Author/s:
Nathanael, Joses Grady

Title:
The oxidative damage of biological molecules by air pollutants NO2● and NO3●

Date:
2019

Persistent Link:
https://hdl.handle.net/11343/225057

The oxidative damage of biological molecules by air pollutants $NO_2^{\bullet}$ and $NO_3^{\bullet}$

Joses Grady Nathanael

Submitted in total fulfilment of the requirements of the degree
Doctor of Philosophy

February 2019

School of Chemistry
The University of Melbourne

ORCID: 0000-0002-5373-2833

Abstract

Air pollution is perceived as the world's greatest environmental risk to human health. According to the World Health Organization (WHO), air pollution is responsible for the deaths of about 7 million people each year. In the industrialised urban environment, nitrogen dioxide ($NO_2^{\bullet}$) and ground-level ozone (O_3) are the most oxidising air pollutants. Exposure to these gases has been associated with increased respiratory health problems, such as exacerbation of existing asthma and allergies. While the adverse health effects of air pollution are clear, the precise underlying mechanism through which the pollutants affect biological systems is not well understood. It has been speculated that nitrate radicals ($NO_3^{\bullet}$), which are formed from the reaction of $NO_2^{\bullet}$ and O_3, play an important role in the oxidative damage of biological systems. Therefore, this thesis explores reactions involving $NO_3^{\bullet}$ and biomolecules such as proteins through a combination of kinetic, computational and product studies, in order to gain a better understanding of the fundamental chemical pathways that lead to oxidative damage in biological systems upon exposure to air pollution.

The first section of this thesis investigates the reaction of $NO_3^{\bullet}$ with aliphatic amino acids and peptides. From laser flash photolysis experiments, it was found that $NO_3^{\bullet}$ reacts with aliphatic amino acids and peptides at multiple sites through proton-coupled electron transfer (PCET) at the amide nitrogen, and hydrogen atom transfer (HAT) at the α-carbon or the activated C–H moiety (e.g., tertiary carbons) with the rate of about 1×10^6 M^{-1} s^{-1}.

Following the above finding, this thesis proceeds to examine the reaction of $NO_3^{\bullet}$ with aromatic amino acids and peptides. A faster rate by a factor of 5–6 suggests that the reaction occurs at the aromatic ring through electron transfer

(ET). An unprecedented amide neighbouring group effect was discovered, by which the rate of aromatic ring oxidation is increased considerably when the ring is flanked by two amide groups, instead of one amide and one ester group. Due to this effect, phenylalanine can potentially act as a relay amino acid in a long-distance ET even though the aromatic ring in phenylalanine is not readily oxidisable under biochemical conditions.

The third section of this thesis explores $NO_3^{\bullet}$ reactions involving proline, where its side chain is covalently bound to the α-amino group. This unique structure increases electron density at the nitrogen and significantly accelerates the rate of ET at this nitrogen by a factor of about 600 compared to the other aliphatic substrates. However, when the amide moiety in proline residue is involved in the amide neighbouring group effect, accelerating the rate of aromatic ring oxidation, the rate of ET at this nitrogen was found to decrease significantly.

The final part of this thesis studies the reaction of $NO_2^{\bullet}$ with various biological molecules, including short peptide and cholesterol derivatives. It was found that contrary to the widely accepted radical pathway, the reaction of $NO_2^{\bullet}$ with these molecules involves an ionic pathway through the dissociation of N_2O_4 into NO^+ and NO_3^-.

Declaration

This is to certify that

I) the thesis comprises only my original work towards the PhD except where indicated in the Preface;

II) due acknowledgement has been made in the text to all other material used; and

III) the thesis is less than 100 000 words in length, exclusive of tables, references and appendices

Joses G. Nathanael

February 2019

Preface

All work reported herein has been conducted by the candidate Joses G. Nathanael except where indicated. In accordance with regulations of The University of Melbourne, I acknowledge that portions of this work were performed in a collaborative manner (at the Bio21 Institute, The University of Melbourne, unless otherwise state):

Chapter 2 Synthesis of Phth-Gly and Ac-Tyr-OMe was carried out by Dr. Luke Gamon.

