1) the introduction of a Brønsted acid catalyst to facilitate the protonation of the 2-ketoacid substrates and the hydroxyl-leaving group as well as the deprotonation of the carboxylic acid prior to decarboxylation, (2) a comprehensive set of curved arrows indicating electron movement and (3) the correction of hydrogen and charge balance.

Equation 5: Preliminary balanced reaction for 2-keto acids with ammonia.

$$R^1C_2H_1O_3 + R^2C_2H_1O_3 + H_3N + 2HX \rightleftharpoons R^1C_2H_2O_3^+ + R^2C_2H_2O_3^+ + NH_3 + 2X^-$$
$$\rightleftharpoons R^1C_2H_5NO_3^+ + R^2C_2H_2O_3^+ + 2X^- \rightleftharpoons R^1C_2H_4NO_3 + R^2C_2H_2O_3^+ + HX + X^-$$
$$\rightleftharpoons R^1R^2C_4H_6NO_6^+ + HX + X^- \rightleftharpoons R^1R^2C_4H_4NO_5^+ + H_2O + HX + X^-$$
$$\rightleftharpoons R^1R^2C_3H_3NO_3^+ + CO_2 + H_2O + 2HX \rightleftharpoons R^1R^2C_3H_3NO_3 + CO_2 + H_2O + 2HX$$

4.11.2 Geometric optimisation

The proposed reactants, reaction intermediates and product structures (Figure 25) were drawn using the Visualizer module in the Materials Studio software. These structures were roughly optimised (by using the software's built-in 'auto-clean' function) to find low-energy structures that made chemical sense. The structures were further geometrically optimised using the DFT module (DMol^3) in Materials Studio. All calculations (including those in the sections below) were performed on the Lengau Cluster at the Centre for High Performance Computing (CSIR, Rosebank, Cape Town). The generalised gradient approximation (GGA), PW91 functional with Ortmann, Bechstedt and Schmidt (OBS) dispersion correction and double-numerical basis set with polarisation functions (DNP) basis file 4.4 was selected as a calculation model. It is important to note that selecting a calculation model is a compromise between accuracy and calculation time. GGA/PW91 with OBS correction and DNP 4.4 provided acceptable accuracy, calculation time and model fit for the current study. Refer to Annexure E for published crystallography data for pyruvate, compared to results obtained for the current calculation model. Electron spin was set to unrestricted and formal spin was used as an initial value. Charge was set separately for each molecule depending on its formal charge as depicted in the preliminary reaction mechanism in Figure 25.

Convergence tolerance was set to medium, which relates to a maximum energy change of 2 x 10^{-5} Hartree (Ha), a maximum force of 0.004 Ha/Å and a maximum displacement of 0.005 Å for each iteration during the geometric optimisation. The maximum number of iterations (when calculation would stop even if the convergence criteria were not met) was set to 1 000. The maximum allowable step size for any change in Cartesian coordinates was 0.3 Å. Integration accuracy was set to medium, relating to a thousand grid points for each atom in the calculations. Direct inversion in interactive subspace (DIIS) was used to speed up self-consistent field (SCF) convergence. The SCF tolerance (which sets the threshold for density convergence) was set to medium, which relates to a value of 1 x 10^{-5}. If the largest component of the DIIS density error matrix was less than this value, the SCF procedure was considered converged. DIIS size was set to 6. The maximum number of SCF iterations was set to 1 000.

The maximum angular momentum function (multipolar expansion) was set to octupole. Density mixing charge and spin were set to 0.2 and 0.5, respectively. To speed up convergence, thermal smearing with a value of 0.005 Ha was selected. All electrons were included in the calculations. The orbital cut-off of the atomic basis set was set to a quality of medium and the exact value was dependant on the elements present in the current structure. The conductor-like screening model (COSMO) continuum solvation model (CSM) was selected to simulate a solvent environment for the modelling. Water was selected as solvent with the dielectric constant set to 78.54. Vibrational frequency analysis was performed as part of the DMol^3 calculations. Coarse-grained parallelisation was used for the numerical displacement frequency calculations. After convergence of the DMol^3 calculations, the vibration analysis tool was used to determine whether any imaginary vibrations were present.

The above-mentioned calculations provided the E_e of a rigid structure, i.e. not taking vibrations into account. By performing frequency calculation, ΔG corrections were calculated, which take vibrations into account at specific temperature intervals. The ΔG corrections showed a polynomial relationship with temperature. From this, polynomial regression was used to derive the ΔG correction at 37°C. The total relative Gibbs free energy (ΔG_{rel}) at 37°C was calculated as shown in Equation 6. The ΔG_{rel} of the geometrically optimised structures was used to construct a preliminary reaction energy profile.

Equation 6: Calculation of ΔG at 37°C.

$$\Delta G_{rel} = E_e + \Delta G_{37}$$

After the structures were geometrically optimised, a single-point energy calculation was performed to determine the molecules' relative E_e. The calculated value was based on a zero value for a system where all the electrons and nuclei are separated to infinity. The calculated value was generally a large negative value and should not be confused with binding energy (which is the energy required to separate individual atoms). To simplify the calculated values to a more understandable system, the E_e of all the proposed reactant, reaction intermediate and product structures were normalised relative to one molecule (the starting 2-keto acid reactant). To be able to plot the relative E_e change between each step of the reaction and TS(s), the total relative E_e of all the reacting species needed to be considered at each step in the reaction (i.e. all reactants in the reaction pool). In order to do this, the reaction needed to be balanced, paying special attention to electron balance, since E_e was measured (refer to Equation 5).

4.11.3 Energy calculations

For the energy calculations, the DMol^3 setup was kept identical to that of the geometric optimisation, with the exceptions that convergence tolerance was not applicable and vibrational frequency analysis was not repeated. In addition to the energy calculations, the following properties were calculated only to assist in the interpretation of the mechanism and was therefore not reported in the results: total electron density, ESP, Fukui functions (nucleophilic and electrophilic), orbitals and population analysis. Orbitals included the HOMO, LUMO and two additional levels above and below the Femi level. Population analysis consisted of Mulliken analysis set to atomic charge, Hirshfeld analysis set to charge and ESP charges were selected. All other properties were calculated with grid resolution set to medium, which correlates to a grid interval of 0.25 Å and border set to 3.0 Å.

4.11.4 TS searches with PES scans

TS structures between two reaction intermediates were determined by constraining one property on the starting structure (usually a bond length or proximity between two atoms) and increasing/decreasing the value of the property with a set number of increments and incremental sizes. This would result in a bond forming or breaking, thereby altering the starting structure until it resembles the final structure of the second intermediate. The in-between structures were geometrically optimised after each incremental change, keeping the constrained property constant and the resulting E_e was determined. The E_e values were plotted against the incremental changes in the constrained property's value, i.e. the reaction coordinate. The trend of the energy profile was used to determine whether a TS was involved in the current reaction step. A TS was expected at a point in the profile where the first derivative of the function for E_e to bond length approached zero (where the curve flattened) and where the second derivative was a negative value (curve sloped down to both sides). If the energy profile resembled a Morse potential, such as that in Figure 16, it was assumed that no TS was present and that a bond simply broke or formed without additional energy requirements beyond that of the reaction intermediates in question. The 'calculate bonds' tool in Materials Studio was used to determine the point where bonds were formed or broken.

Calculations were set up via a manual script file, which was subsequently submitted for calculations (refer to Annexure D for an example). Similar to the setup for geometric optimisations, DFT with GGA, PW91 functional with OBS dispersion correction and DNP basis file 4.4 was selected as a calculation model. Electron spin was set to unrestricted and charge was set for the initial structure as depicted in the preliminary reaction mechanism in Figure 25. The orbital cut-off of the atomic basis set was set to 3.7. The COSMO CSM was set to 78.54 for

water, the maximum number of iterations was set to 1 000 and the quality was set to fine. To speed up convergence, thermal smearing was selected with a value of 0.005 Ha.

The distance between the two atoms in question was measured and constrained, together with any other properties that needed to remain constant during the geometric optimisation steps (e.g. bond angle). Based on the measured distance, initial and final values were calculated for the constrained bond length. These parameters were selected beyond the measured and expected values; this approach was applied to ensure that the resulting energy profile would be complete. The distance of the step size was calculated based on the initial and final distances, ensuring 50 increments for each PES scan (refer to Equation 7).

Equation 7: Calculating distance step size for PES scans.

$$Step\ size\ = \frac{final\ distance - initial\ distance}{50}$$

The atoms were chosen based on the bonds to be broken or formed as depicted in the proposed reaction mechanism (refer to Figure 25). The initial and final parameters were selected based on the following criteria:

(1) To form a bond, the initial value was calculated as 25% more than the measured distance (from the geometrically optimised structure) between the atoms in question. The final value was calculated as 25% less than the expected bond length for the element type, bond type, and hybridisation in question. For example (refer to Figure 26), to form a C-C bond in a system where the atoms in the geometrically optimised structure were originally 2.419 Å apart, the initial and final distances were calculated as follows:

$$Initial = 2.419\ \text{Å} + \left(2.419\ \text{Å} \times 25\%\right) =\ 3.024\ \text{Å}$$

$$Final = 1.512\ \text{Å} - \left(1.512\text{Å} \times 25\%\right) = 1.134\ \text{Å}$$

(2) To break a bond, the initial value was calculated as 25% less than the measured value and the final value was calculated as twice the measured value. For example (refer to Figure 26), to break a standard C-C bond, initial and final distances were calculated as follows:

$$Initial = 1.512\ \text{Å} - \left(1.512\text{Å} \times 25\%\right) = 1.134\ \text{Å}$$

$$\text{Final} = 1.512\text{Å} \times 2 = 3.024\text{Å}$$

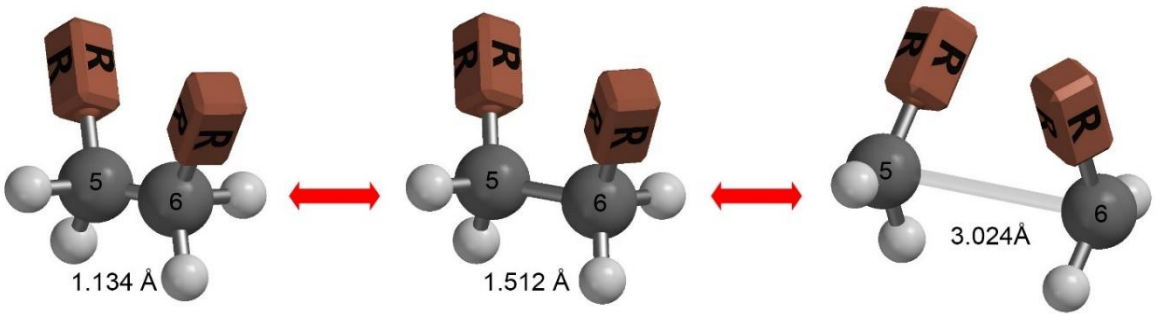

Figure 26: Example of constrained property during PES scan.

In the figure above, the labelled atoms' numbers do not coincide with IUPAC nomenclature, but represent the order in which the atoms were drawn in Spartan software (v14 Wavefunction Inc., Irvine, USA). 'R' represents the rest of the molecule. During the PES scan, the distance between atoms 5 and 6 (in this example) was constrained and incrementally changed, while the rest of the molecule was geometrically optimised and the E_e reported. The resulting energy profile was similar to that depicted in Figure 16.

When the trend derived from a PES scan's energy profile was indicative of a TS, a TS search was performed. This was achieved by placing the optimised starting reaction intermediate in a document together with an intermediate reacting with it, paying special attention to the orientation of bonding orbitals. For an example refer to Figure 27: Pyruvate together with a hydronium ion was placed in a document; the proton to be exchanged was selected and the bonding orbitals were orientated. The set was then geometrically optimised, followed by a TS search using the same setup as that for the geometric optimisation, with task set to 'TS search'. After a TS had been determined, refined TS optimisation calculations were performed on the structure to determine E_e. This was also achieved with the same calculation setup, with task set to 'TS optimisation'. Energy values obtained from both geometric and TS optimisations were used to construct a total energy profile for the proposed reaction mechanism.

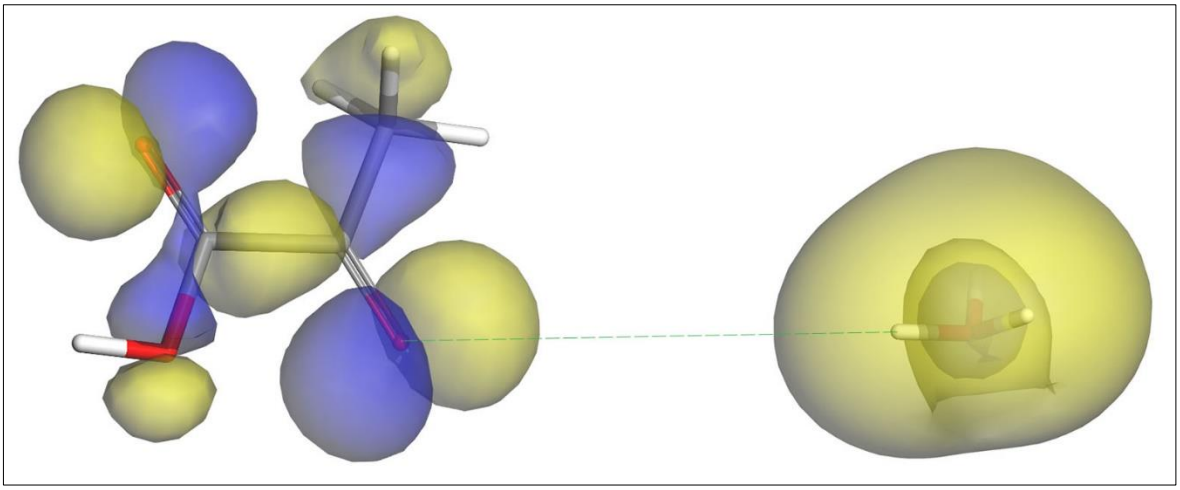

Figure 27: Starting setup for transition state search.

In the figure above, the blue and yellow overlay on pyruvate (left) indicates the HOMO, whereas the overlay on the hydronium ion (right) indicates the LUMO. Yellow and blue indicate opposing lobes (positive/negative) of the same orbital.

CHAPTER 5: RESULTS AND DISCUSSION

5.1 Phase 1: Standardisation of new methods

5.1.1 Standardisation of hydrolysis of N-AAA conjugates

The initial approach included the testing of various hydrolysis variables mentioned in section 4.7. Complete hydrolysis was achieved with 2.0 mol/L NaOH with an incubation period of 1 h at 110°C. This, however, proved problematic when moving to the next phase of the study, namely the separation of enantiomers via GC-MS (refer to section 4.8). After the hydrolysate was dried with nitrogen, the dried sample consisted of a large mass ratio of precipitated NaOH to lysed amino acids (approximately 2 000 to 1). This made derivatisation with a small volume of *R*-2-butanolic HCl impractical. The following criteria were set for an ideal hydrolysate:

1. Very little (approximately 100 µg) hydrolysis reagent or derivatives thereof had to be present in the final dried preparation.

2. The hydrolysate should contain no constituents that could interfere with derivatisation with *R*-2-butanol or MTBSTFA with 1% t-BDMCS (e.g. chloride ions).

3. If (1) and (2) could not be avoided, a method to purify the lysed amino acids from the hydrolysate would have to be practicable and then standardised.

Since separating very small amounts of amino acids from large amounts of ionic precipitate proved near impossible, reagents were chosen that would decompose upon heating or that could evaporate completely, leaving little or no residue when dried (at 100°C and standard atmospheric pressure). Despite being effective hydrolysis reagents, NaOH and KOH were excluded from the new criteria mentioned above, since they do not evaporate at 100°C. The weak acids and alkalis tested (ammonium hydroxide, glacial acetic acid and formic acid) were easier to evaporate, but proved to be ineffective hydrolysis reagents. Sulphuric acid, a strong acid with a low vapour pressure, could not be dried effectively at 100°C. Since HCl(aq) is a gaseous solution and a strong acid, it completely evaporated at a 100°C and proved to be an effective hydrolysis reagent. It was therefore selected as the hydrolysis reagent. Using less than 2.0 mol/L HCl was ineffective, while using more did not improve results. This was also true for an incubation temperature of 110°C and an incubation time of 1 h.

Figure 28 to Figure 32 demonstrate the effectiveness of the final method: Figure 28 shows the organic acid analysis result of N-Ac-L-Ala prior to hydrolysis, indicating a sufficient quantity of this conjugate for the procedure. Figure 29 shows the amino acid analysis result of the same

sample, confirming that the sample was not contaminated with detectable amounts of Ala prior to hydrolysis. Figure 30 shows another organic acid analysis result of N-Ac-L-Ala, but this time after hydrolysis. Only a trace amount of the mono-trimethylsilyl (TMS) derivative of N-Ac-L-Ala was detectable, indicating complete hydrolysis of N-Ac-L-Ala. Figure 31 shows an amino acid analysis of Ala, also after hydrolysis, confirming the presence of sufficient amounts of Ala, indicating that the procedure did not significantly decompose the desired end product. After standardising the method with N-Ac-L-Ala, the method was also tested with all the available N-AAA conjugates and found to be effective in all cases (results not shown). Figure 32 shows the enantiomeric separation of L- and D-Ala after the hydrolysis of N-Ac-L-Ala, indicating an acceptably low amount of D-Ala. It is unclear, however, whether the detected amount of D-Ala originated from a slightly enantio-impure N-Ac-L-Ala standard and/or racemisation induced during hydrolysis and derivatisation. The observed D-Ala might also be the opposing L-Ala diastereomer, originating from a slightly enantio-impure *R*-2-butanolic HCl derivatisation reagent or a combination of all the aforementioned factors. Nonetheless, the intensity of the second peak is still low enough to deduce whether the original sample consisted of a racemic or enantiomeric excess of the original N-AAA conjugate.

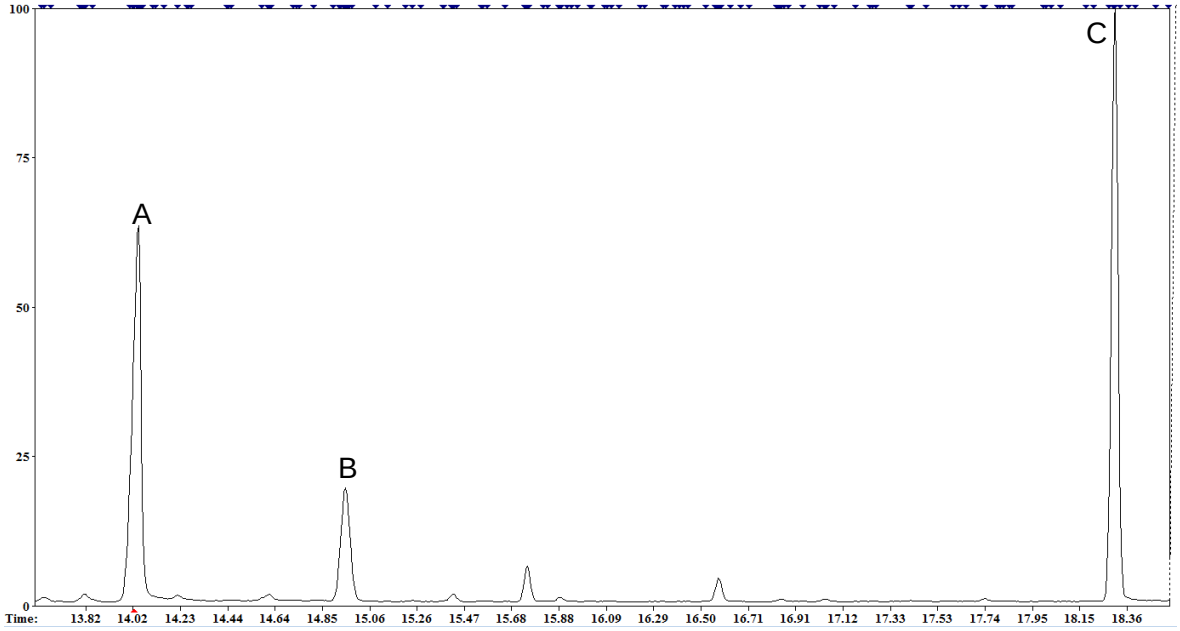

Figure 28: Organic acid analysis of N-Ac-Ala prior to hydrolysis.

Compound A is the mono-TMS derivative of N-Ac-Ala, B is the di-TMS derivative of N-Ac-Ala and C is the internal standard, 3-phenylbutyric acid. The y-axis (abundance) is scaled to 100% of C and the x-axis (retention time in minutes) is cropped to between 13.8 and 18.4 min to show only the section of interest.

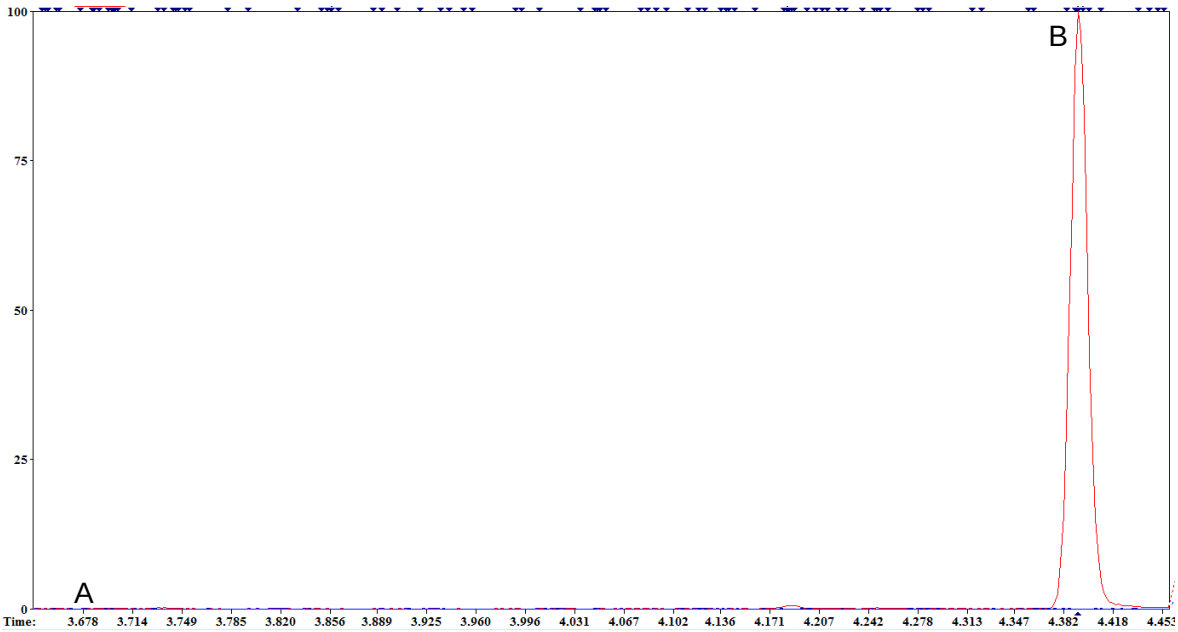

Figure 29: Amino acid analysis of free Ala prior to hydrolysis.