Chapter 3 Ac-Leu-NH_2 was prepared by Julius Kerstien. Computational studies were performed by Prof. Uta Wille.

Chapter 4 Product studies of Ac-Phe-Phe-OMe were performed by Dr. Luke Gamon. Computational studies were performed by Prof. Paul R. Rablen (Swarthmore College, PA, USA).

Chapter 5 Computational studies were performed by Prof. Uta Wille. X-ray analysis from product studies of Ac-Pro-Phe-OMe was performed by Prof. Jonathan White.

Chapter 6 Product studies of Ac-Gly-Gly-Gly-Phe-OMe were performed by Dr. Luke Gamon. Product studies of *O*-methyl cholesterol and *O*-acetyl cholesterol were performed by Anton Zalewski.

The remainder of this thesis comprises only my orginal work. Overall, I assess my contribution to be greater than 95%.

Publications arising from this work

Chapter 3

1. "Reaction of Amino Acids, Di- and Tripeptides with the Environmental Oxidant $NO_3^{\bullet}$: A Laser Flash Photolysis and Computational Study"

   Nathanael, J. G.; Hancock, A. N.; Wille, U. *Chem. Asian J.* **2016**, *11*, 3188–3195.

2. "Oxidative Damage in Aliphatic Amino Acids, Di- and Tripeptides by the Environmental Free Radical Oxidant $NO_3^{\bullet}$: The Role of the Amide Bond Revealed by Kinetic and Computational Studies"

   Nathanael, J. G.; Wille, U. *J. Org. Chem.* **2019**, *84,* 3405–3418.

Chapter 4

3. "Amide Neighbouring-Group Effects in Peptides: Phenylalanine as Relay Amino Acid in Long-Distance Electron Transfer"

   Nathanael, J. G.; Gamon, L. F.; Cordes, M.; Rablen, P. R.; Bally, T.; Fromm, K. M.; Giese, B.; Wille, U. *Chem. Bio. Chem.* **2018**, *19*, 922–926.

Chapter 6

4. "Fragmentation-Rearrangement of Peptide Backbones Mediated by the Air Pollutant $NO_2^{\bullet}$"

   Gamon, L. F.; **Nathanael, J. G.**; Taggert, B. I.; Henry, F. A.; Bogena, J.; Wille, U. *Chem. Eur. J.* **2015,** *21,* 14924–14930.

5. "Oxidation of cholesterol and O-protected derivatives by the environmental pollutant $NO_2^{\bullet}$"

   Zalewski, A. N.; **Nathanael, J. G.**; White, J. M.; Wille, U. *Chem. Commun.* **2016,** *52,* 4060–4063.

Acknowledgments

Prof. Uta Wille

Thank you so much for all your guidance and support in the last four years. I am amazed at how much I have grown under your wings. If research life is a warzone, you are the kind of commander who cares for our growth as an individual and fights along our side in every battle we face. Thank you for equipping us with 'guns' and 'fangs' so that when we go on our own way into the real world, we know how to stand for ourselves. I thank you from the bottom of my heart.

The Wille Group & Bio21 members

Thank you to all members (past and present): Luke, Emma, Beth, Kay, Desi, Bo, Parv, Sibel, Ben, Tom, Jessie, Anton, Cam and all who have visited during my time. Thank you for being a mentor, a colleague and a friend to me. For all the fun times we had in the lab, I will cherish the memories in my heart. Thank you especially to Howard, Kay and Ethane for the companies and encouragement you have provided. Playing board games and having hot pots have been an important part of my PhD journey. Thank you also to Myint, Finas and Tram who have been there to help and support me.

Friends and family

Thank you for all the love and support you have given me. To my parents who have taken care of me since little with patience and understanding and let me pursue my dreams. To my brother, Jason, the best and worst brother I will ever have, thank you for your support and your cheekiness that always put smiles on my face and irritate my heart at the same time. A special thank you to Joshua, Jeremy, Mikha, David A, Nat B, Hajin, Johan, Noel, and David S who have been there for moral supports. Thank you to Patricia, who not only has been my senior demonstrator, but also my mentor and someone I look up to.

Along the way, I have met a lot of people who have changed me as a person in so many ways and I am grateful for each one of you.

List of symbols and abbreviations

α	alpha
Å	Ångström
aa	amino acid
Ac	acetyl
Aib	2-aminoisobutyric acid
Ala	alanine
aq.	aqueous
atm	atmosphere (pressure)
β	beta
bis(TRIS)	bis(2-hydroxyethyl) amino-tris(hydroxy methyl)methane
Boc	*tert*-butoxycarbonyl
Bn	benzyl
BSA	bovine serum albumin
Bu	butyl
Bz	benzoyl
c	centi-
°C	degree(s) Celcius
calcd.	calculated
CAN	cerium(IV) ammonium nitrate
CT	charge-transfer
CumO	cumyloxyl
δ	delta, chemical shift
Δ	delta, heat
$\Delta E^{\ddagger}$	activation barrier
ΔE	reaction energy
d	doublet (NMR)
d	deuterium
DCC	*N*,*N*'-dicyclohexyl carbodiimide
DCM	dichloromethane
dd	doublet of doublets (NMR)
ddd	doublet of doublet of doublets (NMR)
DIPEA	*N*,*N*-diisopropylethyl amine
DMAP	4-dimethylamino pyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dt	doublet of triplets (NMR)
dp	doublet of pentets (NMR)
dq	doublet of quartets (NMR)
dqd	doublet of quartet of doublets (NMR)
ε	epsilon, extinction coefficient
e	electron (charge)
EDC	*N*-(3-dimethyl aminopropyl)-*N*'-ethylcarbodiimide
EDG	electron-donating group
EPR	electron paramagnetic resonance
eq	equation

equiv.	equivalent(s)
E_{rel}	relative energy
ESI	electrospray ionisation
Et	ethyl
ET	electron transfer
EWG	electron-withdrawing group
γ	gamma
g	gram(s)
GABA	γ-aminobutyric acid
GC	gas chromatography
Glp	pyroglutamic acid
Glu	glutamic acid
Gly	glycine
h	hour(s), hexet (NMR)
HAT	hydrogen atom transfer
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
hept	heptet (NMR)
heptd	heptet of doublets (NMR)
His	histidine
HOBt	1-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
Hz	hertz (NMR)
i	iso-
Ile	isoleucine
IR	infrared
J	joule(s)
J	coupling constant (NMR)
k	kilo-
k	absolute second-order rate coefficient
K	Kelvin(s)
K_a	acid dissociation constant
KIE	kinetic isotope effects
k_{obs}	pseudo-first-order rate coefficient
λ	wavelength
L	litre(s)
Leu	leucine
Lys	lysine
μ	micro-
m	milli-, metre(s), multiplet (NMR)
m	meta
M	molecular ion (HRMS), mega- (NMR), molar, metal
m/*z*	mass-to-charge ratio
Me	methyl
min	minute(s)
MNP	2-methyl-2-nitrosopropane
mol	mole(s)
n	nano-
NADP	nicotinamide adenine dinucleotide phosphate
NBS	*N*-bromosuccinimide

NHE	normal hydrogen electrode
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
o	ortho
OD	optical density
π	pi
p	pentet (NMR)
p	para
P5C	pyrroline-5-carboxylate
PCET	proton-coupled electron transfer
Pet.	petroleum
Ph	phenyl
Phe	phenylalanine
Phth	phthaloyl
Piv	pivaloyl
PM	particulate matter
PMA	phosphomolybdic acid
PMT	photomultiplier tube
ppb	parts-per-billion
ppm	parts-per-million
ppt	parts-per-trillion
Pr	propyl
Pro	proline
q	quartet (NMR)
R	alkyl group
ROS	reactive oxygen species
RSE	radical stabilisation energy
RT	room temperature
RTLF	respiratory tract lining fluid
σ	sigma
s	second(s), singlet (NMR)
sat.	saturated
sec	secondary
Ser	serine
t	triplet (NMR)
t	tertiary
td	triplet of doublets (NMR)
tert	tertiary
TFA	trifluoroacetic acid
Tfa	trifluoroacetyl
THF	tetrahydrofuran
TLC	thin layer chromatography
TOF	time-of-flight
Trp	tryptophan
TS	transition state
Tyr	tyrosine
UV	ultraviolet
V	volt(s)
Val	valine
VOC	volatile organic compound
wt.	weight