Compound A (trace amount) is Ala and B is the internal standard, Nva. The y-axis (abundance) is scaled to 100% of B and the x-axis (retention time in minutes) is cropped to between 3.5 and 4.5 min to show only the section of interest.

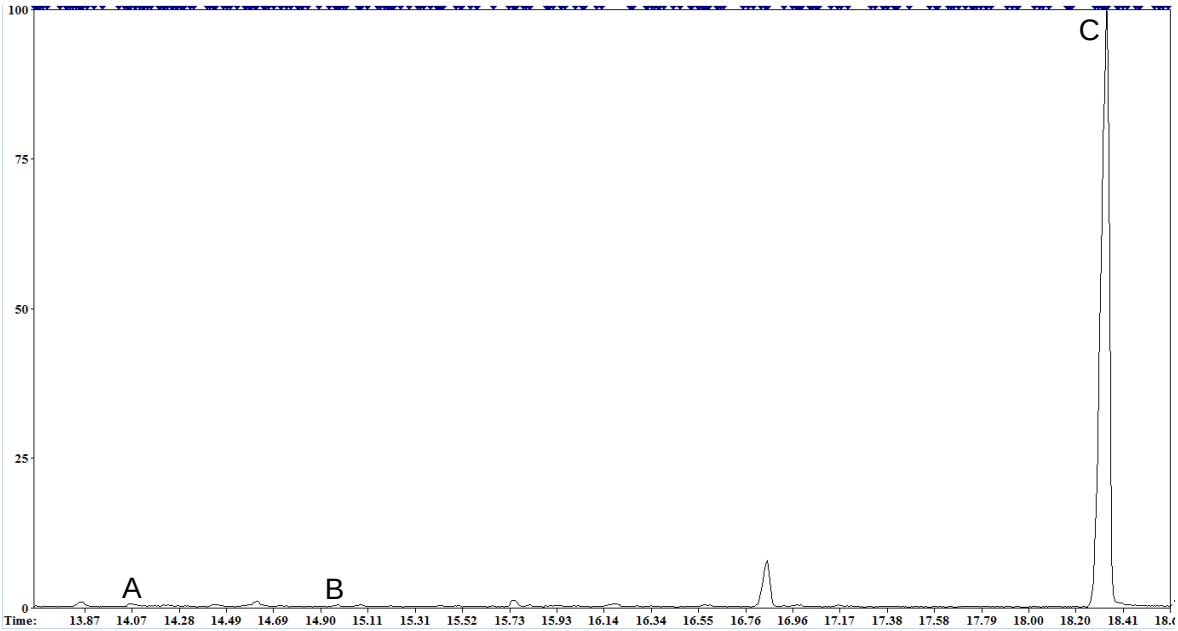

Figure 30: Organic acid analysis of N-Ac-Ala after hydrolysis.

Compound A is the mono-TMS derivative of N-Ac-Ala, B is the di-TMS derivative of N-Ac-Ala and C is the internal standard, 3-phenylbutyric acid. The y-axis (abundance) is scaled to 100% of C and the x-axis (retention time in minutes) is cropped to between 13.8 and 18.4 min to show only the section of interest.

Figure 31: Amino acid analysis of free Ala after hydrolysis.

Compound A is Ala and B is the internal standard, Nva. The y-axis (abundance) is scaled to 100% of B and the x-axis (retention time in minutes) is cropped to between 3.5 and 4.5 min to show only the section of interest. The graph is an overlay (blue and red lines) of the ions used during amino acid single ion monitoring.

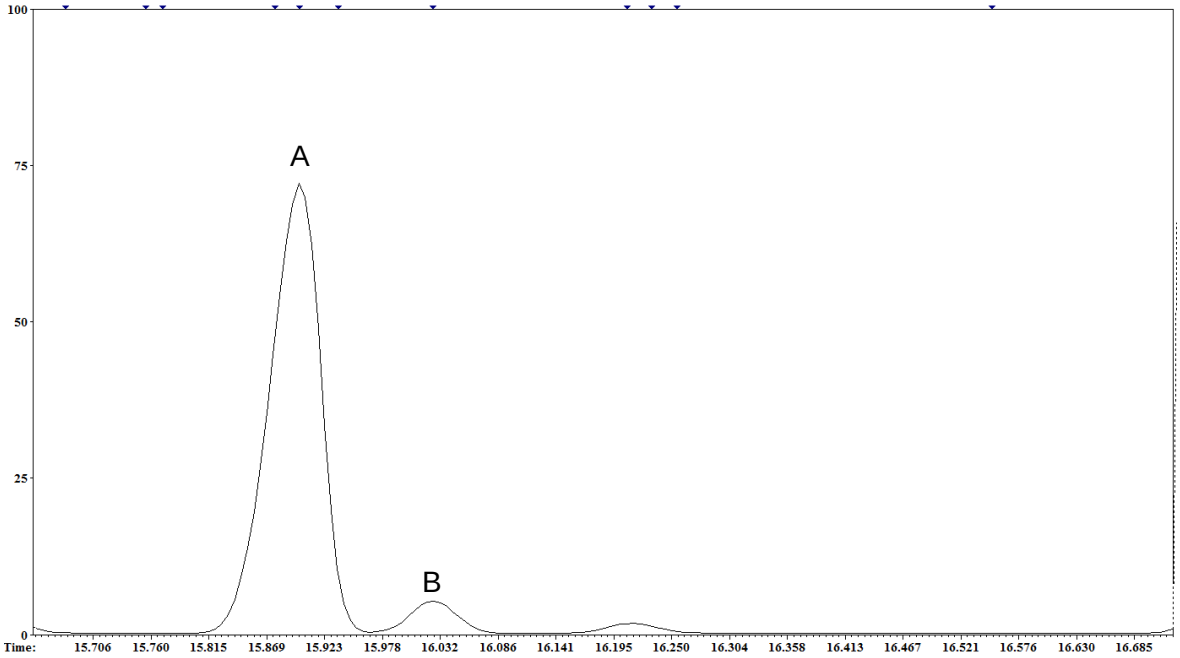

Figure 32: Hydrolysis- and/or derivatisation-induced racemisation of L-Ala.

Compound A is L-Ala and B is D-Ala and/or the opposing diastereomeric derivative of L-Ala. The y-axis (abundance) is scaled to 100% of the di-tBDMS ester of Ala (not shown) and the x-axis (retention time in minutes) is cropped to between 15.7 and 16.7 min to show only the section of interest.

5.1.2 Separation of racemic amino acids via GC-MS (qualitative)

Based on the method described by Bremer *et al*. (1981:483) and other routine methods utilised in-house with similar derivatisation techniques, initial reaction conditions were selected as follows: A sample containing approximately 100 µg DL-Val was derivatised with 150 µL 3 mol/L *R*-2-butanolic HCl for 90 min at 110°C. Thereafter the sample was evaporated to dryness under a steady stream of nitrogen at 40°C for approximately 30 min and subsequently derivatised with 100 µL MTBSTFA with 1% t-BDMCS and 50 µL pyridine at 75°C for 30 min.

This resulted in three relevant peaks on the chromatogram, which was unexpected (refer to Figure 33). Futher investigation revealed that the first two compounds were the expected L- and D-*N*-*tert*-butyldimethylsilyl-*O*-2-butanyl derivatives of Val, while the third compound was the *N,O*-di-*tert*-butyldimethylsilyl ester of Val (refer to Figure 34). It seems likely that the first *O*-esterification step was either incomplete or that the 2-butyl group was replaced by the more reactive *tert*-butyldimethylsilyl (tBDMS) group. In either case, since the objective was only to develop a qualitative method and the third compound proved useful to validate the identification of other compounds, complete derivatisation of the hydroxy side chain with *R*-2-butanol was not pursued. To elaborate (refer to Figure 35), even though D-aIle and L-Ile did not separate with

this method, the DL-*N,O*-di-tBDMS esters of aIle and Ile could and did give an indication as to which may be present in a sample, e.g. an enantiopure standard of D-Ile was not available at the time of method standardisation, but only the mixture DL-Ile. Previous experience showed that many samples of DL-Ile standards might not be *allo*-free (and *vice versa*), even if stated by the manufacturer. The absence of an *N,O*-di-tBDMS ester of aIle proved that the DL-Ile standard was indeed *allo*-free.

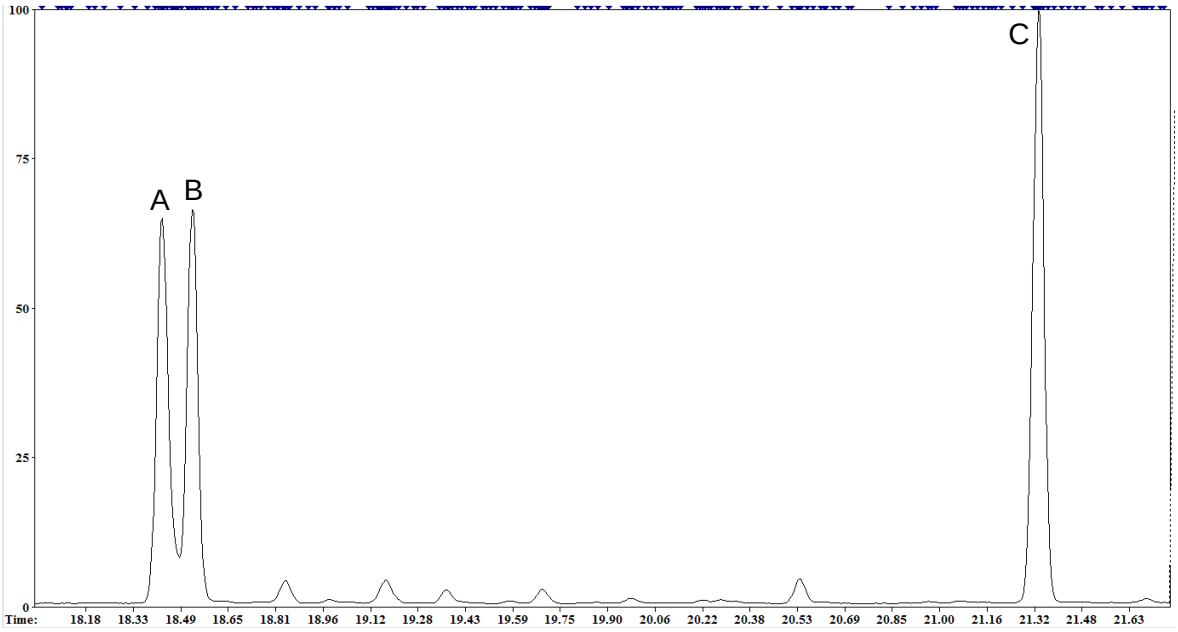

Figure 33: Separation of DL-Val via gas chromatography-mass spectrometry.

Compound A is L-Val, B is D-Val and C is the di-tBDMS ester of DL-Val. The y-axis (abundance) is scaled to 100% of C and the x-axis (retention time in minutes) is cropped to between 18.1 and 21.7 min to show only the section of interest. tBDMS: tert-butyldimethylsilyl.

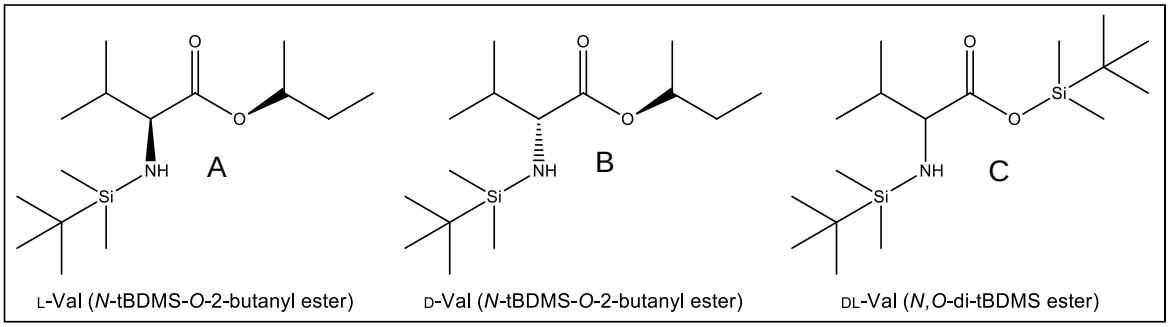

Figure 34: Derivatives of L- and D-Val after treatment with R-2-butanol and MTBSTFA with 1% t-BDMCS.

*A, B and C indicate the chemical structure of the compounds correlating to the peaks in Figure 33. MTBSTFA: N-*tert*-butyldimethylsilyl-N-methyltrifluoroacetamide; t-BDMCS:* tert*-Butyldimethylchlorosilane; tBDMS:* tert*-butyldimethylsilyl.*

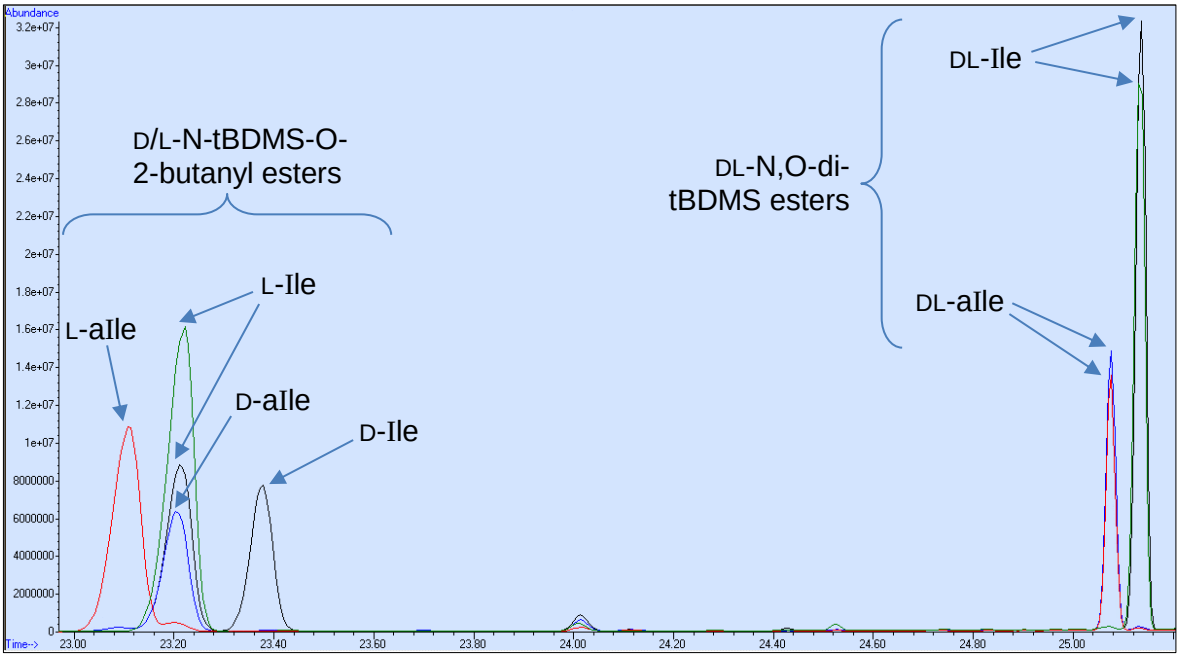

Figure 35: Chromatogram overlay of L- and D-Ile, and L- and D-aIle.

The chromatogram is an overlay of four different runs to illustrate the retention time of certain compounds: red-lined peaks are from a standard of L-aIle, blue peaks are from D-aIle, green peaks are from L-Ile and black peaks are from DL-Ile. The absence of a black-lined peak at the same retention time of the red and blue DL-aIle DL-N,O-di-tBDMS esters shows that the DL-Ile standard is indeed allo-free. The y-axis (abundance) is scaled to 100% of the di-tBDMS ester of DL-Ile and the x-axis (retention time in minutes) is cropped to between 23.0 and 25.2 min to show only the section of interest.

Slightly better results were obtained by increasing derivatisation time for the first step of derivatisation with *R*-2-butanol to 2 h. Lowering the incubation temperature, time or concentration of *R*-2-butanol yielded poor results and higher than 3.0 mol/L *R*-2-butanol, higher temperatures and longer incubation times did not significantly improve results. Refer to Figure 36 for an illustration of the final method with DL-Ala, -Nva and -BCAAs. Note that even though Ala is not a BCAA, it is included in all results in the remainder of this dissertation, as it was extensively used to standardise all methods. Baseline separation of DL-Leu was not achieved (refer to Figure 37), but with an enantiomeric excess of either, the opposite enantiomer was still distinguishable from the other (e.g. if L-Leu was highly elevated compared to D-Leu, D-Leu was still distinguishable from L-Leu).

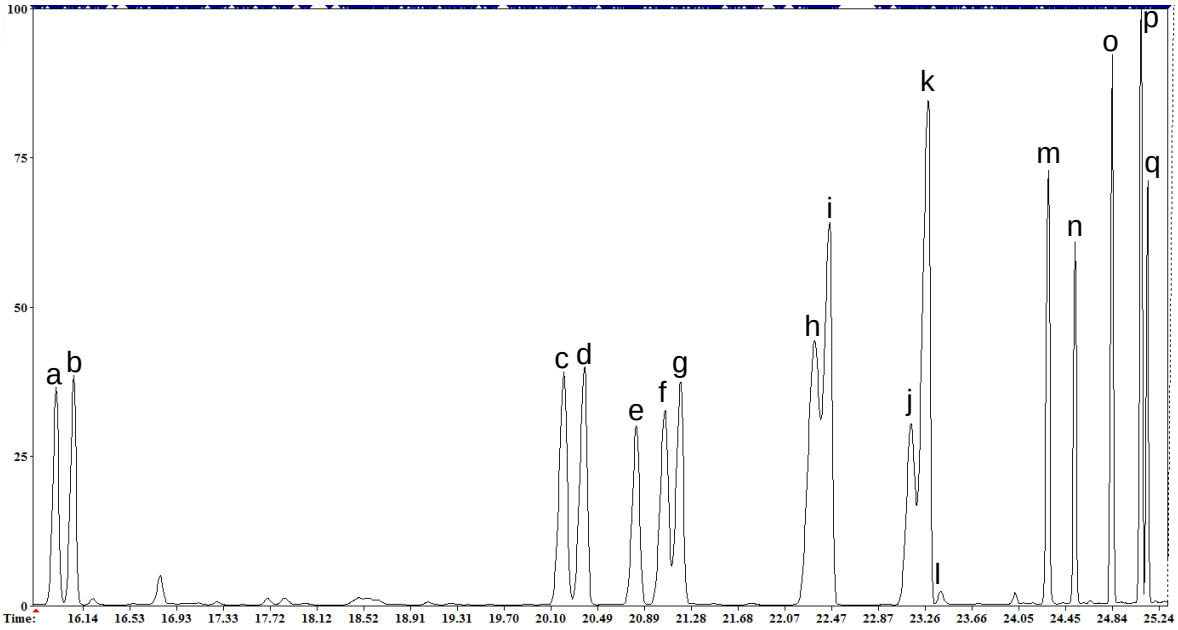

Figure 36: Separation of racemic amino acid via gas chromatography-mass spectrometry.

The compound list is as follows: (a) L-Ala, (b) D-Ala, (c) L-Val, (d) D-Val, (e) the di-tBDMS ester of Ala, (f) L-Nva, (g) D-Nva, (h) L-Leu, (i) D-Leu, (j) L-aIle, (k) D-aIle and L-Ile, (l) D-Ile and (m–q) the di-tBDMS esters of Val, Nva, Leu, aIle and Ile, respectively. The y-axis (abundance) is scaled to 100% of (p) and the x-axis (retention time in minutes) is cropped to between 15.5 and 25.3 min to show only the section of interest. It is interesting to note that with the tested amino acids, the L-isomer eluted before the D-isomers in all cases.

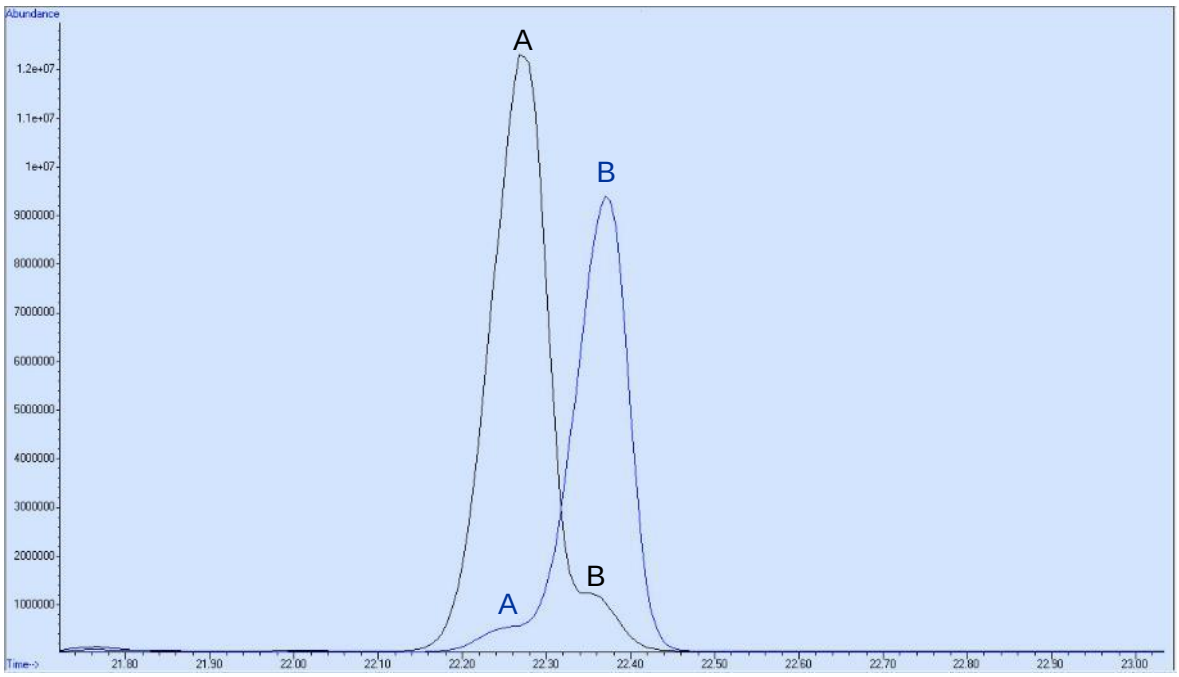

Figure 37: Separation of DL-Val via gas chromatography-mass spectrometry.

Compound A was identified as L-Leu and B as D-Leu. The black line originated from a sample containing L-Leu and the blue line from that containing D-Leu. The y-axis (abundance) is scaled to 100% of the black-lined A and the x-axis (retention time in minutes) is cropped to between 21.7 and 23.0 min to show only the section of interest.

5.1.3 Non-enzymatic *in vitro* synthesis of N-AAA conjugates

Refer to Table 12: The predicted N-AAA conjugate end products reported in Table 11 (in section 4.9), were generated in all cases except for the formation of N-(*p*-hydroxyphenylacetyl)-(*p*-hydroxy-Phe) from *p*-hydroxyphenylpyruvate and ammonia. A mass spectrum was not available for this compound and one could not be derived from the surprisingly cluttered chromatogram. The best results were obtained with ammonium sulphate, at pH <2, with an incubation time of 1 h at 37°C. No result was obtained at pH 7 and above, whereas good results were obtained at pH 3.5, which further improved at pH <2. With no incubation, trace amounts of N-AAA conjugates formed, which significantly improved at 1 h; increasing incubation times to 24 h and more resulted in poor repeatability. No results were obtained at room temperature, whereas increasing the temperature to above 37°C did not significantly improve results. At 110°C and at pH <2 hydrolysis of the final product occurred.

Table 12: N-AAA conjugate yield from *in vitro* synthesis.

2-Keto acid substrate	Product	Yield (µmol/L)
2-Keto-3-methylvaleric acid	N-2-Methylbutyry-Ile	Trace ~ 0.6
2-Ketobutyric acid	N-Propionyl-2-aminobutyric acid	10.12
2-Ketocaproic acid	N-Pentanoyl-*nor*leucine	14.70
2-Ketoisocaproic acid	N-Isovaleryl-Leu	1.35
2-Ketoisovaleric acid	N-Isobutyryl-Val	20.32
2-Ketovaleric acid	N-Butyryl-*Nva*	Trace ~ 1.0
Phenylpyruvate	N-(Phenylacetyl)-Phe	5.22
p-Hydroxyphenylpyruvate	No result	Not detected
Pyruvate	N-Ac-Ala	2.46

Omitting the low pH of the stomach (since ACY1 in the gut would hydrolyse N-AAA conjugates upon absorption [Perrier *et al*., 2005:677]), the low pH required for the successful synthesis of the tested N-AAA conjugates excluded the first criterion mentioned in section 4.9 (physiological conditions). However, the method satisfied the second and third criteria. These findings confirmed the non-enzymatic *in vitro* synthesis of N-AAA conjugates from 2-keto acids and ammonia, but the low yields and reaction conditions required increased doubt as to whether this may be achievable *in vivo* via the proposed mechanism.

Ammonium acetate, ammonium carbonate and ammonium bicarbonate did not produce any products. Refer to Figure 38: Incubation with urea resulted in the formation of significant amounts of the related hydantoin conjugates with almost complete depletion of the substrates, but resulted in no N-AAA conjugate formation. It is possible that the hydantoin conjugates formed via a similar mechanism to that of N-AAA conjugate formation. A mechanism for the non-enzymatic formation of hydantoin conjugates from 2-keto acids and ammonia is proposed in Figure 39. Refer to Figure 40 for the net reaction. By using this net reaction, the resulting hydantoin conjugate end product could be predicted in each case. If ammonium sulphate and a 2-keto acid were added to a urine control, all the available 2-keto acids reacted with the native urea in the urine first, resulting in large amounts of hydantoin conjugates and only trace amounts of N-AAA conjugates. This observation is in contrast to the biochemical image of sample 11, which consisted of large amounts of various N-AAA conjugates and only trace amounts of hydantoin conjugates. Substantiated by the results presented in the sections below, a plausible explanation of the latter observation is presented in section 6.4.1.

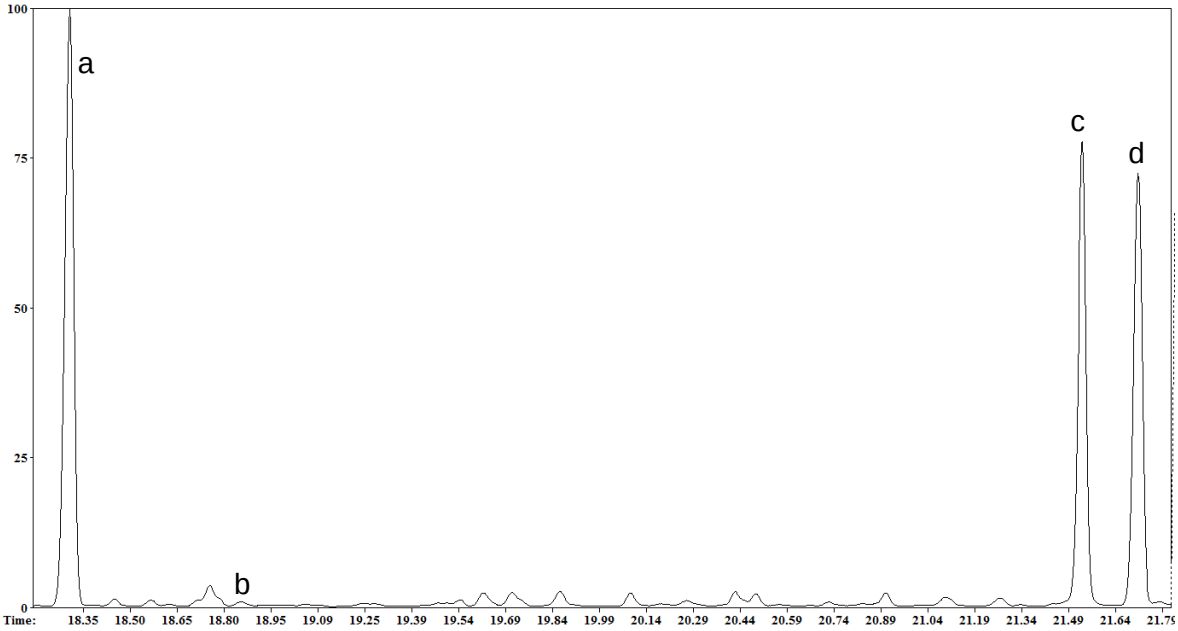

Figure 38: Formation of 5-hydroxy-5-isopropyl hydantoin from 2-ketoisovaleric acid and urea.

Compound (a) is the internal standard, 3-phenylbutyric acid, (b) is N-isobutyryl-Val, (c) contaminant and (d) is 5-hydroxy-5-isopropyl hydantoin. The substrate 2-ketoisovaleric acid (not shown) was not detectable. The y-axis (abundance) is scaled to 100% of (a) and the x-axis (retention time in minutes) is cropped to between 18.2 and 21.8 min to show only the section of interest.

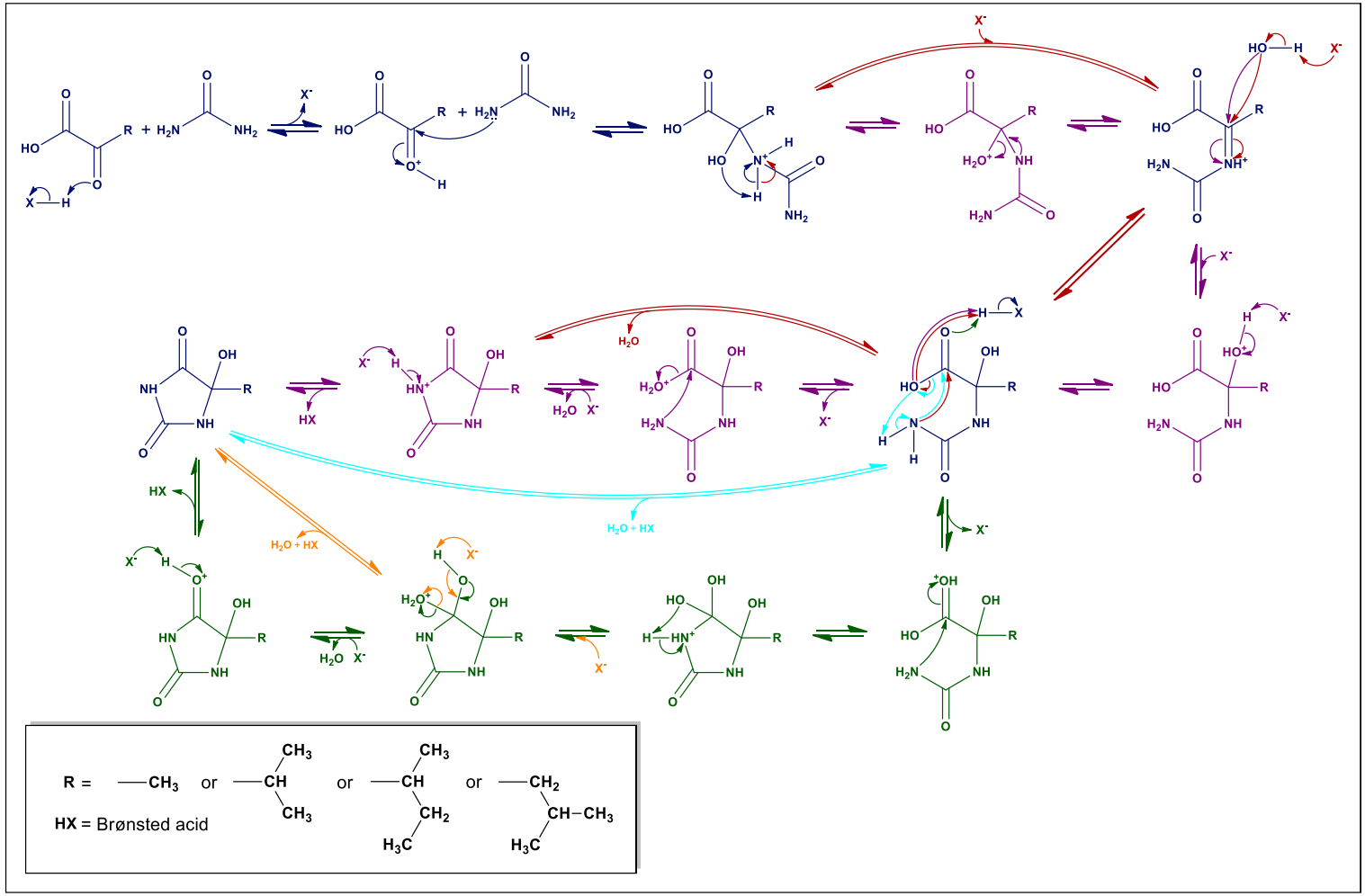

Figure 39: Proposed mechanism for the non-enzymatic formation of hydantoin conjugates from 2-keto acids and urea.

R represents a methyl, isopropyl, 2-butyl or isobutyl side chain. The different colored arrows and structures represents alternative reaction mechanisms for the formation of hydantoin conjugates. The symbol 'HX' is used to indicate a Brønsted acid.

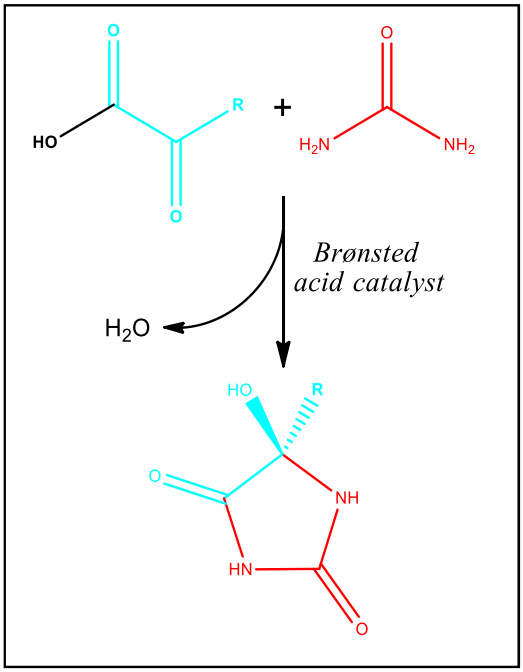

Figure 40: Net reaction for the non-enzymatic formation of hydantoin conjugates from 2-keto acids and urea.

5.2 Phase 2: Optimisation and rehearsal of the analytical strategy

5.2.1 Determining whether N-AAA conjugates were present in controls

Using the library described at the end of section 4.9 and given in Annexure C-Table 18, none of the N-AAA conjugates were detected in the 10 control samples. This was expected since N-AAA conjugates are not detectable in urine samples of healthy controls (Gerlo *et al*., 2006:199). This is also evident in the lack of liberated BCAAs when the samples were treated with the hydrolysis reagent mentioned in section 5.2.2. This meant that the control samples could be used to further test the analytical strategy without the interference of native N-AAA conjugates.

5.2.2 Determining whether any compounds in the controls would liberate BCAAs when hydrolysed

Refer to Figure 41 and Figure 42: From the 10 controls that were used to determine whether any other compounds would liberate BCAA upon hydrolysis, only one sample (Figure 42) showed negligible traces of the di-tBDMS derivatives, which may also have been carried over from previous injections of standards. Although not definitive, this indicates a low probability of other conjugates (i.e. other than the observed N-AAA conjugates) in sample 11 producing amino acids upon hydrolysis.

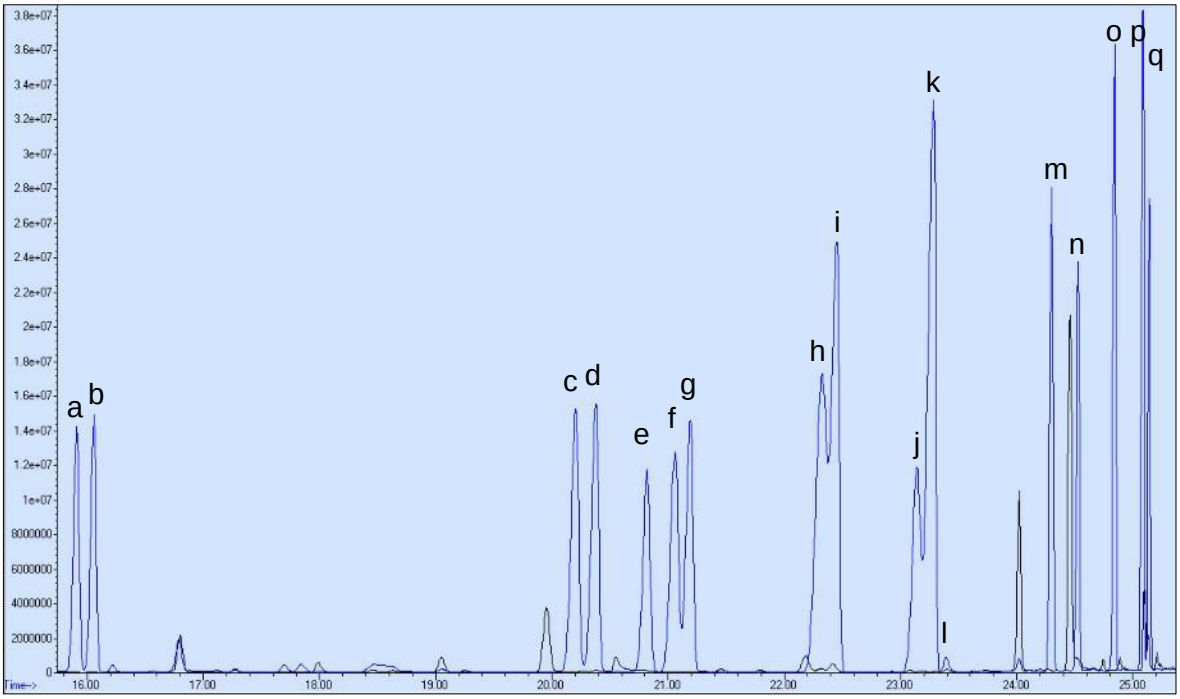

Figure 41: Example 1 of branched-chain amino acid (BCAA) enantiomer profile in a control sample with no additional N-AcAA conjugates or added BCAAs.

The example above is merged with Figure 36 (blue chromatogram), this serves to indicate where D/L-amino acids would have been present. The black chromatogram is the control sample. The y-axis (abundance) is scaled to 100% of (p) and the x-axis (retention time in minutes) is cropped to between 15.7 and 25.3 min to show only the section of interest. The compound list includes the following: (a) L-Ala, (b) D-Ala, (c) L-Val, (d) D-Val, (e) the di-tBDMS ester of Ala, (f) L-Nva, (g) D-Nva, (h) L-Leu, (i) D-Leu, (j) L-aIle, (k) D-aIle and L-Ile, (l) D-Ile and (m–q) the di-tBDMS esters of Val, Nva, Leu, aIle and Ile, respectively.

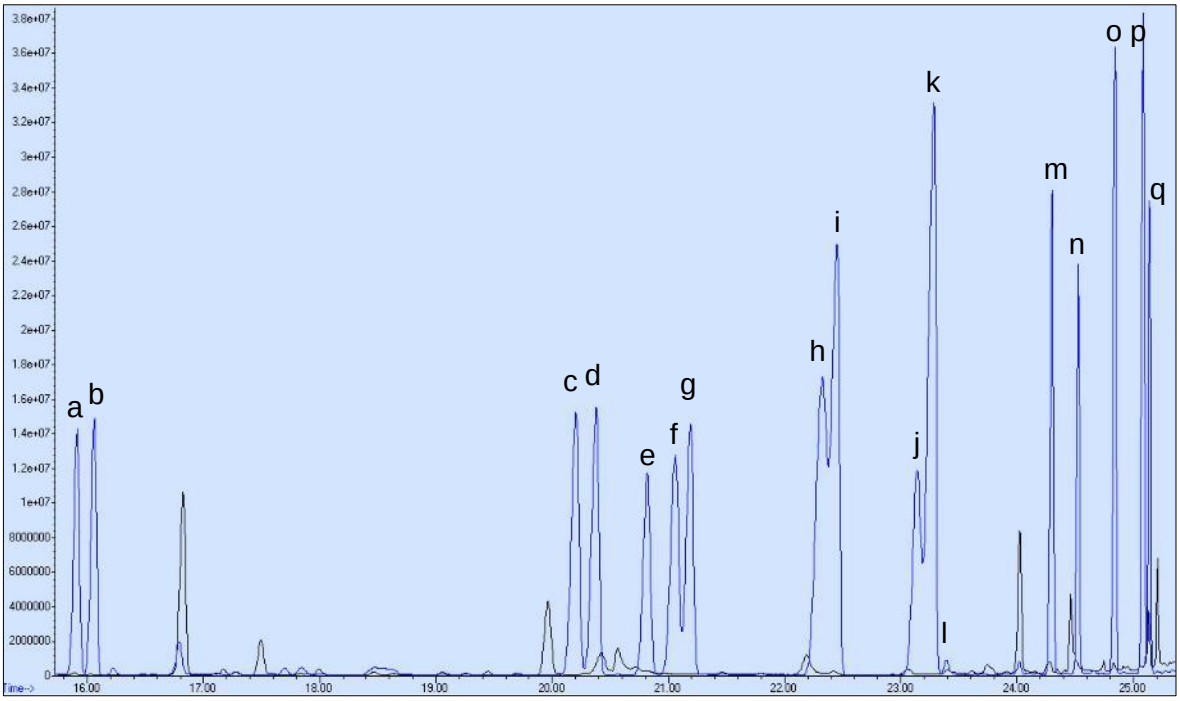

Figure 42: Example 2 of branched-chain amino acid (BCAA) enantiomer profile in control with no additional N-AcAA conjugates or added BCAAs.

The example above is merged with Figure 36 (blue chromatogram), this serves to indicate where D/L-amino acids would have been present. The black chromatogram is the control sample. The y-axis (abundance) is scaled to 100% of (p) and the x-axis (retention time in minutes) is cropped to between 15.7 and 25.3 min to show only the section of interest. The compound list is as follows: (a) L-Ala, (b) D-Ala, (c) L-Val, (d) D-Val, (e) the di-tBDMS ester of Ala, (f) L-Nva, (g) D-Nva, (h) L-Leu, (i) D-Leu, (j) L-aIle, (k) D-aIle and L-Ile, (l) D-Ile and (m–q) the di-tBDMS esters of Val, Nva, Leu, aIle and Ile, respectively.

5.2.3 Confirming that BCAAs were not extracted by the organic acid extraction method

Refer to Table 13 and Figure 43: To determine whether any amount of detected BCAAs originated from an impure isotope mixture or column carryover, a reagent blank (containing ddH_2O instead of sample) was prepared and analysed. This was compared to the amount of amino acids detected in the organic acid extract of the 10 control samples after spiking with a 600 µmol/L mixture of amino acids. No significant amount of amino acids was detected in any of the samples. This indicated that the organic acid extraction method did not extract any significant amounts of amino acids. This also means that subjecting a sample, with a high concentration of amino acids (as is the case with MSUD) to the analytical strategy, will not result in the extraction of significant amounts of these L-amino acids, which could present a pseudo-enantiomeric excess.

Table 13: Comparison between a reagent blank and amino acids detected in an organic acid extract.

Amino acid	Reagent blank (µmol/L)	Control average (n = 10) (µmol/L)
Alanine	2.88	2.81 ± 0.08
Valine	1.39	1.53 ± 0.51
Leucine	1.31	1.23 ± 0.09
Isoleucine	0.73	0.68 ± 0.06

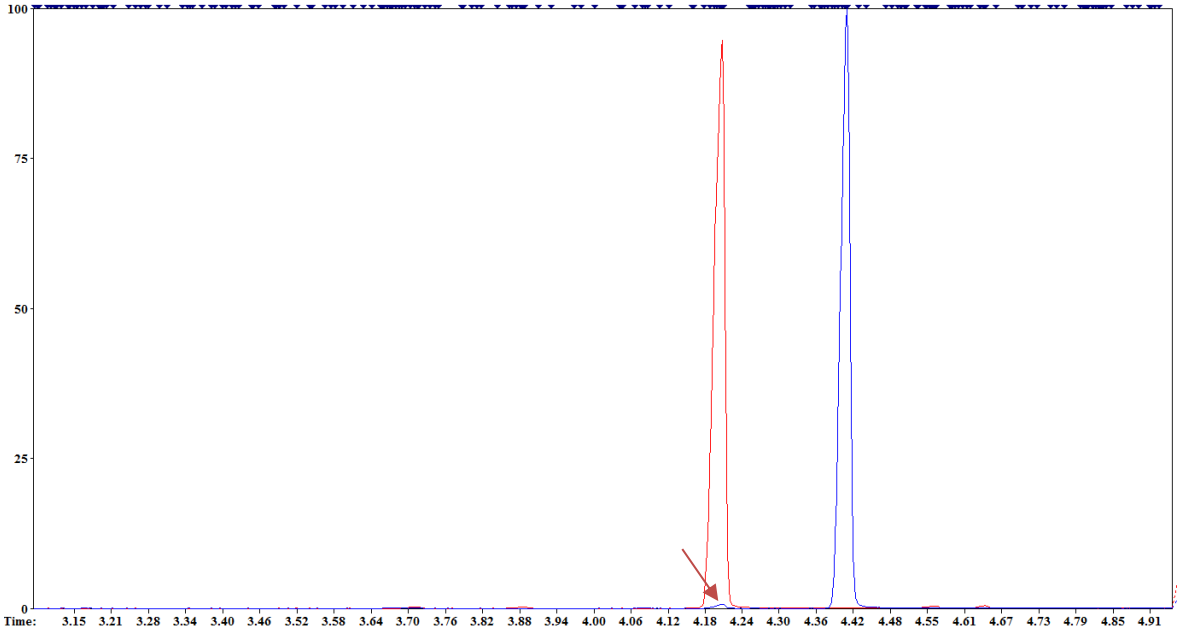

Figure 43: Chromatogram of an amino acid reagent blank.