Table of contents

List of figures

List of schemes

List of tables

Chapter One: Introduction

1.1 Pollution in the atmosphere

Air pollution has been identified as a major environmental health issue that is associated with increasing mortality. A number of unfortunate events since the early 20th century have been reported in the literature, which demonstrate lethality of air pollution.[1] For example, in 1930 the smog that resulted from the burning of coal and affected the Meuse Valley in Belgium was recognised as the cause of over 6000 respiratory problems and 60 deaths after only a few hours of illness.[1a,f] Sulfur dioxide (SO_2) and sulfuric acid (H_2SO_4), which result from oxidation of SO_2 in the air, were identified as the main culprit of this catastrophe.[1a] Another major incident took place in 1952 in London, UK following a four-day smog, where the death toll was estimated to be over 4000 during the ensuing two-week period.[1b,c] The incident resulted from a combination of very cold weather and a large increase in burning of cheap coal, leading to formation of sulfate aerosols and acid rain (due to H_2SO_4). A considerable increase in numbers of deaths was observed even on the first day of the incident. Deaths assigned to bronchitis and pneumonia were reported to increase, eight times and three times respectively, in just one week.[1b] After this episode, the death rate remained high for a few more months.[1h] Data incorporating this extended period, revealed that the number of additional deaths due to acute and persistent effects of the 1952 London smog reached 12000.[1g] Clearly, there is a strong correlation between exposure to pollution (smog) and the mortality rate.

In response, various efforts and legislation have sought to eliminate most of the old methods which produced air pollutant.[1h] Despite this successful abatement of air pollution derived from combustion of the traditional (sulfur-containing)

fossil fuel, there is a steady growth in emission of traffic-associated pollutants.[2] In fact, air pollution is still perceived as the world's greatest environmental risk to human health. The World Health Organization (WHO) estimates that in 2012 the deaths of about 7 million people were attributed to the effects of household (indoor) and ambient (outdoor) air pollution.[3] This is equivalent to one in eight of total reported global deaths in that year.[3b] Therefore, understanding how air pollution causes health issues is vital to combatting its adverse effects and potentially saving millions of lives.

Under the National Environmental Protection Measure (NEPM) in Australia, the National Environmental Protection Council (NEPC) has set ambient air quality standards and identified six major air pollutants (Figure 1.1-1).[4a]

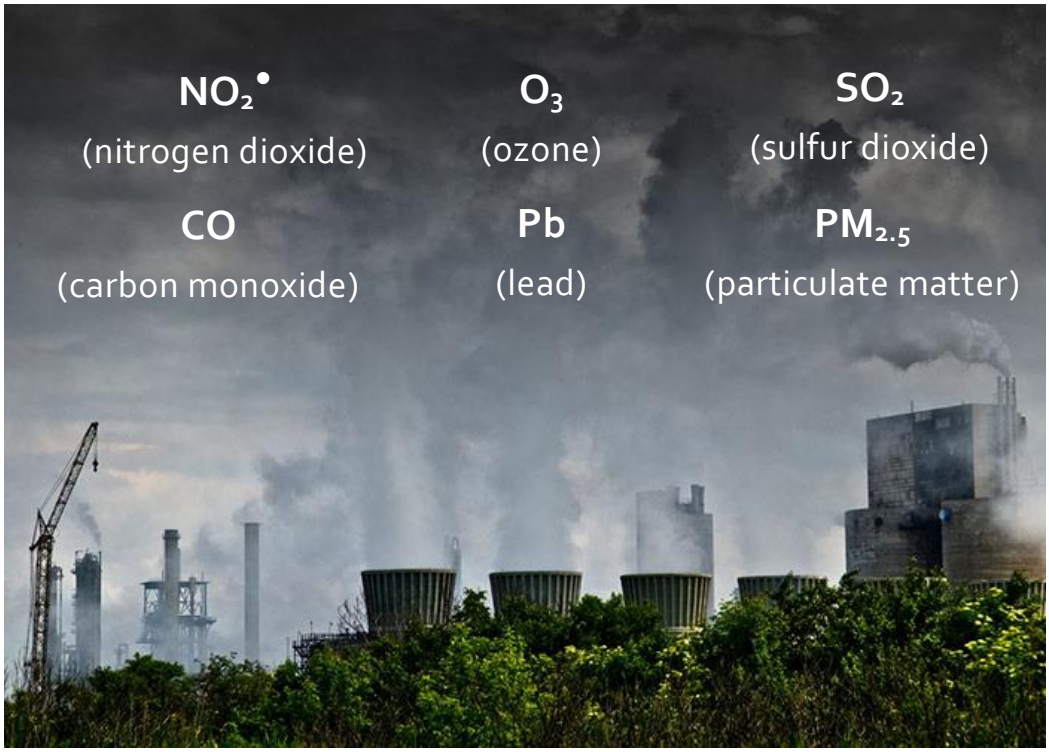


Figure 1.1-1. The six common pollutants according to the NEPC,[4a] figure taken from India Celebrating.[4b]

These major air pollutants range from toxic gases ($NO_2^{\bullet}$, O_3, SO_2 and CO) to tiny solid materials, such as Pb and $PM_{2.5}$ (particles with a diameter of less than 2.5 μm), suspended in the air. They are generally toxic and act on humans in various

ways. For example, CO replaces O_2 in haemoglobin, which consequently deprives the body parts of getting enough O_2, and in high concentrations, CO can be lethal to humans.

Among these major air pollutant components, nitrogen dioxide ($NO_2^{\bullet}$) and ground-level ozone (O_3) are of particular concern, especially in the industrialised urban environment. Not only can they cause oxidative damage due to their oxidising nature [E^o ($NO_2^{\bullet}/NO_2^-$) = 0.9 V; E^o (O_3/ $O_3^{\bullet -}$) = 1.6 V],[5] but studies have also shown that increases in respiratory health problems, such as asthma and allergies, are strongly correlated to increased $NO_2^{\bullet}$ and O_3 concentrations in the atmosphere.[6] Primary sources of $NO_2^{\bullet}$ include the combustion of fossil fuels in motor vehicles, electricity generation, mining industries, burning of forests or vegetation and manufacturing (Figure 1.1-2, left chart).[7]

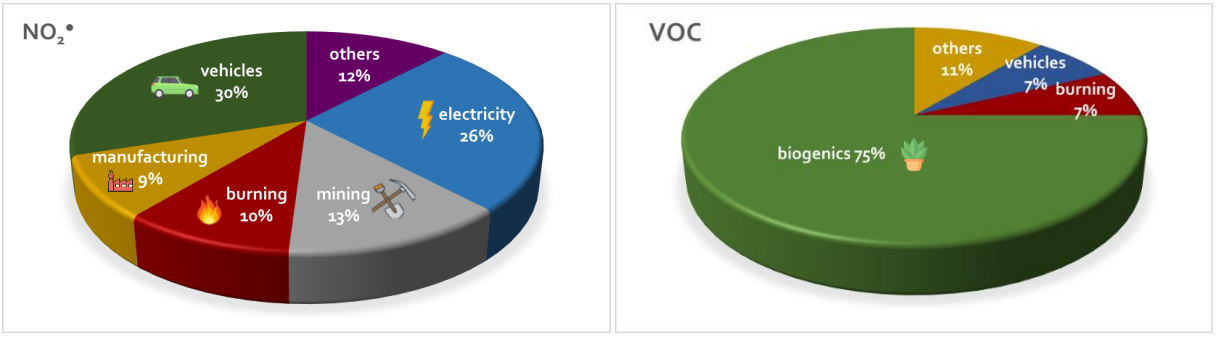


Figure 1.1-2. Source of tropospheric $NO_2^{\bullet}$ (left chart) and VOC (right chart) as the O_3 precursor in Australia in 2016/2017; data taken from the National Pollutant Inventory.[7b]

Ground-level O_3 exists in the troposphere as a result of a complex series of chemical reactions involving $NO_2^{\bullet}$ and volatile organic compounds (VOCs) in the presence of sunlight.[1h] The majority of the VOCs in Australia (up to 75%) are emitted from vegetation or other living organisms (Figure 1.1-2, right chart).