The large red peak indicates the U-^{13}C amino acid isotope of Val and the large blue peak indicates that of the internal standard Nva. The red arrow points to the trace amount of non-isotopic Val present in the isotope mixture. This was also the case with Ala, Leu and Ile. This corresponds with the trace amounts of amino acids detected in the reagent blank and controls reported in Table 13. The y-axis (abundance) is scaled to 100% of Nva and the x-axis (retention time in minutes) is cropped to between 3.1 and 5.0 min to show only the section of interest.

5.2.4 Determining whether any constituents in the control samples would overlap with the peaks of BCAA enantiomers on the GC chromatograms

Refer to Figure 44: From the 10 samples that were used to test for overlapping peaks, only one showed a low concentration compound overlapping with L-Leu (indicated in Figure 44 with a red arrow). Fortunately, the compounds' mass spectra were sufficiently different and subsequently separated during post-processing.

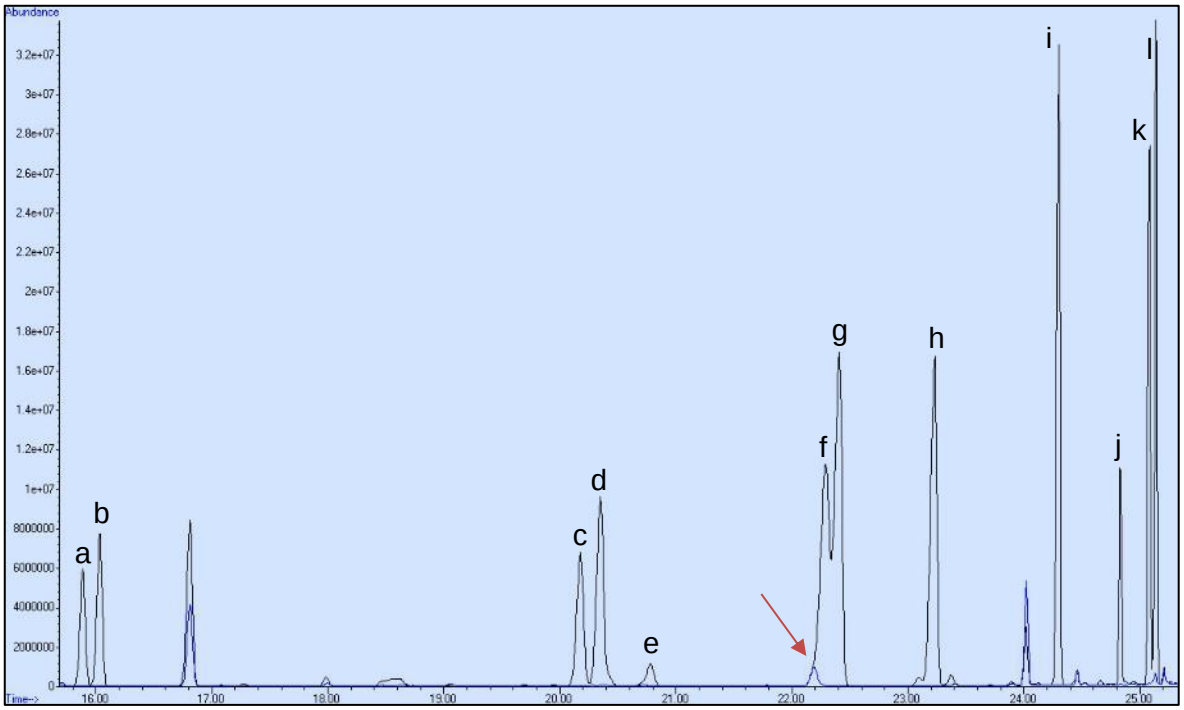

Figure 44: Branched-chain amino acid enantiomer profile of a control spiked with N-AcAA conjugates compared to that of a non-spiked control.

The red arrow indicates a compound that overlaps with L-Leu. The black line represents the chromatogram of the control spiked with N-Ac-DL-Ala, -DL-Val, -DL-Leu, -L-Ile and -D-aIle. The blue line represents the chromatogram of the same non-spiked control. The compound list is as follows: (a) L-Ala, (b) D-Ala, (c) L-Val, (d) D-Val, (e) the di-tBDMS ester of Ala, (f) L-Leu, (g) D-Leu, (h) D-aIle and L-Ile and (i–l) the di-tBDMS esters of Val, Leu, aIle and Ile, respectively. The y-axis (abundance) is scaled to 100% of (l) and the x-axis (retention time in minutes) is cropped to between 15.7 and 25.3 min to show only the section of interest.

The results presented in sections 5.2.1 to 5.2.4 show that the urine controls were indeed suitable for this study and that the analytical strategy was achievable without significant interferences. Furthermore, since the analytical strategy was laborious and precarious, the

control samples were subsequently used to adequately rehearse the analytical strategy, before performing the experiment on the low-volume sample of interest.

5.3 Phase 3: Assessment of the non-enzymatic reaction between 2-keto acids and ammonia

Due to the low yields obtained (reported in Table 12), the enantiomeric composition of the synthesised conjugates could not be determined. However, it is clear from the literature reviewed in section 2.5 that non-enzymatically synthesised chiral compounds formed from prochiral reagents would always result in a racemic end product (McMurry, 2004:295-297).

5.4 Phase 4: Determining the enantiomeric composition of N-AAA conjugates observed in sample 11

5.4.1 Confirming the presence of N-AAA conjugates in sample 11

Refer to Figure 45 and the corresponding figures below (Figure 46–Figure 51) for a visual result and Table 16 in Annexure A for quantitated values. Sample 11 showed significant amounts of various 2-keto acids, N-AAA conjugates and trace amounts of hydantoin conjugates.

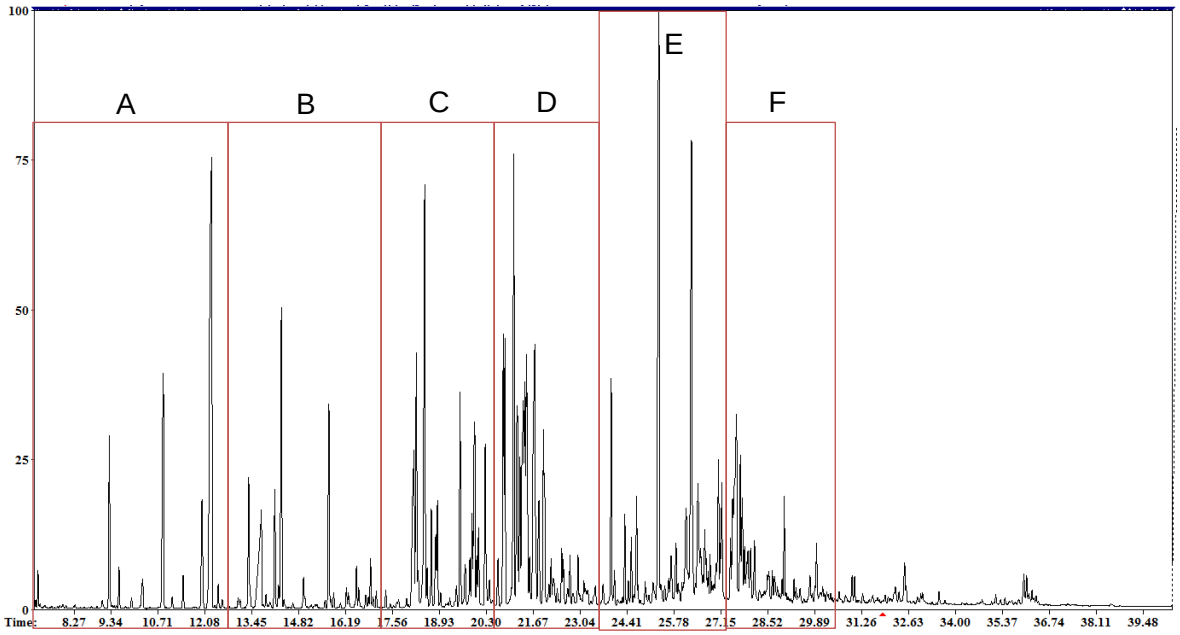

Figure 45: Chromatogram of the organic acid analysis of sample 11.

Due to the complexity of the chromatogram above, it has been divided into six segments (A–F) and given in more detail in the corresponding figures below. The y-axis is the percentage abundance (relative to the most abundant peak in the total ion chromatogram), and the x-axis is the retention time in minutes.

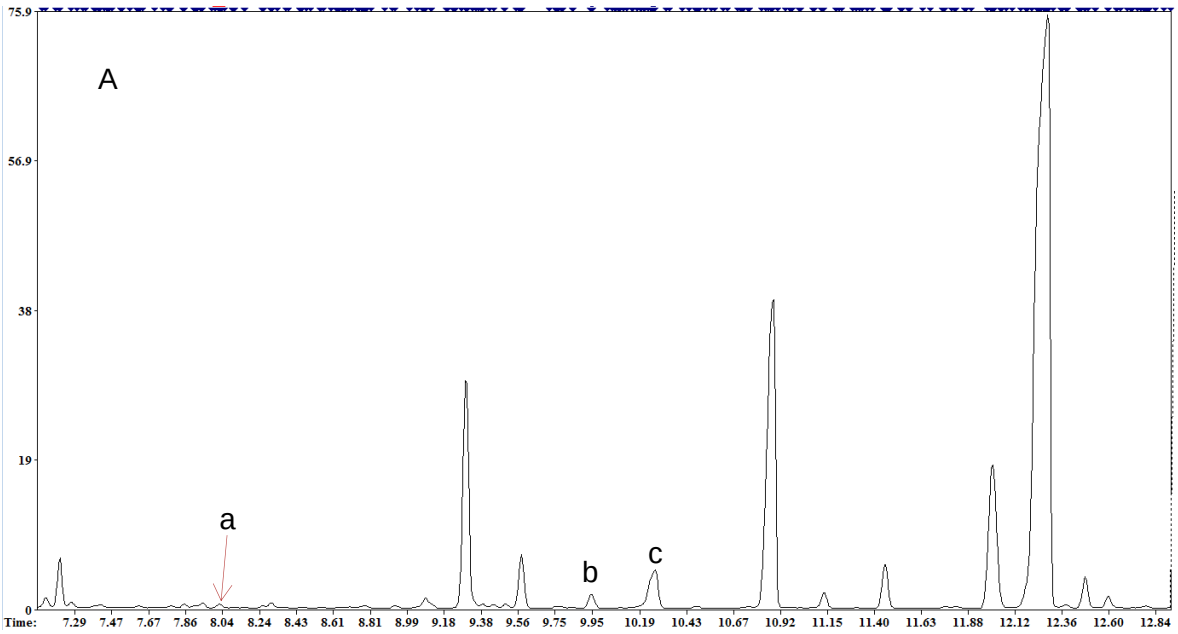

Figure 46: Segment A of Figure 45.

The compound list is as follows: (a) 2-ketoisovaleric acid (mono-TMS ester), (b) pyruvate and (c) 2-keto-3-methylvaleric acid. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment A as indicated in Figure 45.

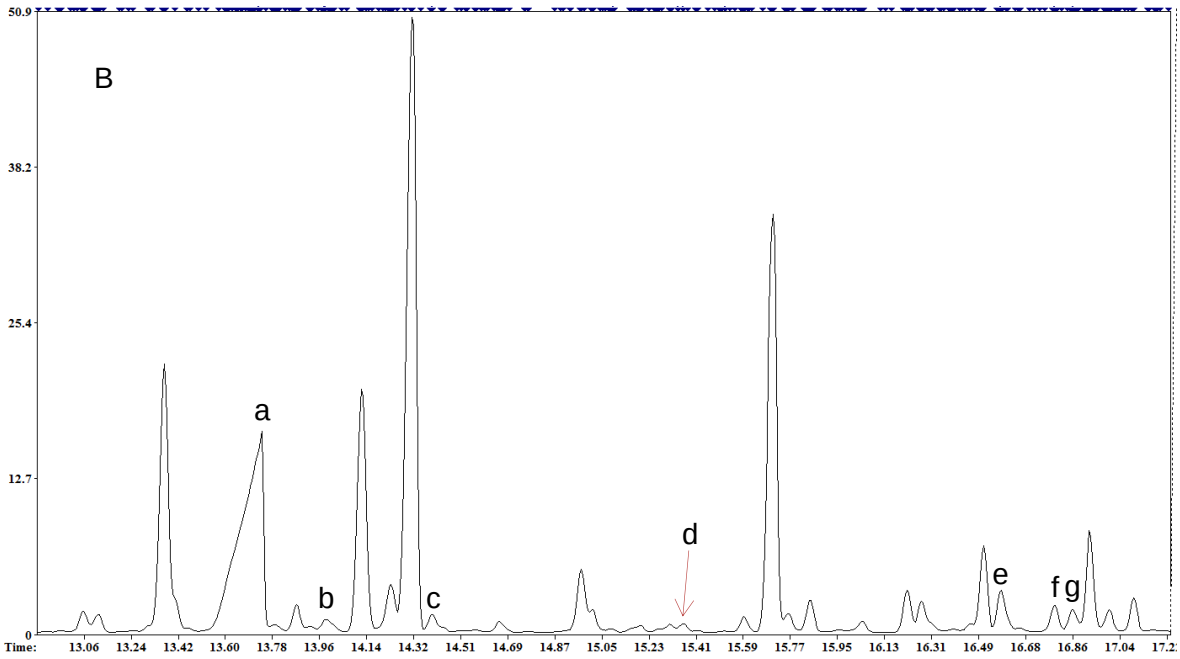

Figure 47: Segment B of Figure 45.

The compound list is as follows: (a) urea, (b) N-Ac-Ala, (c) 2-ketoisocaproic acid, (d) 2-ketoisovaleric acid (di-TMS ester), (e) 2-ketovaleric acid, (f) N-Ac-Val (mono-TMS ester) and (g) N-Ac-Gly. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment B as indicated in Figure 45.

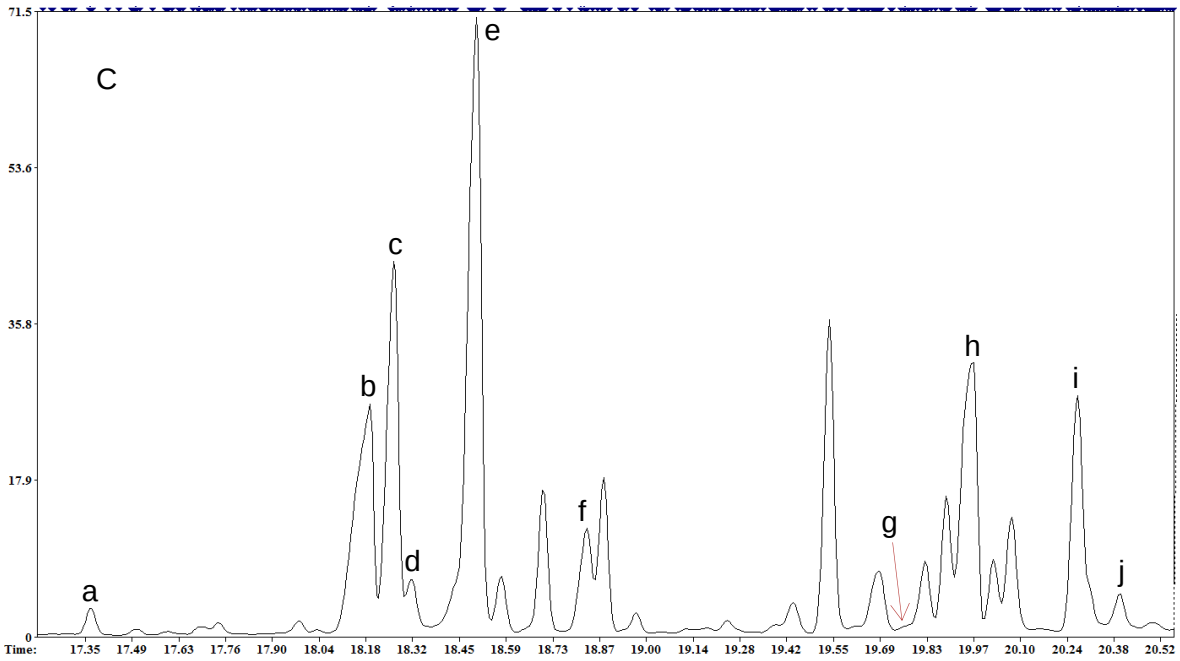

Figure 48: Segment C of Figure 45.

The compound list is as follows: (a) N-Ac-Val (di-TMS ester), (b) N-Ac-Leu, (c) the internal standard, 3-phenylbutyric acid, (d) N-Ac-Ile, (e) N-hexanoyl-Gly, (f) N-isobutyryl-Val, (g) 5-hydroxy-5-methylhydantoin, (h) N-isobutyryl-Leu, (i) N-isobutyryl-Ile and (j) N-2-methylbutyryl-Val. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment C as indicated in Figure 45.

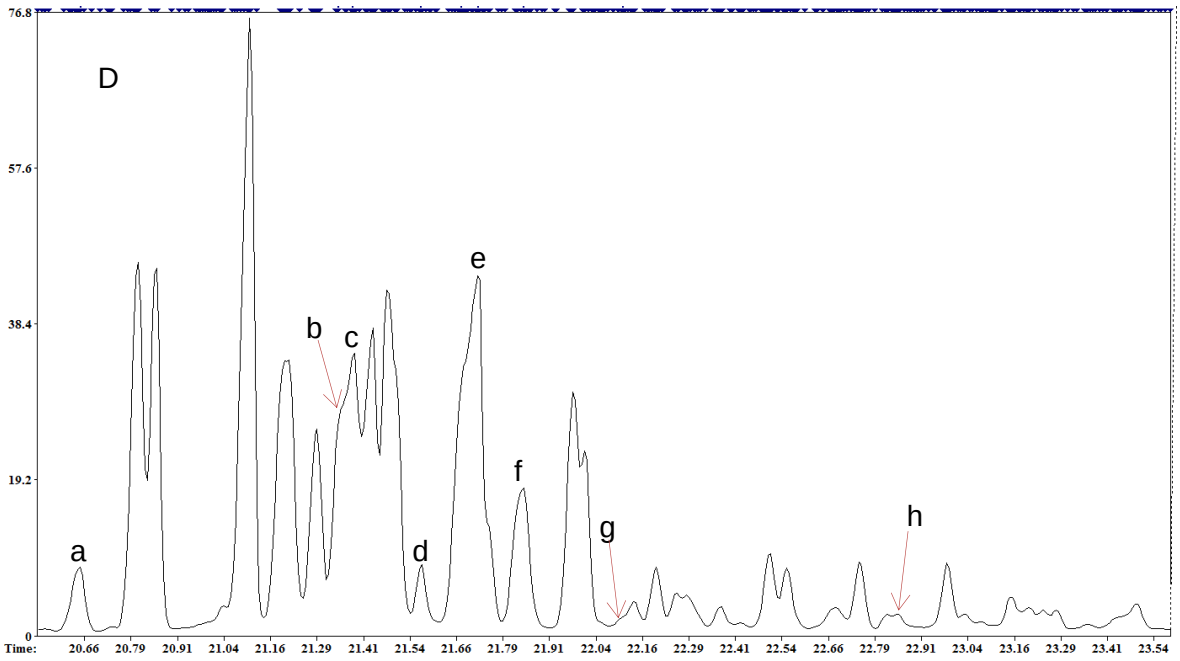

Figure 49: Segment D of Figure 45.

The compound list is as follows: (a) N-isovaleryl-Val, (b) N-isovaleryl-Leu (mono-TMS ester), (c) N-2-methylbutyryl-Leu, (d) N-isovaleryl-Ile, (e) N-2-methylbutyryl-Ile, (f) N-isovaleryl-Leu (di-TMS ester), (g) 5-hydroxy-5-isopropylhydantoin and (h) N-lactyl-Ile. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment D as indicated in Figure 45.

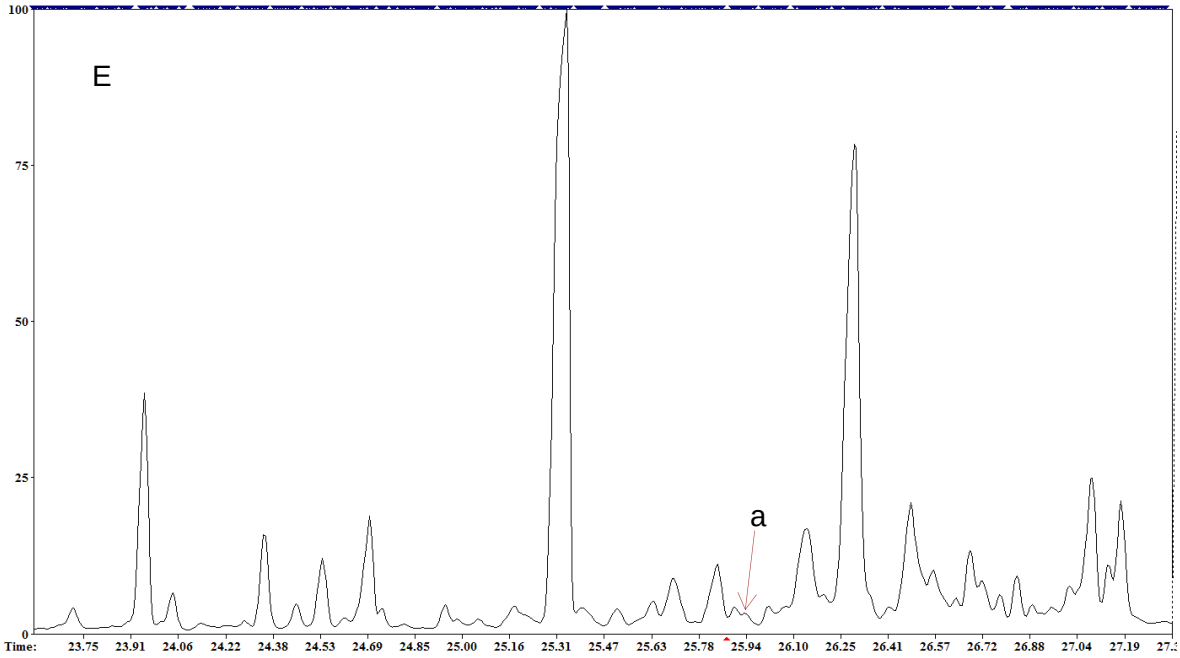

Figure 50: Segment E of Figure 45.

Compound (a) indicates N-isovaleryl-Gln. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment E as indicated in Figure 45.

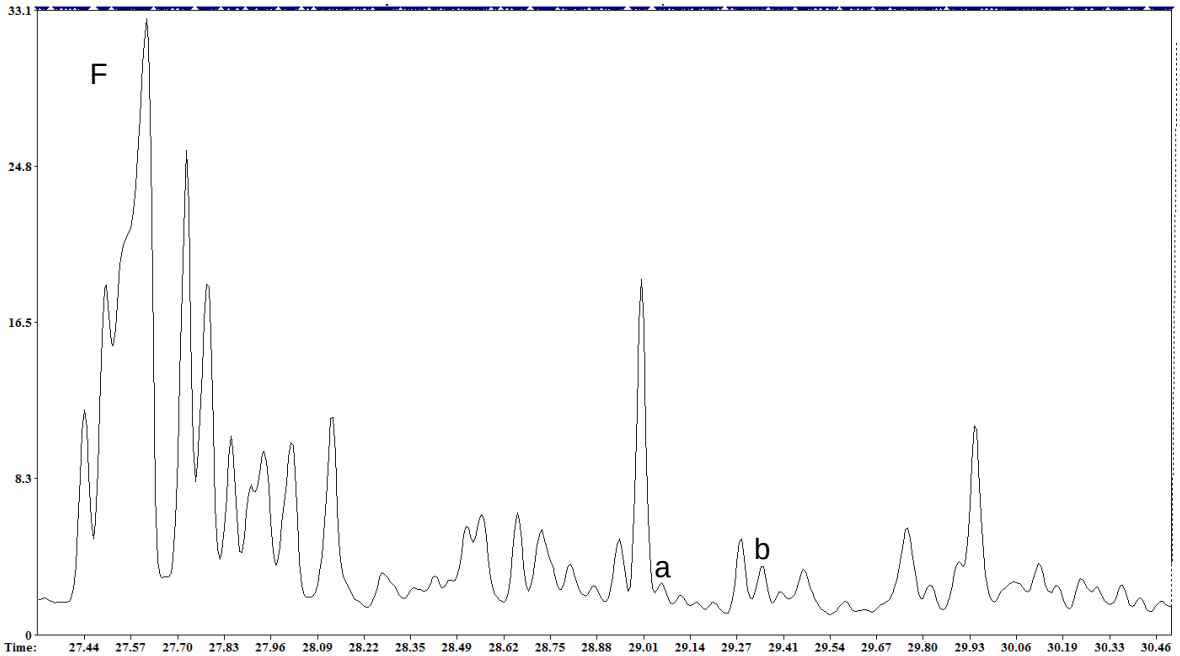

Figure 51: Segment F of Figure 45.

The compound list is as follows: (a) N-2-methylbutyryl-Ile and (b) N-phenylacetyl-Gln. The y-axis (abundance) is scaled to the most abundant peak of this segment and the x-axis (retention time in minutes) is cropped to only show segment F as indicated in Figure 45.

In summary, the results above indicate significant amounts of the N-acylated-BCAA conjugates: N-isovaleryl-, N-2-methylbutyryl- and N-isobutyryl-BCAA conjugates. A significant amount of N-Ac-Leu was also present, as were lower levels of N-Ac-Val, -Ala and -Ile. Trace amounts of 5-hydroxy-5-methyl hydantoin and 5-hydroxy-5-isopropyl hydantoin conjugates were also detectable. It is worth noting that these were the only two hydantoin conjugates with available mass spectra and more might have been present in the sample.

5.4.2 Determining enantiomeric composition of N-AAA conjugates in sample 11

After sufficient rehearsal and scrutiny of the selected methods, the protocol outlined in Table 9 was executed. Refer to Figure 52: Together with sample 11, three controls were analysed to assess each critical step in the protocol: an extraction control (a urine control sample spiked with a mixture of N-AcAA conjugates), a hydrolysis control (a dried preparation of N-AcAA conjugate standards) and a derivatisation control (a dried preparation of DL-amino acids standards).

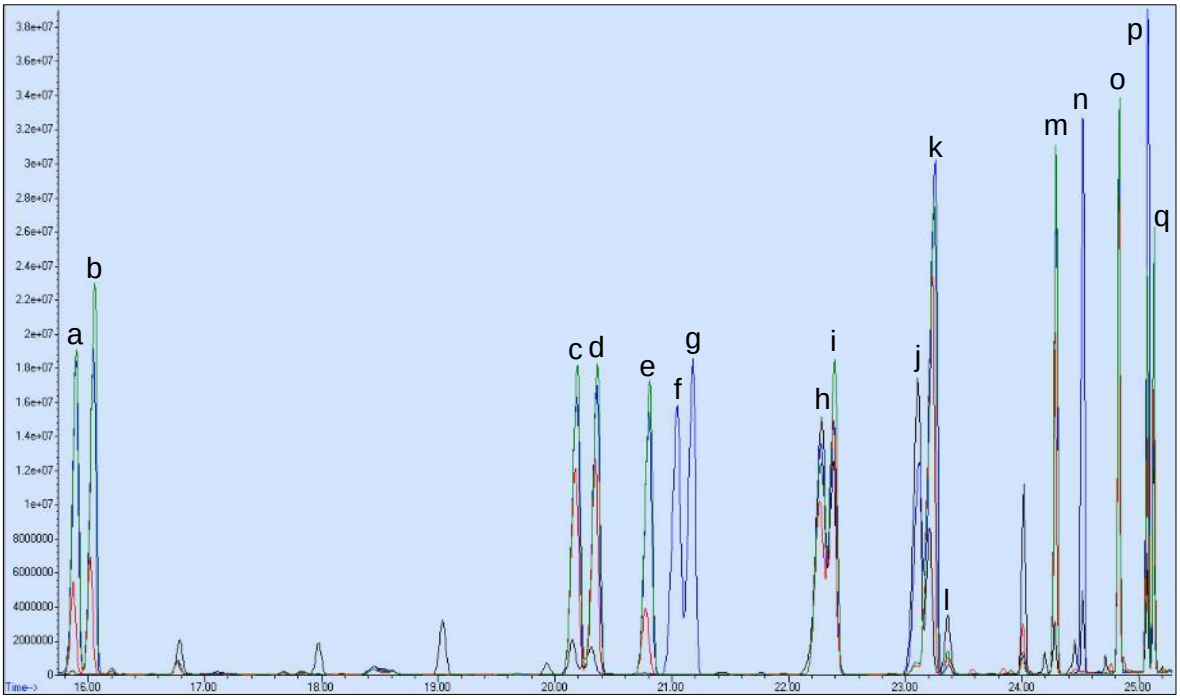

Figure 52: Branched-chain amino acid enantiomer composition of sample 11 and an overlay of controls.

The black chromatogram is that of sample 11, the blue is that of the derivatisation control, the red is that of the extraction control and the green is that of the hydrolysis control. Note that the double blue peaks without a corresponding peak from any of the other controls or sample 11 is Nva, the amino acid internal standard. The compound list is as follows: (a) L-Ala, (b) D-Ala, (c) L-Val, (d) D-Val, (e) the di-tBDMS ester of Ala, (f) L-Nva, (g) D-Nva, (h) L-Leu, (i) D-Leu, (j) L-aIle, (k) D-aIle and L-Ile, (l) D-Ile and (m–q) the di-tBDMS esters of Val, Nva, Leu, aIle and Ile, respectively. The y-axis (abundance) is scaled to 100% of (p) and the x-axis (retention time in minutes) is cropped to between 15.7 and 25.3 min to show only the section of interest.

For clarity, Figure 53 to Figure 57 respectively show a single ion extract of each amino acid pair of sample 11.

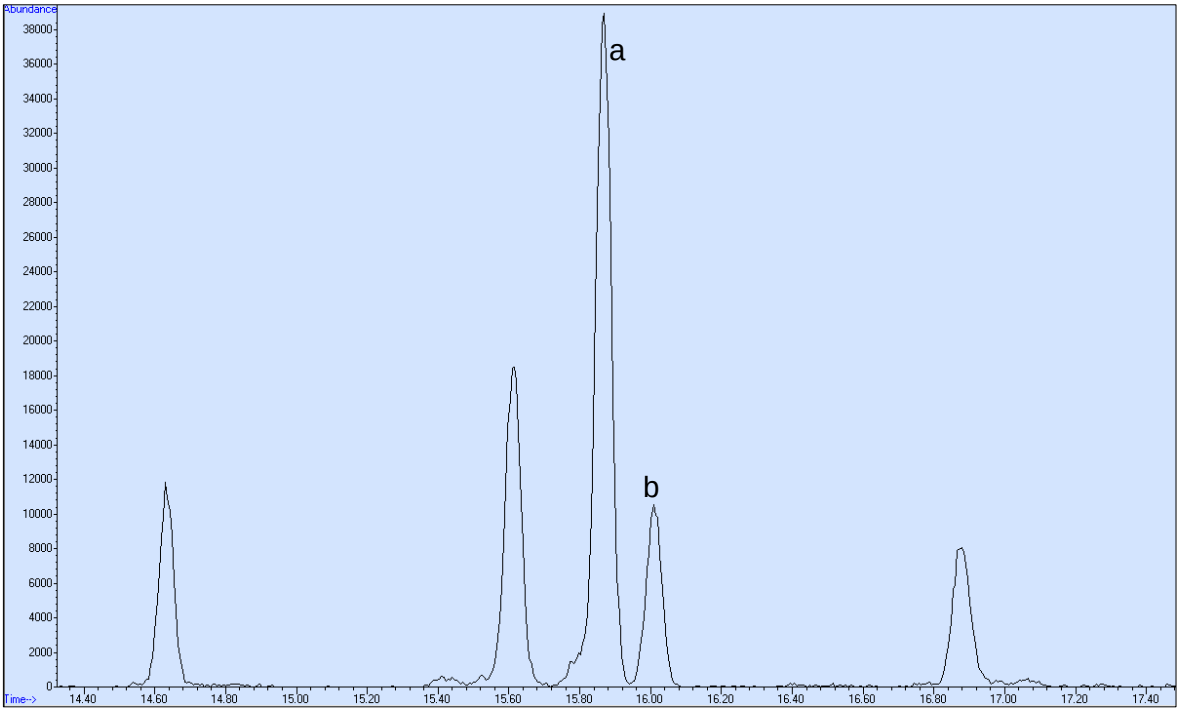

Figure 53: Sample 11, Ala.

The ion extract (m/z 158.1) of (a) L-Ala and (b) D-Ala. The y-axis (abundance) is scaled to 100% of (a) and the x-axis (retention time in minutes) is cropped to between 14.4 and 17.3 min to show only the section of interest.

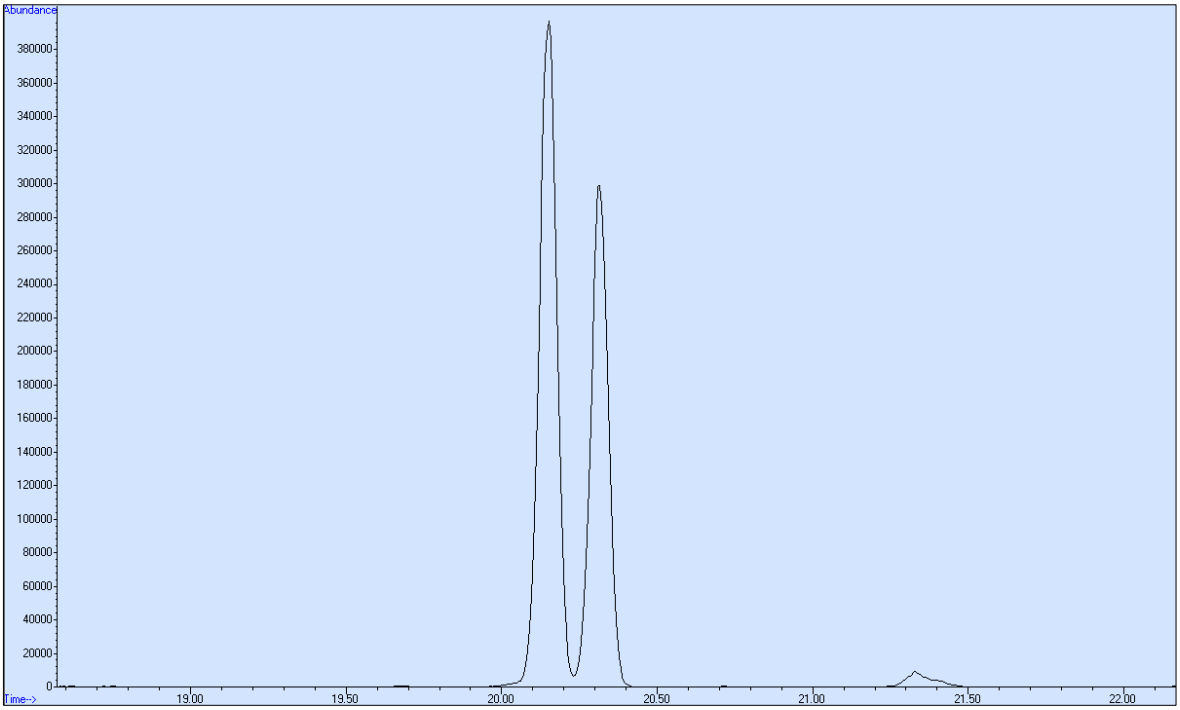

Figure 54: Sample 11, Val.

The ion extract (m/z 186.2) of L-Val (left) and D-Val (right). The y-axis (abundance) is scaled to 100% of L-Val and the x-axis (retention time in minutes) is cropped to between 18.5 and 22.2 min to show only the section of interest.

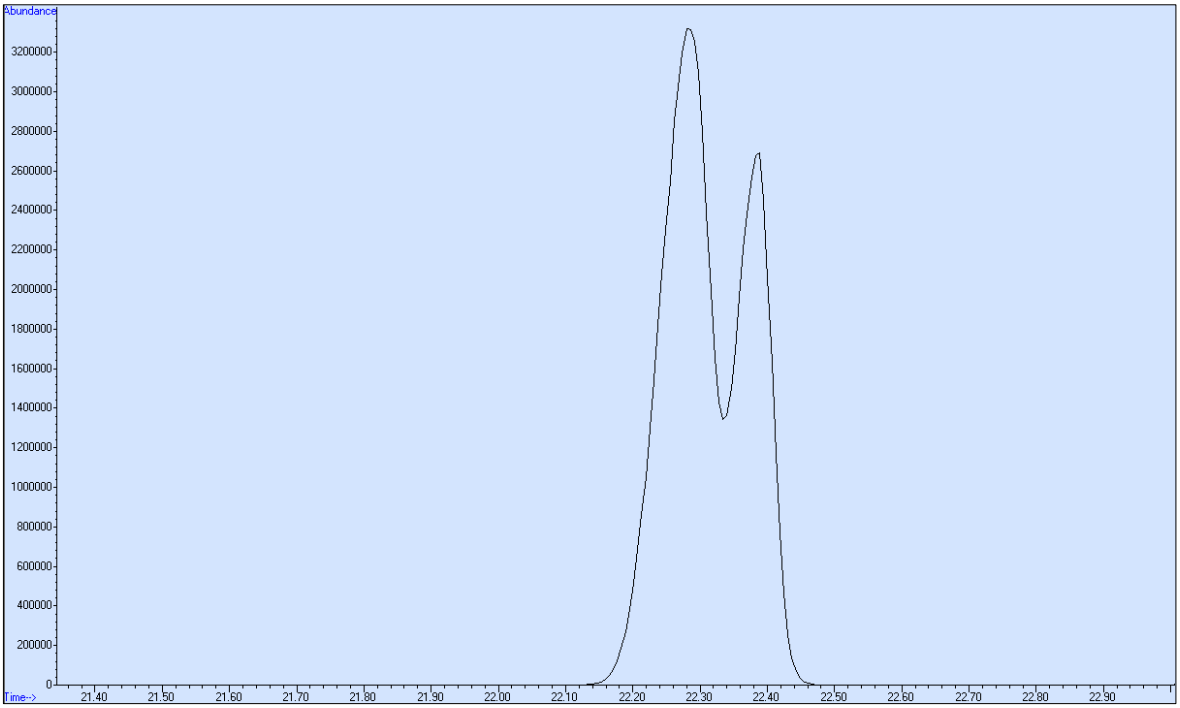

Figure 55: Sample 11, Leu.

The ion extract (m/z 200.2) of L-Leu (left) and D-Leu (right). The y-axis (abundance) is scaled to 100% of L-Leu and the x-axis (retention time in minutes) is cropped to between 21.3 and 23.0 min to show only the section of interest.

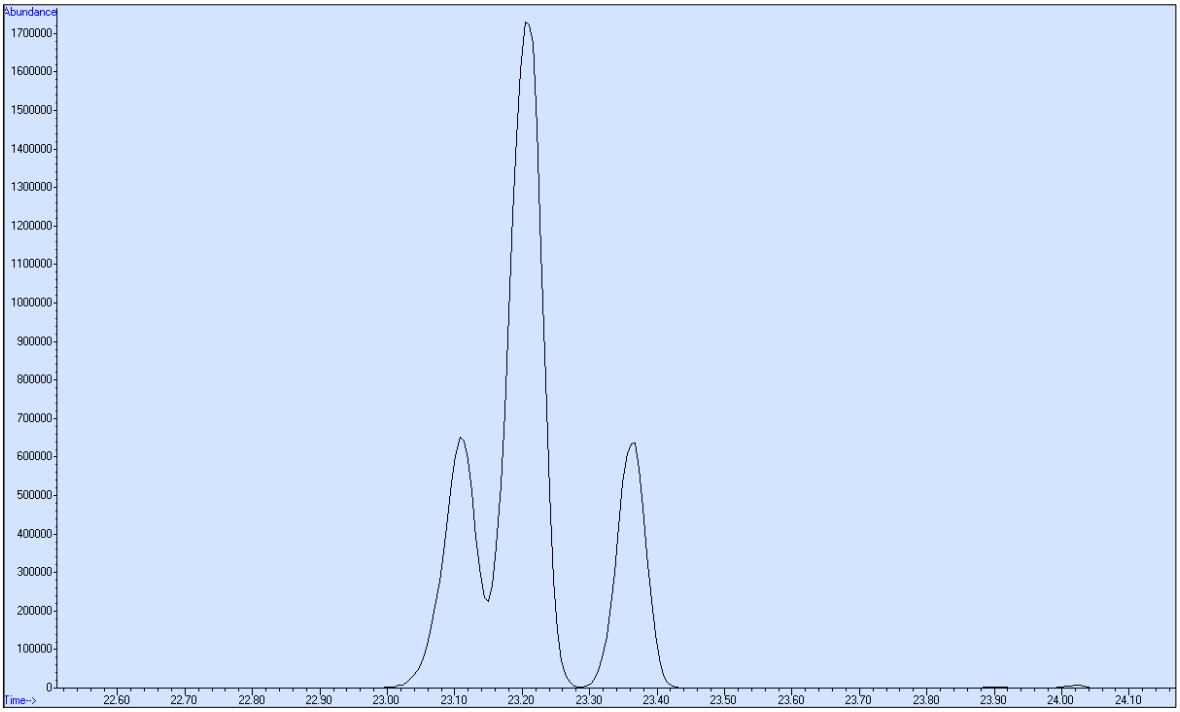

Figure 56: Sample 11, Ile and aIle.

The ion extract (m/z 200.2) of L-aIle (left), D-aIle and L-Ile (middle) and D-Ile (right). The y-axis (abundance) is scaled to 100% of the middle-peak and the x-axis (retention time in minutes) is cropped to between 22.5 and 24.1 min to show only the section of interest.

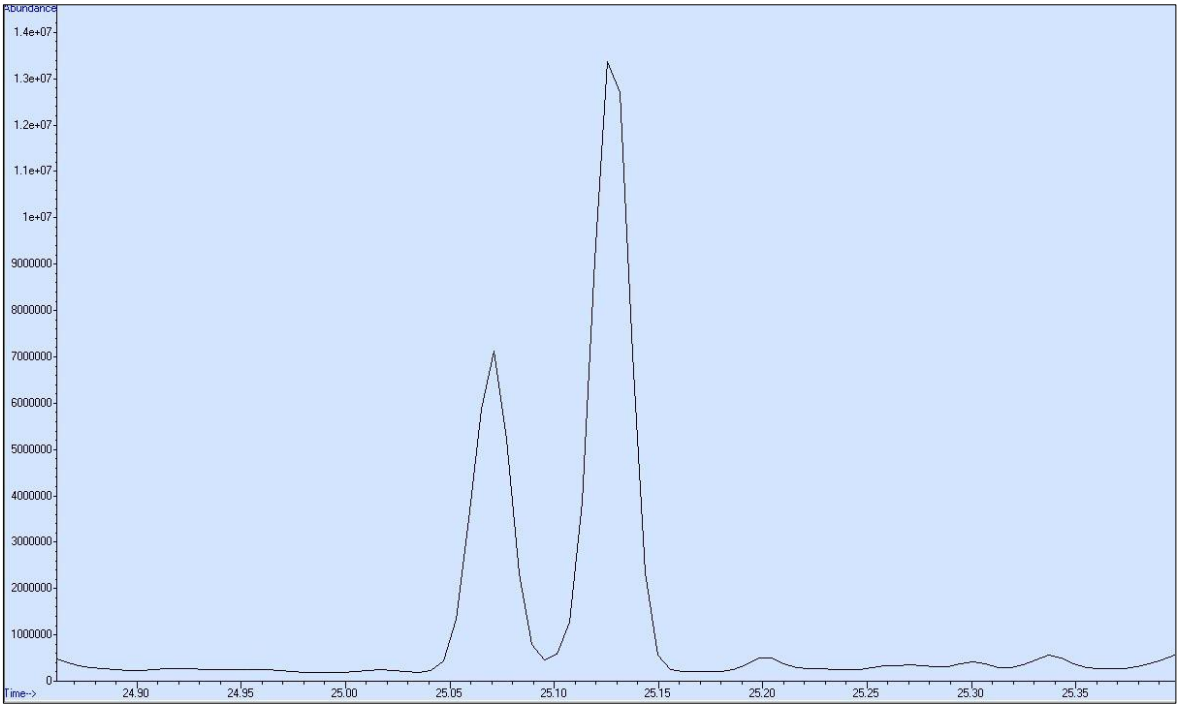

Figure 57: Sample 11, Ile and aIle – supportive di-tBDMS esters.

Total ion chromatogram of the di-tBDMS esters of DL-aIle (left) and DL-Ile (right). The y-axis (abundance) is scaled to 100% of DL-Ile and the x-axis (retention time in minutes) is cropped to between 24.8 and 25.4 min to show only the section of interest.

The results above indicate that the analytical strategy was a success and that the N-AAA conjugates observed in sample 11 consisted of both the L- and D-enantiomers. The latter also indicated that these conjugates originated from a non-enzymatic reaction. Whether this happens *in vivo* or *in vitro* merits further discussion following the results from molecular modelling below.

5.5 Phase 5: Molecular modelling to investigate the non-enzymatic formation of N-AAA conjugates

Refer to Figure 58: Due to the high cost of molecular modelling (processing time on the cluster), only the reaction mechanism of the simplest structural analogue of the N-Ac-BCAAs, N-Ac-Ala formed from pyruvate, was calculated. R^1 and R^2 in the figure thus represent only a methyl group, but the principles could be applied to other 2-keto acids as well. Preliminary results indicated that an acid catalyst was required for the reaction to proceed (this was also indicated in Figure 25). Weak acids, such as carbonic acid, resulted in very high E_a, whereas stronger acids, such as HCl and especially the hydronium ion, produced more acceptable results. Since a strong acid, such as HCl, is almost completely converted to the hydronium ion when diluted in water (Clever, 1963:637), the remainder of the modelling process was focused on the hydronium ion as the representing Brønsted acid. A TS was found for the protonation of

pyruvate (refer to Figure 59 and Figure 60). In addition, the following were found to proceed via two steps:

(1) Reaction step 3 (3-A and 3-B as indicated in Figure 58): When lengthening the bond between the hydrogen and nitrogen, simulating an acid-catalyst removing the hydrogen, the hydrogen instead moves over to the hydroxy group. This results in a water molecule leaving the compound instead of a just a hydrogen atom. In the next step (3-B), the acid-catalyst removes a hydrogen atom from a water molecule resulting in a hydroxide ion. The latter can then execute a nucleophilic attack on the beta carbon of #G1, which in turn rearranges its electrons to alleviate the charge on the nitrogen atom. This converts the beta carbon to a sp^3 hybridised state. The more electronegative nitrogen is then ready for a nucleophilic attack on the C2 carbon of #I.

(2) Reaction step 5 (5-A and 5-B as indicated in Figure 58): Rather than two units of catalysts removing a hydrogen and hydroxy group simultaneously, a single catalyst first removes the N-H hydrogen, with the remaining electrons alleviating the positive charge on the nitrogen atom, followed by a second catalyst removing a hydroxy group from a beta carbon in the following step. The possibility of the N-H hydrogen instead moving over to a Cβ-hydroxy group, resulting in a water molecule leaving the molecule (instead of a two-step acid catalyst), was also investigated. This seemed probable due to the close proximity of the latter two groups in a geometrically optimised structure, but a stable PES scan was not obtained.

(3) Reaction step 6 (6-A and 6-B as indicated in Figure 58): After the acid-catalyst removes the hydrogen from the carboxy group, the electron first stabilises on the oxygen atom before the molecule decarboxylates. The rest of the molecule undergoes rearrangement ending with a negatively charge oxygen on the remaining carboxyl group.

In step 7, the final rearrangement was found to occur without the aid of a catalyst as shown in Figure 58.

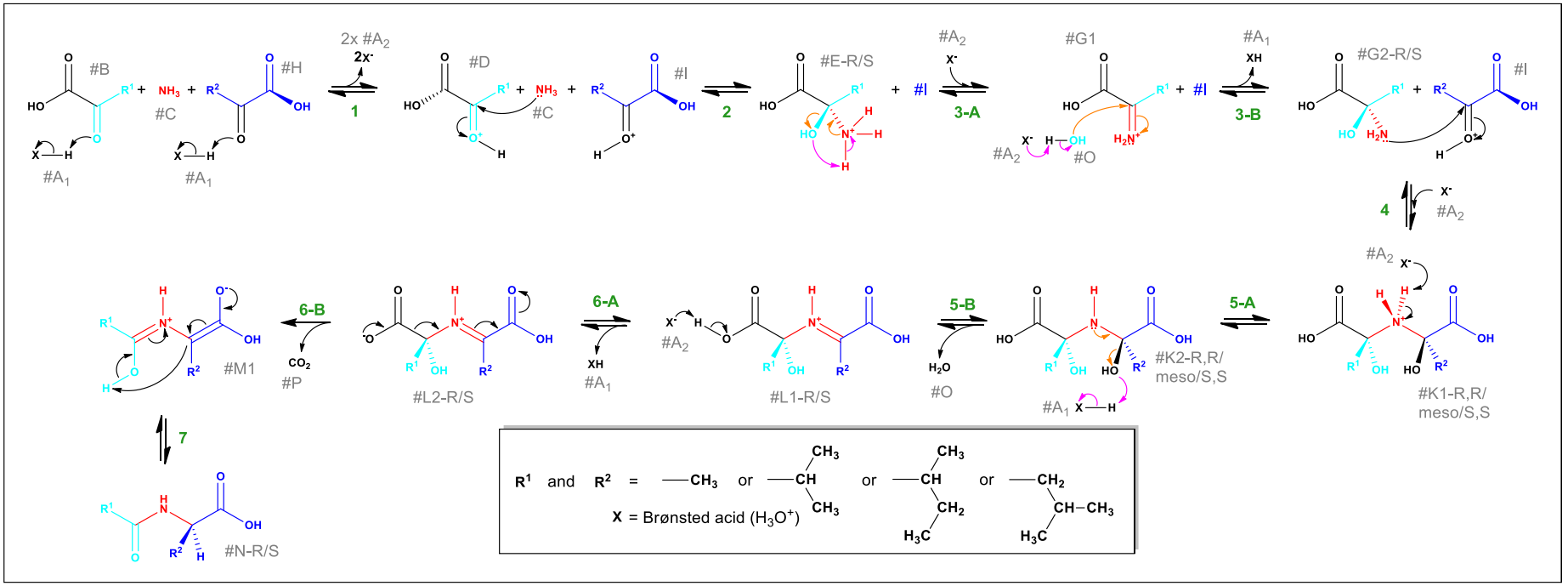

Figure 58: Newly proposed reaction mechanism for the non-enzymatic formation of N-AAA conjugates, constructed from molecular modelling observations.

The red, light-blue and dark-blue coloured groups correlate to the net reaction shown in Figure 23. The reaction steps are numbered in green, while reaction intermediates are labelled with the grey letters preceded by a hash symbol. The symbol 'HX' is used to indicate the Brønsted acid, H_30^+. Were colored arrows are used (pink and orange) the reaction may possibly proceed via 2 separate steps. As a side note: The meso*-compounds indicated above (#K1 and #K2) are not true* meso*-compounds. Even though a C_2 symmetry-axis is apparent in a Natta-projection or when building a physical three-dimensional model of the compounds, intra-molecular repulsion forces between atoms, angular torsion and hydrogen bonding cause a change in conformation after geometric optimisation, which nullifies the plane of symmetry in the optimised structures.*

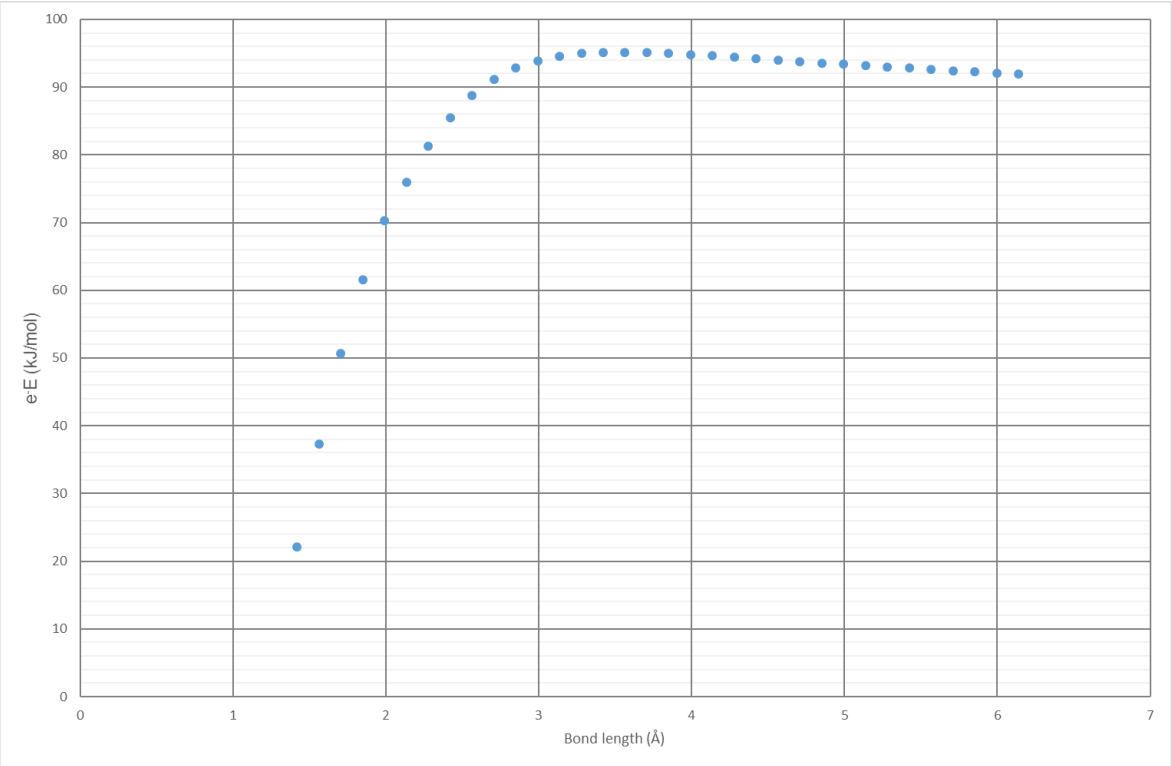

Figure 59: Potential energy surface scan for the protonation of pyruvate.

The figure above was generated as follows: Starting at #D (refer to Figure 58) the O-H bond on the oxonium ion was constrained. This bond length (x-axis) was then increased incrementally with geometric optimisation in-between. The resulting E_e of the geometric optimisation is plotted on the y-axis. The first derivative of the slope approaches zero at around 3.57 Å, indicating a possible TS at this point. E_e: Electron energy.

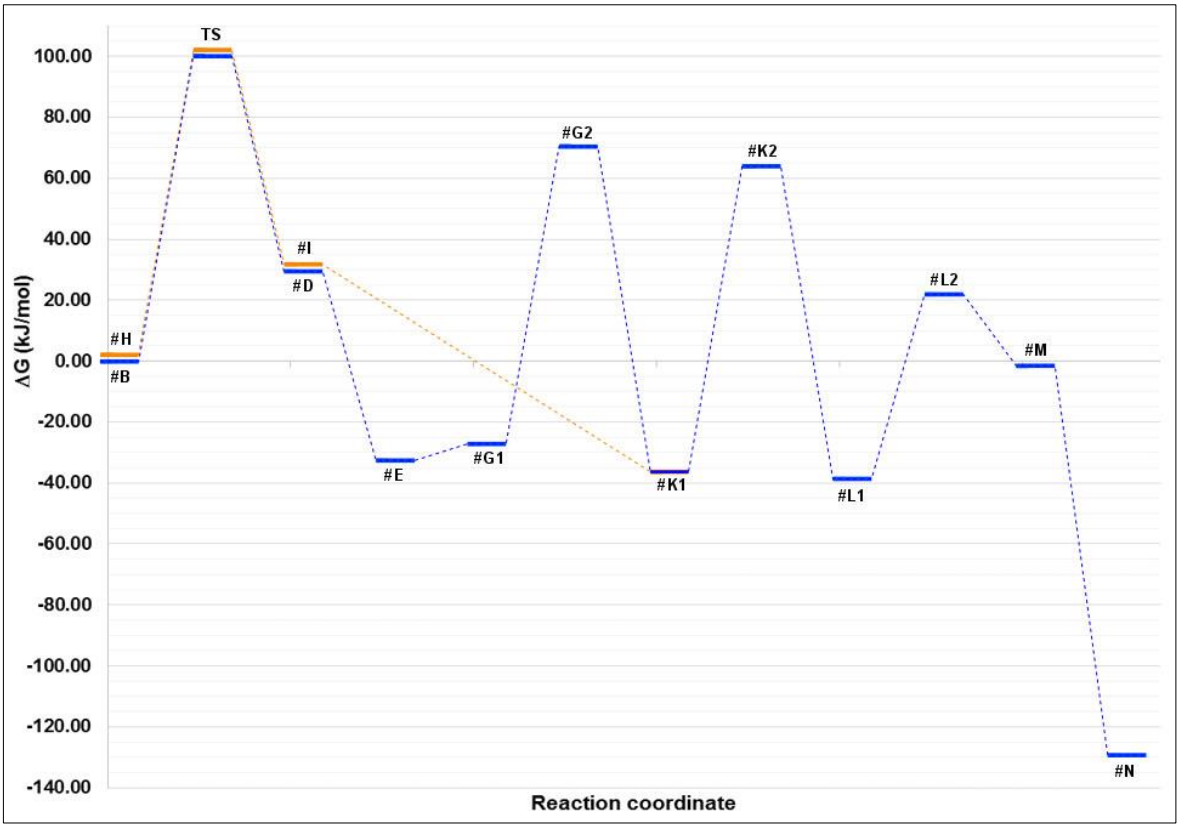

Figure 60: Energy profile for the non-enzymatic formation of N-Ac-Ala from pyruvate and ammonia.

To simplify the figure above, each alpha-numeral represents the most significant reaction intermediate in a balanced pool of reaction intermediates (refer to Table 14 and Figure 58). The orange lines represent the reaction path of the second pyruvate molecule, starting with a TS during protonation and ending with the condensation reaction with #G2 to form #K1. These values are intentionally slightly offset to show the blue line underneath. ΔG: Gibbs free energy.

Table 14: Composition and energy of each reaction pool in the final balanced reaction mechanism.

Representing intermediate	Pool content	Total energy of reaction pool (kJ/mol)
#B and #H	2x #A, #B, #H and #C	0.00
TS	TS (#B-#D, #H-#I and #A-#O) and #C	100.11
#I and #D	2x #O, #D, #I and #C	29.60
#E	2x #O, #E and #I	-32.69
#G1	3x #O, #G1 and #I	-27.21
#G2	#A, #O, #G2 and #I	70.69
#K1	#A, #O and #K1	-36.37
#K2	2x #A and #K2	63.87
#L1	#A, 2x #O and #L1	-38.62
#L2	2x #A, #O and #L2	21.92
#M	2x #A, #O, #M and #P	-1.86
#N	2x #A, #O, #N and #P	-129.48

Refer to Figure 58 for alpha-numerals preceded by hash simbols. TS: Transition state.

The results in Figure 60 indicate a slightly higher E_a when compared to the E_a values observed in the literature (refer to Table 6). Given the energy requirements of most enzymatic reactions, the E_a obtained for the non-enzymatic formation of N-Ac-Ala was still within an acceptable range of what could be expected of an *in vivo* non-enzymatic reaction. N-AAA conjugates may also form *in vitro*, which may be exposed to higher temperatures during sample preparation (up to 75 °C), keeping in mind that a higher E_a does not prevent a reaction from occurring but only slows the rate of the reaction (Keller *et al.*, 2015:153). This correlates to the trace amounts of N-AAA conjugates that were formed in the non-enzymatic *in vitro* experiment. Yanagawa *et al.* (1982:2089) also reported that slow hydrolysis occurs during the formation of N-oxalyl-Gly, which may have also contributed to the low yields.

CHAPTER 6: CONCLUSION

6.1 The importance of enzymatic versus non-enzymatic reactions

A summary of key concepts that are apparent from the literature is as follows: The metabolic significance of N-AAA conjugates is clear from its involvement in protein function (Perrier *et al.*, 2005:674-677), urea cycle regulation (Caldovic & Tuchman, 2003:281-285), myelin synthesis (Wijayasinghe *et al.*, 2014:4970), biotransformation and medicine (Mazaleuskaya *et al.*, 2015:416-417). The urinary excretion of N-AAA conjugates is also associated with various IEMs as indicated in Table 1. The enzymatic formation of N-AAA conjugates can be attributed to acyltransferases with acyl-CoAs and amino acids as substrates (Table 2), whereas the breakdown is catalysed by aminoacylases (Perrier *et al*., 2005:677 679). These enzymatic reactions use/yield only the L-amino acid enantiomer (Cava *et al.*, 2010:818). The non-enzymatic formation of some N-AAA conjugates was described by Yanagawa *et al*. (1982), whereas Dry (1997), Deysel (2001) and Mienie *et al*. (2001) investigated N-AAA conjugates in MSUD.

It is important to distinguish between enzymatic and non-enzymatic reactions due to the various influences each has on the metabolism as described in section 2.4.4. In summary, enzymatic reactions are highly selective and subjected to rigorous regulating mechanisms ensuring homeostasis and targeted product formation. Non-enzymatic reactions proceed unchecked and can result in harmful metabolite accumulation, but some non-enzymatic reactions are supplementary and even essential to life. This study successfully demonstrated three approaches to differentiate enzymatic and non-enzymatic reactions: an analytical strategy (determining chiral composition), chemical synthesis in the absence of enzymes and molecular modelling to confirm sufficiently low energy requirements in the applicable environment and to validate proposed reaction mechanisms. The data also support this methodology being applied in other cases to assess biological reactions.

6.1.1 Effectiveness of the chiral analytical approach and care required in each step

The general stereo-nonspecific nature of non-enzymatic reactions and the phenomenon of racemic mixtures being the product of prochiral substrates (McMurry, 2004:295-297) have enabled this approach. By determining the enantiomeric composition of target BCAAs, the origin (enzymatic/non-enzymatic) of the reported N-AAA conjugates was resolved. The results presented in Chapter 5 illustrate the effectiveness of the chiral strategy. The successful execution of the chiral analytical strategy depends on the necessary detail given in the steps to follow: N-AAA conjugates were extracted from urine without simultaneous extraction of native

BCAAs. This ensured that native BCAAs would not account for pseudo-enantiomeric excess. The N-AAA conjugates were hydrolysed without significant hydrolysis-induced racemisation, which could also have accounted for a pseudo-enantiomeric excess if left unchecked. The extracted N-AAA conjugates were converted to volatile diastereomers for separation via GC-MS, again, without inducing racemisation during derivatisation. With this qualitative method, the enantiomeric composition (enantiopure/racemic) of N-AAA conjugates could effectively be determined, indicating the origin of the said conjugates (enzymatic/non-enzymatic)

6.2 *In vitro* synthesis at near-physiological conditions

The low pH required for the *in vitro* non-enzymatic synthesis of the N-AAA conjugates makes *in vivo* occurrence improbable, but matches the *in vitro* reaction conditions during sample preparation for an organic acid analysis. Yanagawa *et al.* (1982:2088) synthesised N-Ac-Ala non-enzymatically at a temperature of 27°C and a more acceptable pH of 4 (also with low yields), with slightly better results at higher temperatures. This suggests that the *in vivo* non-enzymatic formation of N-AAA conjugates cannot yet be ruled out and further investigation is warranted. Since the objective of this study was only to attempt synthesis in an applicable *in vitro* environment, this was not pursued. Even though only low amounts of N-AAA conjugates were synthesised in this study, the presence of the conjugates in an environment devoid of enzymes proves that a non-enzymatic reaction does take place. Furthermore, the proposed net reaction did succeed in predicting the product in each case where a product was formed, confirming the basis of the proposed reaction mechanism. It is also worth pointing out that this approach is only valid if the appropriate starting substrates are selected (i.e. non-enzymatic synthesis of N-AAA conjugates from acyl-chlorides and amino acids would not have proved anything in this case, since acyl-chlorides are not metabolites).

6.3 Agreement between molecular modelling and analytical findings

With the commencement of the molecular modelling approach, many educated assumptions needed to be made. It is therefore crucial that the relevant expertise and known principles of organic chemical reactivity be applied. Starting with a poorly compiled starting setup will yield inaccurate results. It is also important to continually learn from the results and reapply new knowledge as the process progresses. In conclusion, proper planning, care and re-evaluation is crucial for an effective molecular modelling approach. In this study, molecular modelling was successfully used to validate the proposed reaction mechanism and to determine the E_a requirements. The calculated E_a was relatively high (in the upper range of typical *in vivo* enzymatic reactions), but agreed with the results of the chemical synthesis (low yields). These results (1) support the probability of a slow *in vivo* reaction, which may be sped up by another

(yet to be determined) catalyst or more appropriate reaction conditions or (2) indicate that an acceptably low amount of energy is required for formation during normal sample handling (including sample preparation). The results do not prove or disprove the non-enzymatic reaction occurring *in vivo* or *in vitro*, but support the hypothesis of N-AAA conjugates forming non-enzymatically in urine.

6.4 Limitations

All three approaches (analytical, synthesis and molecular modelling) have their limitations, but with the proper care and attention to the above-mentioned factors, these limitations seem unlikely to affect the final result.

6.4.1 Synthesis and analytical method: *In vivo* versus *in vitro* and qualitative versus quantitative

As mentioned above, it is currently not possible to determine whether the non-enzymatic reaction occurs *in vivo* or *in vitro* in this MSUD case. As noted in section 5.1.3, the large amounts of hydantoin conjugates observed in the *in vitro* synthesis (with urea as a nitrogen source) depleted the available 2-keto acids. This shows a more favourable reaction between 2-keto acids and urea than between 2-keto acids and ammonia under the reaction conditions used in this study. Noting that the reaction conditions do not necessarily represent the conditions of the sample, these findings can still be speculated about. With urine containing large amounts of urea, the formation of N-AAA conjugates after urine filtration seems unlikely. In addition, the N-AAA conjugates observed in the urine of this MSUD case study were at much higher concentrations than in the *in vitro* synthesis experiment and *vice versa* for the hydantoin conjugates. This suggests the non-enzymatic formation of the N-AAA conjugates in the blood, which is then concentrated in the urine. This would also limit the formation of the hydantoin conjugates – as was observed in the urine of sample 11.

A possible solution would have been to analyse the enantiomeric composition of the N-AAA conjugates directly in a blood sample, which would have at least shown whether the conjugates would have formed in circulation or after urine filtration. Unfortunately, a blood sample was not available and one could not be obtained for this study. A pitfall of this approach would be the relatively low concentrations of these conjugates in blood. Another limitation is the qualitative nature of the method. Currently the method only shows whether an N-AAA conjugate is present in an enantiopure or a racemic mixture, but the exact ratio could not be determined. However, the initial aim of this study was only to distinguish between enzymatic and non-enzymatic

reactions. The mere presence of significant amounts of the D-enantiomers in the urine sample confirms the non-enzymatic reaction.

6.4.2 Molecular modelling approach and why the limitations are unlikely to have influenced the final outcome

There are three main approaches when modelling molecules. In gas-phase simulations, molecules are simulated in a vacuum environment, calculations are quick and trends can be derived, but calculated energy values may be inaccurate. In the second approach, the effect of a solvent is estimated using an empirical mathematical expression, called an implicit solvation simulation. Although this provides more refined energy values and more accurate tendencies, energy released to the environment is not accounted for. The implicit solvation simulation used in this study was the COSMO CSM. A third approach is the explicit simulation, also called a periodic system, where the solvent molecules are included in the simulations. The most accurate tendencies and energy values are hereby obtained, but this laborious approach has the drawback of increased calculation time. Since the second aim of this study was only to verify the feasibility of the proposed reaction mechanism, an implicit solvation simulation was selected as an acceptable compromise between calculation time and adequate accuracy for the intended purpose (Leach, 2001:610). Refer to Annexure E for crystallography data supporting the selected model, indicating an acceptable agreement between the selected model and expected values.

An energy profile constructed from compounds in their local minima may result in slightly lower E_a values and a slightly less accurate energy change of the reaction than one constructed from compounds in their global minima. A search for the global minimum of a complex compound is possible but extremely time-consuming (since complex compounds may have multiple minima) and was thus not attempted in this study. Instead, a geometric optimisation that resulted in a local minimum was accepted, but only if the conformation of the reaction intermediate was on the correct reaction path (i.e. a conformation that could participate in the next step of the reaction in terms of stearic hindrance and proximity of atoms). This ensured selection of configurations in local minima with acceptably low energy values (significantly lower than those of compounds without any geometric minimisation) and configurations that made sense in terms of the reaction coordinate. A reaction intermediate may need to pass through a local minimum (instead of a global minimum) to stay on the correct reaction path.

6.4.3 Generalisation limits

The results reported in this study are in no way generalizable to MSUD at this stage, but are certainly generalizable to the investigation of unexpected/novel compounds in future cases. The intention was to demonstrate that the specific methodology can be utilised to distinguish between enzymatic and non-enzymatic reactions. As noted in section 3.4.2, the limited sample size of this study unfortunately limited deductions that can be made from any findings, but at least proved useful to demonstrate the method. Future samples containing N-AAA conjugates, or other chiral products not related to an enzymatic origin, may potentially be subjected to the methods of this study.

6.5 Recommendations on further research

Given the outcome if this study, a number of new research questions and possibilities arose:

(1) Why was this metabolic profile only reported in the three South African MSUD cases? Could these N-AAA conjugates be artefacts of improper sample handling, bacterial contamination or other non-human metabolic factors and/or catalysts? This may be elucidated by determining the origin of the non-enzymatic formation (*in vivo* or *in vitro*) of the N-AAA conjugates.

(2) What is the functionality and clinical implication of the unique and/or elevated levels of N-AAA conjugates? Do the described N-AAA conjugates act as substitutes for natural N-AAA conjugates? For example, N-Ac-Glu is an allosteric activator of the urea cycle. Propionyl-CoA and BCAA-related CoAs have been shown to inhibit the urea cycle by acting as substitutes for N-Ac-Glu (Dercksen *et al.*, 2012). This may lead to compromised ammonia detoxification and may also be of interest in the above patients. Further studies into N-AAA conjugates' influence on the initial steps of the urea cycle are therefore recommended.

(3) Could the observed hydantoin conjugates also form non-enzymatically from 2-keto acids and urea (instead of ammonia) via a similar mechanism? The results presented in section 5.1.3 already substantiates this possibility and a preliminary reaction mechanism is proposed in Figure 39. Further investigation, chiral analysis and molecular modelling are therefore warranted.

In conclusion, all three methods (chiral analysis, synthesis and molecular modelling) proved to be effective ways of assessing whether a reaction proceeds enzymatically or non-enzymatically.

The combination of all three methods supported the results of the others and amounted to the same final conclusion: The non-enzymatic formation of N-AAA conjugates from 2-keto acids and ammonia.

REFERENCE LIST

Ævarsson, A., Chuang, J.L., Wynn, R.M., Turley, S., Chuang, D.T. & Hol, W.G.J. 2000. Crystal structure of human branched-chain α-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure,* 8(3):277-291.

Ævarsson, A., Seger, K., Turley, S., Sokatch, J.R. & Hol, J.G.W. 1999. Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes. *Nature structural biology,* 6(8):785-792.

Aksoy, Y., Ögüs, I.H. & Özer, N. 2001. Purification and some properties of human placental glucose-6-phosphate dehydrogenase. *Protein expression and purification*, 21(2)286-292.

Badenhorst, C.P., Jooste, M. & Van Dijk, A.A. 2012. Enzymatic characterization and elucidation of the catalytic mechanism of a recombinant bovine glycine N-acyltransferase. *Drug metabolism and disposition: the biological fate of chemicals*, 40(2):346-352.

Bauld, N.L. 2001. Cyclohexane conformational analysis. http://research.cm.utexas.edu/nbauld/teach/cyclohex.html Date of access: 26 March 2016.

Badenhorst, C.P.S., Van der Sluis, R., Erasmus, E. & Van Dijk, A.A. 2013. Glycine conjugation: importance in metabolism, the role of glycine N-acyltransferase, and factors that influence interindividual variation. *Expert opinion on drug metabolism & toxicology,* 9(9):1139-1153.

Bartell, L.S. 1959. On the length of the carbon-carbon single bond. *Journal of the American chemical society*, 81(14):3497-3498.

Blau, N., Duran, M. & Gibson, K.M. 2008. Laboratory guide to the methods in biochemical genetics. 1st ed. Heidelberg: Springer-Verlag.

Bremer, H.J., Duran, M., Kamerling, J.P., Przyrembel, H. & Wadman, S.K. 1981. Disturbances of amino acid metabolism: Clinical chemistry and diagnosis. Baltimore: Urban & Schwarzenberg.

BRENDA (Braunschweig enzyme database). 2016a. Information on EC 3.5.1.15 - aspartoacylase. http://www.brenda-enzymes.org/enzyme.php?ecno=3.5.1.15 Date of access: 10 Apr. 2016.

BRENDA (Braunschweig enzyme database). 2016b. Information on EC 3.5.1.14 - N-acyl-aliphatic-L-amino acid amidohydrolase. http://www.brenda-enzymes.org/enzyme.php?ecno=3.5.1.14 Date of access: 10 Apr. 2016.

BRENDA (Braunschweig enzyme database). 2016c. Information on EC 3.5.1.114 - N-acyl-aromatic-L-amino acid amidohydrolase. http://www.brenda-enzymes.org/enzyme.php?ecno=3.5.1.114 Date of access: 10 Apr. 2016.

Caldovic, A. & Tuchman, M. 2003. N-Acetylglutamate and its changing role through evolution. *Biochemical journal,* 372(2):279-290.

Cava, F., Lam, H., De Pedro, A. & Waldor M.K. 2011. Emerging knowledge of regulatory roles of D-amino acids in bacteria. *Cellular and molecular life sciences*, 68(5):817-831.

Chalancon, G., Kruse, K. & Babu, M.M. 2013. Metabolic networks, structure and dynamics. (*In* Encyclopedia of systems biology, 2013:1263-1267).

Cherepanov, A.V. & De Vries, S. 2003. Kinetics and thermodynamics of nick sealing by T4 DNA ligase. *European journal of biochemistry*, 270(21):4315-4325.

Chuang, D.T., Shih, V.E. & Wynn, R.M. 2012. Maple syrup urine disease (branched-chain ketoaciduria). (*In* The online metabolic and molecular bases of inherited disease. http://ommbid.mhmedical.com/content.aspx?sectionid=62675436&bookid=971&jumpsectionID=62675470&Resultclick=2#1121417425 Date of access: 28 Jun. 2017).

Clever, H.L. 1963. The hydrated hydronium ion. *Journal of chemical education,* 40(12):637-641.

Craig, D.B., Arriaga, E.A., Wong, J.C.Y., Lu, W. & Dovichi, N.J. 1996. Studies on single alkaline phosphatase molecules: Reaction rate and activation energy of a reaction catalyzed by a single molecule and the effect of thermal denaturations: the death of an enzyme. *Journal of the American chemical society*, 188(22):5245-5253.

Dancis, J., Levitz, M., Miller, S. & Westall, R.G. 1959. Maple syrup urine disease. *British medical journal,* 1(5114):91-93.

Das, B., Daver, H., Pyrkosz-Bulska, M., Perscha, E., Barman, S.K., Mukherjee, R., Gumienna-Kontecka, E., Jarenmark M., Himo, F., & Nordlander, E. 2014. A dinuclear zinc(II) complex of a new unsymmetric ligand with an N_5O_2 donor set; A structural and functional model for the active site of zinc phosphoesterases. *Journal of inorganic biochemistry,* 132(1):6-17.

Dercksen, M., IJlst, L., Duran, M., Mienie, L.J., Van Cruchten A., Van der Westhuizen F.H. & Wanders, R.J.A. 2014. Inhibition of N-acetylglutamate synthase by various monocarboxylic and dicarboxylic short-chain coenzyme A esters and the production of alternative glutamate esters. *Biochimica et biophysica acta*, 1842(12):2510-2516.

Deysel, M.S.M. 2001. Biochemiese karakterisering van 'n Suid-Afrikaanse MSUD variant. Potchefstroom: Potchefstroomse universiteit vir christelike hoër onderwys. (Skripsie – Honns).

Dry, J. 1997. 'n Studie van geïnduseerde metaboliese weë weens 'n aangebore defek van die vertakte ketting-α-ketosuur dehidrogenase-kompleks. Potchefstroom: Potchefstroomse universiteit vir christelike hoër onderwys. (Verhandeling – M.Sc.)

Eldjarn, L., Jellum, E., Stokke, O., Pande, H. & Waaler, P.E. 1970. Beta-hydroxyisovaleric aciduria and beta-methylcrotonylglycinuria: a new inborn error of metabolism. *Lancet,* 2(7671):521-522.

Farrell, E.K. & Merkler, D.J. 2008. Biosynthesis, degradation, and pharmacological importance of the fatty acid amides. *Drug discovery today,* 13(14):558-568

Fidaleo, M. & Lavecchia, R. 2003. Kinetic study of enzymatic urea hydrolysis in the pH range 4-9. C*hemical and biochemical engineering quarterly,* 17(4):311-318.

Fiehn, O. 2001. Combining genomics metabolome analysis and biochemical modelling to understand metabolic networks. *Comparative and functional genomics,* 2(3):155-168.

Fourches, D., Muratov, E. & Tropsha, A. 2010. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. *Journal of chemical information and modeling*, 50(70):1189-1204.

Garret, R.H. & Grisham, C.M. 2005. Biochemistry. 3rd ed. Belmount: Thomson Brooks/Cole.

Gerlo, E., Van Coster, R., Lissens, W., Winckelmans, G., De Meirleir, L. & Wevers, R. 2006. Gas chromatographic–mass spectrometric analysis of N-acetylated amino acids: The first case of aminoacylase I deficiency. *Analytica chimica acta*, 571(2):191-199.

Gibson, K.M., Breuer, J., Kaiser, K., Nyhan, W.L., McCoy, E.E., Ferreira, P., Greene, C.L., Blitzer, M.G., Shapira, E., Reverte, F., Conde, C., Bagnell, P. & Cole, D.E.C. 1988. 3-Hydroxy-3-methylglutaryl-coenzyme A lyase deficiency: report of five new patients. *Journal of inherited metabolic disease,* 11(1):76-87.

Goodman, S.I., McCabe, E.R.B., Fennessey, P.V., Mace, J.W. 1980. Multiple acyl Co-A dehydrogenase deficiency (glutaric aciduria type II) with transient hypersarcosinemia and sarcosinuria; possible inherited deficiency of an electron transfer flavoprotein. *Pediatric research,* 14(1):12-17.

Gompertz, D., Saudubray, J.M., Charpentier, C., Bartlett, K., Goodey, P.A. & Draffan, G.H. 1974. A defect in L-isoleucine metabolism associated with alpha-methyl-beta-hydroxybutyric and alpha-methylacetoacetic aciduria: quantitative in vivo and in vitro studies. C*linica chimica acta,* 57(3)269.81.

Guder, C.M. 1998. Die gebruik van filtreerpapier vir die versending van urine vir die opsporing van aangebore defekte van die metabolisme van organiese sure. Potchefstroom: NWU. (Verhandeling – Honns).

Hagenfeldt, L., Bollgren, I. & Venizelos, N. 1967. N-acetylaspartic aciduria due to aspartoacylase deficiency-a new etiology of childhood leukodystrophy. *Journal of inherited metabolic disease,* 10(2):135-141.

Hagenfeldt, L. & Naglo, A.S. 1987. New conjugated urinary metabolites in intermediate type maple syrup urine disease. *Clinica chimica acta,* 169(1):77-83.

Harata, K., Sakabe, N. & Tanaka, J. 1976. Pyruvic acid. *Acta crystallographica,* B33(1977):210-212.

Heard, K.J. 2008. Acetylcysteine for acetaminophen poisoning. *The new England journal of medicine*, 359(3):285-292.

HMDB (Human metabolome database). 2016. Database statistics. http://www.hmdb.ca/statistics. Date of access: 05 Apr. 2016.

Haymond, M.W., Ben-Galim, E. & Strobel, K.E. 1978. Glucose and alanine metabolism in children with maple syrup urine disease*. Journal of clinical investigation,* 62(2):398-405.

Jellum, E., Horn, L., Thoresen, O., Kvittingen, E.A. & Stokke, O. 1986. Urinary excretion of N-acetyl amino acids in patients with some inborn errors of amino acid metabolism. *Scandinavian journal of clinical laboratory investment*, 1986(184):21-26.

Jensen, F. 1999. Introduction to computational chemistry. 1st ed. Chichester: John Wiley.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. & Barabási, A.L. 2000. The large-scale organization of metabolic networks. *Nature,* 407:651-654.

Kaiser, K. & Benner, R. 2005. Hydrolysis-induced racemization of amino acids. *Limnology and oceanography: Methods,* 3(8):318-325.

KEGG (Kyoto encyclopaedia of genes and genomes). 2016. Arginine biosynthesis. http://www.genome.jp/kegg-bin/show_pathway?ko00220 Date of access: 13 Oct. 2016.

Keller, M.A., Piedrafita, G., Ralser, M. 2015. The widespread role of non-enzymatic reactions in cellular metabolism. *Current opinion in biotechnology*, 34:153-161.

Kelmer-Bracht, A., Santos, C.P.B., Ishii-Iwamoto, E.L., Broetto-Biazon, A.C. & Bracht, A. 2003. Kinetic properties of the glucose 6-phosphatase of the liver from arthritic rats. *Biochimica et biophysica acta*, 1638(1):50-56.

Kinne-Saffran, E. & Kinne, R.K. 1999. Vitalism and synthesis of urea. From Friedrich Wöhler to Hans A. Krebs. *American journal of nephrology*, 19(2):290-294.

Kniffin, C.L. 2007. N-acetylglutamate synthase; NAGS. http://www.omim.org/entry/608300 Date of access: 13 Oct 2016.

Lambrecht, D. 2011. GC columns for enantiomer separation. http://www.mn-net.com/tabid/11660/default.aspx Date of access: 06 Nov 2015.

Lauterburg, B.H., Corcoran, G.B. & Mitchell, J.R. 1983. Mechanism of action of N-acetylcysteine in the protection against the hepatotoxicity of acetaminophen in rats *in vivo*. *The journal of clinical investigation*, 71(4):980-991.

Leach, A.R. 2001. Molecular modelling principles and applications. 2nd ed. Harlow: Pearson Education.

Lehnert W. & Werle, E. 1988. Elevated excretion of N-acetylated branched-chain amino acids in maple syrup urine disease. *Clinica chimica acta,* 172(1):123-126.

Liska, D., Lyon, M. & Jones, D.S. 2006. Detoxification and biotransformation imbalances. *Explore*, 2(2):122-140.

Londesborough, J. 1980. The causes of sharply bent or discontinuous arrhenius plots for enzyme-catalysed reactions. *European journal of biochemistry*, 105(2):211-215.

LoPachin, R.M., Gavin, T., Petersen, D.R. & Barber, D.S. 2009. Molecular mechanisms of 4-hydroxy-2-nonenal and acrolein toxicity: nucleophilic targets and adduct formation. *Chemical research in toxicology,* 22(9):1499-1508

Luykx, J.J., Bakker, S.C., Van Boxmeer, L., Vinkers, C.H., Smeenk, H.E., Visser, W.F., Verhoeven-Duif, N.M., Strengman, E., Buizer-Voskamp, J.E., De Groene, L., Van Dongen, E.P.A., Borgdorff, P., Bruins, P., De Koning, T.J., Kahn, R.S. & Ophoff. R.A. 2013. D-Amino acid aberrations in cerebrospinal fluid and plasma of smokers. *Neuropsychopharmacology,* 38:2019-2026.

Matalon, R., Michals, K., Sebesta, D., Deanching, M., Gashkoff, P. & Casanova, J. 1988. Aspartoacylase deficiency and N-acetylaspartic aciduria in patients with Canavan disease. *American journal of medical genetics*, 29(2):463-471.

Matschinsky, F.M., Zelent, B., Doliba, N., Li, C., Vanderkooi, J.M., Naji, A., Sarabu, R. & Grimsby, J. 2010. Glucokinase activators for diabetes therapy. *Diabetes care*, 34(2):236-243.

Mazaleuskaya, L.L., Sangkuhl, K., Thorn, C.F., FitzGerald, G.A., Altman, R.B. & Klein, T.E. 2015. Pathways of acetaminophen metabolism at the therapeutic versus toxic doses. *Pharmacogenetics and genomics,* 25(8):416-426.

McMurry, J. 2004. Organic chemistry. 6th ed. Belmount: Thomson Brooks/Cole.

Meister, A. & Abendschein, P.A. 1956. Chromatography of alpha-keto acid, 2,4-dinitrophenylhydrazones and their hydrogenation products. *Analytical chemistry*, 28 (2)171-173.

Menkes, J.H., Hurst, P.L. & Craig, J.M. 1954. A new syndrome: progressive familial infantile cerebral dysfunction associated with an unusual urinary substance. *Pediatrics*, 14(5):462-467.

MIA (Malahyde Information Systems). 2000. ACC™ 200 effervescent tablets. http://home.intekom.com/pharm/hexal/acc200.html Date of access 13 Oct. 2016.

Mienie, L.J., Deysel M.S., Dry, J., Knoll, D.P., Erasmus, E., Van der Westhuizen, F.H., Lippert, M.M., Duran, M. & Pretorius, P.J. 2001. MSUD patient exhibiting an unusual metabolic profile: Possible artefacts and bacterial products. (*In* Deysel, M.S.M., *red*. Biochemiese karakterisering van ’n Suid-Afrikaanse MSUD variant. Potchefstroom: Potchefstroomse universiteit vir christelike hoër onderwys. (Skripsie – Honns). p. 39-63)

NWU Centre for Human Metabolomics. 2017a. Work instruction for amino acid analysis. (Unpublished).

NWU Centre for Human Metabolomics. 2017b. Work instruction for creatinine analysis. (Unpublished).

NWU Centre for Human Metabolomics. 2017c. Work instruction for organic acid analysis. (Unpublished).

Patric, A.D. 1961. Maple syrup urine disease. *Archives of diseases in childhood*, 36:269-272.

Perrier, J., Durand, A., Giardina, T. & Puigserver, A. 2005. Catabolism of intracellular N-terminal acetylated proteins: involvement of acylpeptide hydrolase and acylase. *Biochimie,* 87(8)673-685.

Podlech, J. 2001. Origin of organic molecules and biomolecular homochirality. *Cellular and molecular life sciences,* 58(1):44-60.

Porcelli, A.M., Ghelli, A., Zanna, C., Pinton, P., Rizzuto. R. & Rugolo, M. 2005. pH difference across the outer mitochondrial membrane measured with a green fluorescent protein mutant. *Biochemical and biophysical research communications*, 236(4):799-804.

Rasmussen, K., Ando, T., Nyhan, W.L., Hull, D., Cottom, D., Donnell, G., Wadlington, W. & Kilroy, A.W. 1972. Excretion of propionylglycine in propionic acidaemia. *Clinical science,* 42(6):665-671.

Rinaldo, P., O'Shea, J.J., Coates, P.M., Hale, D.E., Stanley, C.A. & Tanaka, K. 1988. Medium-chain acyl-CoA dehydrogenase deficiency: diagnosis by stable-isotope dilution measurement of urinary n-hexanoylglycine and 3-phenylpropionylglycine. *The new England journal of medicine,* 319(18):1308-1313.

Rogers, K. 2011. The chemical reactions of life: from metabolism to photosynthesis. New York: Britannica.

Salway, J.G. 2004. Metabolism at a glance. 3rd ed. Malden: Blackwell.

Sass, J.O., Mohr, V., Olbrich, H., Engelke, U., Horvath, J., Fliegauf, M., Loges, N.T., Schweitzer-Krantz, S., Moebus, R., Weiler, P., Kispert, A., Superti-Furga, A., Wevers, R.A. & Omran, H. 2006. Mutations in ACY1, the gene encoding aminoacylase 1, cause a novel inborn error of metabolism. *The American journal of human genetics*, 78(3):401-409.

Schäfer, G. & Bode, J.W. 2014. Synthesis of sterically hindered N-acylated amino acids from N-carboxyanhydrides. *Organic letters*, 16(5):1526-1529.

Schlegel, H.B. 2011. Geometry optimization. *Wiley interdisciplinary reviews: Computational molecular* science, 1(5):790-809.

Schummer, C., Delhomme, O., Appenzeller, B.M., Wennig, R., & Millet, M. 2009. Comparison of MTBSTFA and BSTFA in derivatization reactions of polar compounds prior to GC/MS analysis. *Talanta,* 77(4):1473-1482.

Scrutton, M.C. 1974. Pyruvate carboxylase: studies of activator-independent catalysis and of the specificity of activation by acyl derivatives of coenzyme A for the enzyme from rat liver. *The journal of biological chemistry*, 249(22):7057-7067.

Snyderman, S.E., Norton, P.M., Roitman, E. & Holt, L.E. 1964. Maple syrup urine disease, with particular reference to dietotherapy. *Pediatrics,* 34(1):454-472.

Stashenko, E., & Martínez, J.R. 2014. Gas chromatography-mass spectrometry. (*In* Guo, X., *ed.* Advances in gas chromatography. *sine loco*: InTech p. 1-38).

Su, J., Kim, S. & Yan, Y. 2007. Dissecting the pretransitional conformational changes in aminoacylase I thermal denaturation. *Biophysical Journal*, 92(2):578-587.

Sweetman, L., Bates, S.P., Hull, D. & Nyhan, W.L. 1977. Propionyl-CoA carboxylase deficiency in a patient with biotin-responsive 3-methylcrotonylglycinuria. *Pediatric research*, 11(11):1144-1147.

Sweetman, L., Weyler, W., Nyhan, W.L., deCespedes, C., Loria, A.R. & Estrada, Y. 1978. Abnormal metabolites of isoleucine in a patient with propionyl-CoA carboxylase deficiency. *Biomedical mass spectrometry,* 5(3):198-207.

Tanaka, K, & Isselbacher, K.J. 1967. The isolation and identification of N-isovalerylglycine from urine of patients with isovaleric acidemia. *The journal of biological chemistry,* 242(12):2966-2972.

Van Coster, R.N., Gerlo, E.A., Giardina, T.G., Engelke, U.F., Smet, J.E., De Praeter, C.M., Meersschaut, V.A., De Meirleir, L.J., Seneca, S.H., Devreese, B., Leroy, J.G., Herga, S., Perrier, J.P., Wevers, R.A. & Lissens, W. 2005. Aminoacylase 1 deficiency: A novel inborn error of metabolism. *Biochemical and biophysical research communications*, 338(2005):1322-1326.

Sato, T., Tosa, T. 1993. Optical resolution of racemic amino acids by aminoacylase. (*In* Tanaka, A., Tosa, T. & Kobayashi, T., *ed*. Industrial application of immobilized biocatalysts. 1st ed. New York: Marcel Dekker. p. 3-14).

Schlegel, H.B. 2003. Exploring potential energy surfaces for chemical reactions: An overview of some practical methods. *Journal of computational chemistry*, 24(12):1514-1527.

Sigma-Aldrich. 2016. Basics of Chiral HPLC. https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Supelco/Posters/t408109h.pdf Date of access 13 Oct. 2016.

Tsirulnikov, K., Abuladze, N., Bragin, A., Faull, K., Cascio, D., Damoiseaux, R., Schibler, M.J. & Pushkin, A. 2012. Inhibition of aminoacylase 3 protects rat brain cortex neuronal cells from the toxicity of 4-hydroxy-2-nonenal mercapturate and 4-hydroxy-2-nonenal. *Toxicology and applied pharmacology,* 263(3):303-314.

Uttamsingh, V., Baggs, R.B., Krenitsky, D.M. & Anders, M.W. 2000. Immunohistochemical localization of the acylases that catalyze the deacetylation of N-acetyl-L-cysteine and haloalkene-derived mercapturates. *Drug metabolism and disposition,* 28(6):625.632.

Van Damme, P., Hole, K., Pimenta-Marques, A., Helsens, K., Vandekerckhove, J., Martinho, R.G., Gevaert, K. & Arnesen, T. 2011. NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. *PLOS genetics*, 7(7):1-19.

Van der Westhuizen, F.H., Pretorius, P.J. & Erasmus, E. 2000. The utilization of alanine, glutamic acid, and serine as amino acid substrates for glycine N-acyltransferase. *Journal of biochemical and molecular toxicology,* 14(2):102-109.

Van Slyke, D.D., Macfadyen, D.A., & Hamilton, P.B. 1943. The gasometric determination of amino acids in urine by the ninhydrin-carbon dioxide method. *The journal of biological chemistry,* 150(1):251-258.

Van Zyl, L.J., Schubert, W., Tuffin, M.I. & Cowan, D.A. 2014. Structure and functional characterization of pyruvate decarboxylase from *Gluconacetobacter diazotrophicus. BMC structural biology*, 14(21)1-13.

Vickers, T. 2015. Arabidopsis thaliana metabolic network [email]. 21 Sep 2015

Vockley, J., Zschocke, J., Knerr, I., Vockley, C.W. & Gibson, K.M. 2012. Branched chain organic acidurias. (*In* The online metabolic and molecular bases of inherited disease. http://ommbid.mhmedical.com/content.aspx?sectionid=62676787&bookid=971&jumpsectionID=62676789&Resultclick=2 Date of access: 28 Jun. 2017).

Ward, K.D. & Ward T.J. 2014. Recent progress in chiral stationary phase development and current chiral applications. *LCGC,* 32(4):20-23.

Wang, Y. & Chen, A. 2013. Crystallization-based separation of enantiomers. (*In* Andrushko, V. & Andrushko, N., *ed*. Stereoselective Synthesis of Drugs and Natural Products. Bognor Regis: John Wiley & Sons. p. 1663-1682)

Webster, L.T., Siddiqui, U.A., Lucas, S.V., Strong, J.M. & Mieyal, J.J. 1976. Identification of separate acyl-CoA:glycine and acyl-CoA:L-Gln N-acyltransferase activities in mitochondrial fractions from liver of rhesus monkey and man. *The journal of biological chemistry,* 251(11):3352-3358.

Wietmarschen, H. 2012. A systems approach to sub-typing of rheumatoid arthritis. Amsterdam: Leiden University. (Dissertation – MSc).

Wijayasinghe, Y.S., Pavlovsky, A.G. & Viola, R.E. 2014. Aspartoacylase catalytic deficiency as the cause of Canavan disease: a structural perspective. *Biochemistry,* 53(30):4970-4978.

Wilson, K. & Walker, J. 2005. Biochemistry and molecular biology. 6th ed. New York, NY: Cambridge university press.

Wolfender, J., Marti, G., Thomas, A. & Bertrand, S. 2015. Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of chromatography A,* 1382(2015):136-164.

Visser, W.F., Verhoeven-Duif, N.M., Ophoff, R., Bakker, S., Klomp, L.W., Berger, R. & de Koning T.J. 2011. A sensitive and simple ultra-high-performance-liquid chromatography-tandem mass spectrometry based method for the quantification of D-amino acids in body fluids. *Journal of chromatography A,* 1218(40):7130-7136.

Yanagawa, H., Makino, Y., Sato, K., Nishizawa, M. & Egami, F. 1982. Novel formation of α-amino acids and their derivatives from oxo acids and ammonia in an aqueous medium. *Journal of biochemistry*, 91(6):2087-2090.

Yoshioka, T. & Uematsu, T. 1993. Formation of N-hydroxy-N-arylacetamides from nitroso aromatic compounds by the mammalian pyruvate dehydrogenase complex. *Biochemical Journal,* 15(290):783-790.

Zorich, N., Jonas, A. & Pownall, H.J. 1985. Activation of lecithin cholesterol acyltransferase by human apolipoprotein E in discoidal complexes with lipids. *The journal of biological chemistry*, 260(15):8831-8837.

ANNEXURE A: MSUD BIOCHEMICAL IMAGE

Table 15: Urine biochemical image typical of maple syrup urine disease, summarised from Dry (1997:1-186) and other sources.

Metabolites	Status in MSUD	Additional references
Branched-chain amino acids (BCAAs)		
Isoleucine (Ile)	Elevated	Dancis *et al.*(1959:91)
Valine (Val)	Elevated	Dancis *et al.* (1959:91)
Leucine (Leu)	Highly elevated	Dancis *et al.* (1959:91)
allo-Ile	Elevated	Snyderman *et al.* (1964:454)
Other amino acids		
Glycine (Gly)	N.D.†	Dry (1997:1-186)
Alanine (Ala)	Deficient	Haymond *et al.* (1978:398)
Decarboxylated amines		
2-Methylbutylamine	N.D.	Dry (1997:1-186)
Isobutylamine	Trace	Dry (1997:1-186)
Isoamylamine	N.D.	Dry (1997:1-186)
2-Keto acids		
2-Keto-3-methylvaleric acid	Elevated	Dancis *et al.* (1959:91); Meister and Abendschein (1956:171)
2-Keto-isovaleric acid	Elevated	Dancis *et al.* (1959:91); Meister and Abendschein (1956:171)
2-Keto-isocaproic acid	Elevated	Dancis *et al.* (1959:91); Meister and Abendschein (1956:171)
2-Hydroxy acids analogous		
2-Hydroxy-3-methylvaleric acid	Elevated	Meister and Abendschein (1956:171)
2-Hydroxy-isovaleric acid	Highly elevated in E3 deficiency	Meister and Abendschein (1956:171)
2-Hydroxy-isocaproic acid	Elevated	Meister and Abendschein (1956:171)
Decarboxylated CoAs		
N-2-Methylbutyryl-CoA	Deficient	Dry (1997:1-186)
N-Isobutyryl-CoA	Deficient	Dry (1997:1-186)

Metabolites	Status in MSUD	Additional references
N-Isovaleryl-CoA	Deficient	Dry (1997:1-186)
2-Hydroxy acid conjugates		
N-2-Hydroxy-3-methylvaleryl-Ile	Trace	Dry (1997:1-186)
N-2-Hydroxy-3-methylvaleryl-Val	Trace	Dry (1997:1-186)
N-2-Hydroxy-3-methylvaleryl-Leu	Trace	Dry (1997:1-186)
2-Hydroxy-isovaleryl-Ile	Trace, elevated in variant	Hagenfeldt and Naglo (1987:77)
2-Hydroxy-isovaleryl-Val	Trace, elevated in variant	Hagenfeldt and Naglo (1987:77)
2-Hydroxy-isovaleryl-Leu	Trace, elevated in variant	Hagenfeldt and Naglo (1987:77)
2-Hydroxy-isocaproyl-Ile	N.D.	Dry (1997:1-186)
2-Hydroxy-isocaproyl-Val	N.D.	Dry (1997:1-186)
2-Hydroxy-isocaproyl-Leu	Trace	Dry (1997:1-186)
N-2-Hydroxy-3-methylvaleryl-Gly	Trace	Dry (1997:1-186)
2-Hydroxy-isovaleryl-Gly	N.D.	Hagenfeldt and Naglo (1987:77)
2-Hydroxy-isocaproyl-Gly	N.D.	Hagenfeldt and Naglo (1987:77)
Lactyl conjugates		
N-Lactyl-BCAA	Elevated	Hagenfeldt and Naglo (1987:77)
N-Lactyl-Gly	Trace	Hagenfeldt and Naglo (1987:77)
N-AcAA conjugates		
N-Acetyl-BCAA	Elevated	Dry (1997:1-186)
N-Acetyl-Gly	Trace	Dry (1997:1-186)
N-Acetyl-Ala	N.D.	Dry (1997:1-186)
N-AAA (of decarboxylated 2-keto acids)		
N-2-Methylbutyryl-BCAA	Trace, elevated in variant	Dry (1997:1-186)
N-Isobutyryl-BCAA	Trace, elevated in variant	Dry (1997:1-186)
N-Isovaleryl-BCAA	Trace, elevated in variant	Dry (1997:1-186)
N-2-Methylbutyryl-Gly	N.D.	Dry (1997:1-186)
N-Isobutyryl-Gly	N.D.	Dry (1997:1-186)
N-Isovaleryl-Gly	N.D.	Dry (1997:1-186)

Metabolites	Status in MSUD	Additional references
Hydantoins		
5-Hydroxy-5-isopropyl-hydantoin	Trace, N.D. in variant	Dry (1997:1-186)
5-Hydroxy-5-methylpropyl-hydantoin	Trace, N.D. in variant	Dry (1997:1-186)
5-Hydroxy-5-isobutyl-hydantoin	Trace, N.D. in variant	Dry (1997:1-186)
Other metabolites		
Lactic acid	Elevated in E3 deficiency	Dry (1997:1-186)
Uric acid	Elevated in moderate phenotype	Dry (1997:1-186)
Pyruvate	Elevated in E3 deficiency	Dry (1997:1-186)
2-Ketoglutaric acid	Elevated in E3 deficiency	Dry (1997:1-186)
2-Hydroxyglutaric acid	Elevated in E3 deficiency	Dry (1997:1-186)
3-Hydroxyisovaleric acid	Elevated	Dry (1997:1-186)

†N.D. = Not detected

Table 16: Relevant organic acids in Sample 11.

Compound	Concentration (mg/g creatinine)
2-Keto-3-methylvaleric acid (di-TMS† ester)	N.D. ††
2-Keto-3-methylvaleric acid (mono-TMS ester)	32.71
2-Ketobutyric acid (di-TMS ester)	N.D.
2-Ketocaproic acid (di-TMS ester)	N.D.
2-Ketocaproic acid (mono-TMS ester)	N.D.
2-Ketoisocaproic acid (di-TMS ester)	6.41
2-Ketoisocaproic acid (mono-TMS ester)	4.69
2-Ketoisovaleric acid (di-TMS ester)	2.97
2-Ketoisovaleric acid (mono-TMS ester)	1.74
2-Ketovaleric acid (di-TMS ester)	16.22
2-Ketovaleric acid (mono-TMS ester)	N.D.
4-Hydroxyphenylpyruvate (tri-TMS ester)	4.81
Phenylpyruvate (di-TMS ester)	N.D.

Pyruvate (di-TMS ester)	5.10
Pyruvate (mono-TMS ester)	32.37
5-Hydroxy-5-isopropylhydantoin	2.05
5-Hydroxy-5-methylhydantoin	0.52
N-(Phenylacetyl)-(phenylalanine) (mono-TMS ester)	N.D.
N-2-Methylbutyrylglycine (di-TMS ester)	N.D.
N-2-Methylbutyrylglycine (mono-TMS ester)	N.D.
N-2-Methylbutyrylisoleucine (mono-TMS ester)	102.70
N-2-Methylbutyrylleucine (mono-TMS ester)	193.02
N-2-Methylbutyrylvaline (mono-TMS ester)	14.54
N-Butyryl-norvaline (mono-TMS ester)	N.D.
N-Isobutyrylglycine (di-TMS ester)	N.D.
N-Isobutyrylglycine (mono-TMS ester)	N.D.
N-Isobutyrylisoleucine (mono-TMS ester)	59.95
N-Isobutyrylleucine (mono-TMS ester)	207.38
N-Isobutyrylvaline (mono-TMS ester)	23.60
N-Isovalerylglycine (di-TMS ester)	N.D.
N-Isovalerylglycine (mono-TMS ester)	N.D.
N-Isovalerylisoleucine (mono-TMS ester)	28.32
N-Isovalerylleucine (mono-TMS ester)	404.40
N-Isovalerylvaline (mono-TMS ester)	46.87
N-Pentanoyl-norleucine (mono-TMS ester)	54.07
N-Propionyl-2-aminobutyric acid (mono-TMS ester)	N.D.
N-Acetylalanine (di-TMS ester)	0.30
N-Acetylalanine (mono-TMS ester)	3.55
N-Acetylglycine (mono-TMS ester)	7.73
N-Acetylisoleucine (di-TMS ester)	27.37
N-Acetylisoleucine (mono-TMS ester)	N.D.
N-Acetylleucine (mono-TMS ester)	249.22
N-Acetylvaline (di-TMS ester)	15.15
N-Acetylvaline (mono-TMS ester)	11.23

[†]TMS = trimethylsilyl; [††]N.D. = not detected.

ANNEXURE B: REAGENTS

Table 17: Composition of the 16 algal amino acids mixture from Cambridge Isotope Laboratories (Andover, Massachusetts, USA).

Amino acid name	Approximate percentages (mass of total)
L-Alanine	7
L-Arginine	7
L-Aspartate	10
L-Glutamate	10
Glycine	6
L-Histidine	2
L-Isoleucine	4
L-Leucine	10
L-Lysine	14
L-Methionine	1
L-Phenylalanine	4
L-Proline	7
L-Serine	4
L-Threonine	5
L-Tyrosine	4
L-Valine	5

ANNEXURE C: AMDIS N-AAA CONJUGATE MS LIBRARY

Table 18: Mass spectra available for N-AAA conjugates at the time of this study.

Compound	Spectra source
N-2-Methylbutyrylglycine (di-TMS ester)	In-house
N-2-Methylbutyrylglycine (mono-TMS ester)	In-house
N-2-Methylbutyrylisoleucine (mono-TMS ester)	Literature
N-2-Methylbutyrylleucine (mono-TMS ester)	Literature
N-2-Methylbutyrylvaline (mono-TMS ester)	Literature
N-Acetylglycine (mono-TMS ester)	In-house
N-Acetylalanine (di-TMS ester)	Standard
N-Acetylalanine (mono-TMS ester)	In-house and standard
N-Acetylglycine (mono-TMS ester)	In-house
N-Acetylisoleucine (mono-TMS ester)	In-house and literature
N-Acetylleucine (mono-TMS ester)	Literature
N-Acetylvaline (di-TMS ester)	Standard
N-Acetylvaline (mono-TMS ester)	Standard and literature
N-Butyryl-norvaline (mono-TMS ester)	Derived from sample
N-Isobutyrylglycine (di-TMS ester)	In-house
N-Isobutyrylglycine (mono-TMS ester)	In-house
N-Isobutyrylisoleucine (mono-TMS ester)	Literature
N-Isobutyrylleucine (mono-TMS ester)	Literature
N-Isobutyrylvaline (mono-TMS ester)	In-house and literature
N-Isovalerylglycine (di-TMS ester)	In-house
N-Isovalerylglycine (mono-TMS ester)	In-house
N-Isovalerylisoleucine (mono-TMS ester)	Literature
N-Isovalerylleucine (mono-TMS ester)	In-house and literature
N-Isovalerylvaline (mono-TMS ester)	Literature
N-Pentanoyl-*nor*leucine (mono-TMS ester)	Derived from sample
N-Phenylacetyl-phenylalanine (mono-TMS ester)	Derived from sample
N-Propionyl-2-aminobutyric acid (mono-TMS ester)	Derived from sample

In the table above: 'Standard' refers to mass-spectra that were derived from a pure standard derivatised with BSTFA and TMCS, 'In-house' spectra were taken from the in-house organic acid compound library (NWU Centre for Human Metabolomics, 2017), 'Literature' spectra were manually added to the study library by printing and measuring peak heights and calculating relative abundances, 'Derived from sample' spectra were taken from samples where the compounds of interest were expected (products of synthesis reactions) and analysed with MS Interpreter version 2.0 – a part of the NIST mass spectral database version 1.6/1.7/2.0.

ANNEXURE D: EXAMPLE INPUT FILE FOR PES SCAN

```
#!perl

use strict;
use Getopt::Long;
use MaterialsScript qw(:all);



################################################################################################
######
# User editable settings

my $filename = "#E-R"; # without the xsd extention!

my $distanceConstraintName = "O-H"; # select the constraint in the xsd file, and check the name in the
properties window
my $initialDistance = 0.978;
my $finalDistance = 4.093;
my $distanceStepSize = 0.05; # positive value increases distance, negative value decrease distance

# Level of Theory

my $dmol3 = Modules->DMol3;
$dmol3->ChangeSettings([
	# Functional & DFTD
	TheoryLevel => "GGA",
	NonLocalFunctional => "PW91",
	UseDFTD => "Yes",
	DFTDMethod => "OBS",

	# Basis set
	BasisFile => "4.4",
	SpinUnrestricted => "Yes",
	AtomCutoff => 3.7, #  3.7 corresponds to default for medium quality for organic chemistry, with
regards to phosphorus.
	UseSmearing => "Yes",
	Smearing => 0.005,

	# Other
	UseSymmetry => "No",
	Quality => "Fine",
	MaximumSCFCycles => 1000,
	Charge => 1,
	UseCosmo => "Yes",
	SolventDielectric => 78.54,
#	COSMO_Dielectric=>78.5400->Water,

	# Computational
	MaxMemory => 12000, #+- 500 MB per core used.
]);


# End user editable settings
################################################################################################
######
```

```
print "The script has started.\n\n";

my $doc = $Documents{"$filename.xsd"};
my $distanceConstraintMonitor = $doc->Distances("$distanceConstraintName"); # assigns the distance constraint to a variable


# Creates and prepares a study table into which the output can be stored
my $studyTable = Documents->New("$filename"."PES.std")->ActiveSheet;

$studyTable -> ColumnHeading(0) = "Structure";
$studyTable -> ColumnHeading(1) = "Distance";
$studyTable -> ColumnHeading(2) = "Energy";

my $studyTableRowCounter = 0; # variable that is used to populate the study table


# Begins the for loop that does the subsequent geometry optimizations
for (my $currentDistance = $initialDistance; $currentDistance <= $finalDistance; $currentDistance+=$distanceStepSize) {

    $distanceConstraintMonitor->Distance = $currentDistance; # changes the length of the constraint to the current value in the loop

    print "Processing frame $studyTableRowCounter. Current distance is $currentDistance.\n";

    my $output;

    eval {$output = $dmol3 -> GeometryOptimization -> Run($doc);};

    if ($@) {
        print "ERROR. Frame $studyTableRowCounter did not converge.";
        print $@;
    } else {
        my $energy = $output -> TotalEnergy;
        $studyTable -> Cell($studyTableRowCounter, 0) = $doc;
        $studyTable -> Cell($studyTableRowCounter, 1) = $currentDistance;
        $studyTable -> Cell($studyTableRowCounter, 2) = $energy;
    } # end if else block

    ++$studyTableRowCounter;

} #end for

print "\n\nThe script has completed.";
```

ANNEXURE E: CRYSTALLOGRAPHY DATA SUPPORTING MOLECULAR MODELLING

The tables below (Table 19 and Table 20) show crystallography data obtained from Harata *et al.* (1976:211) ('Literature') compared with values obtained from molecular modelling ('Calculated values'). The bond notation for bond distances and angles correlates to the labels given in Figure 61.

Table 19: Comparison of bond distances between those obtained from molecular modelling and literature for pyruvic acid.

Bond distances (Å)	Literature	Calculated values
C(1)-C(2)	1.49	1.49
C(2)-O(1)	1.21	1.22
C(3)-O(3)	1.22	1.22
C(1)-H(C1)	0.93	1.09
C(1)-H'(Cl)	1.20	1.10
C(1)-H''(Cl)	1.12	1.10
C(2)-C(3)	1.53	1.54
C(3)-O(2)	1.31	1.34
O(2)-H(O2)	1.19	0.98

Table 20: Comparison of bond angles between those obtained from molecular modelling and literature for pyruvic acid.

Bond angles (°)	Literature	Calculated values
C(1)-C(2)-C(3)	116.0	115.1
C(3)-C(2)-O(1)	120.2	119.9
C(2)-C(3)-O(3)	121.2	122.9
C(1)-C(2)-O(1)	123.8	125.1
C(2)-C(3)-O(2)	114.5	112.1
O(2)-C(3)-O(3)	124.3	125.0

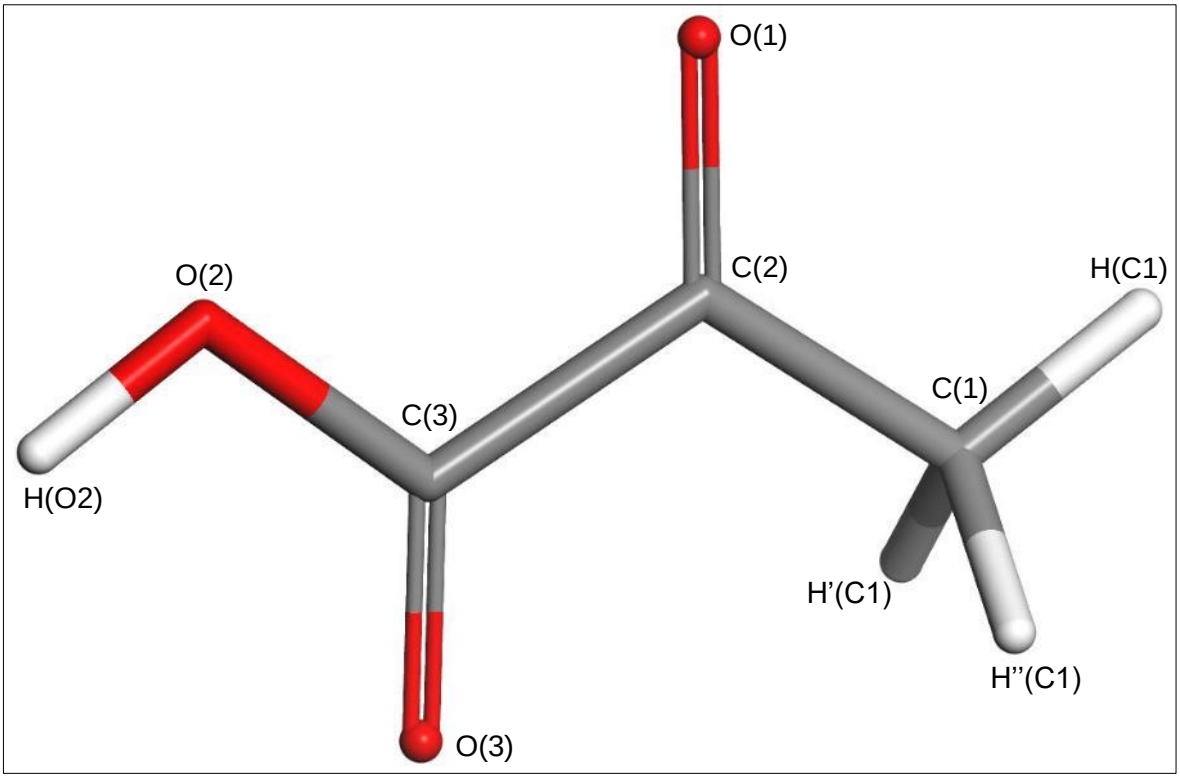

Figure 61: Pyruvic acid model with labels.

The labels used in the figure above [e.g. O(2) and C(3)] correlate to those used in Table 19 and Table 20. The labelled atoms' numbers do not coincide with IUPAC nomenclature but rather to the numbering scheme of Harata et al. (1976:211).