In typical populated industrialised major cities, such as Los Angeles, London and Beijing, $NO_2^{\bullet}$ and O_3 always coexist at higher concentrations than the global average (Figure 1.1-3).

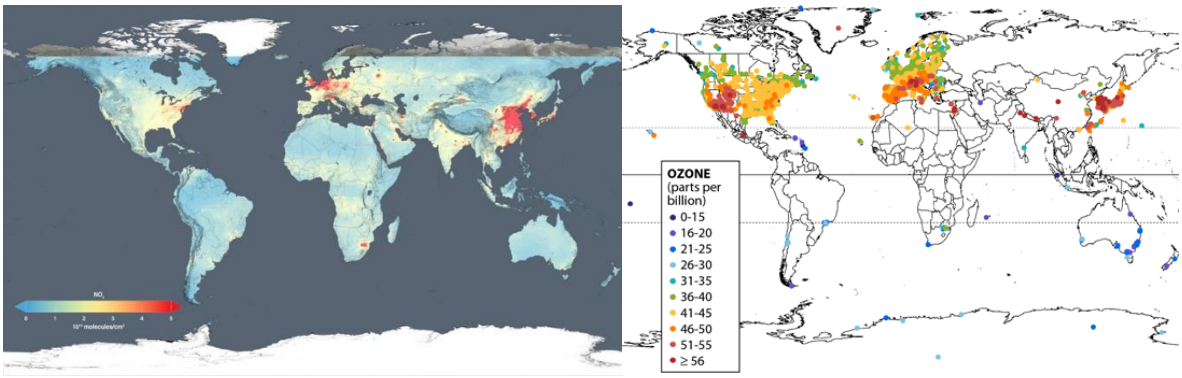


Figure 1.1-3. Average global tropospheric $NO_2^{\bullet}$ concentrations in 2014 (left image, red colour denotes the region where concentration is higher than average), source: Aura satellite via NASA Goddard Space Flight Center.[8a] Average global tropospheric O_3 concentrations in 2010–2014 (right image, red colour denotes the region where concentration is higher than average), source: National Oceanic and Atmospheric Administration (NOAA).[8b]

Significant synergistic effects upon exposure to $NO_2^{\bullet}$ and O_3 in combination have been reported in the literature.[9,10] For example, a clear increase of lipid peroxidation in the lungs of guinea pigs and a considerable growth of antioxidative enzyme activities in rats were observed when these animals were exposed to the mixtures of $NO_2^{\bullet}$ and O_3 as opposed to the single gases alone.[9c] Additionally, protein nitration was observed in birch pollen proteins upon exposure to $NO_2^{\bullet}/O_3$ mixtures.[10] Even though $NO_2^{\bullet}$ alone has been shown to effectively nitrate birch pollen proteins, the quantity of the nitrated proteins was found to be one order of magnitude higher when both $NO_2^{\bullet}$ and O_3 were present.[10b] This nitration can induce a stronger allergenic potential of the pollens, which may promote exacerbation of asthma and other respiratory symptoms.[10c]

These important findings could be taken as an indication for the formation of a highly reactive intermediate under these conditions, which could play a significant role in aggravating oxidative damage in biological systems. It was speculated that this reactive species could be the nitrate radical ($NO_3^{\bullet}$), which is formed from the reaction of $NO_2^{\bullet}$ and O_3 (eq. 1.1).[11]

$$NO_2^{\bullet} + O_3 \rightarrow NO_3^{\bullet} + O_2 \qquad \text{(eq. 1.1)}$$

During the daytime, $NO_3^\bullet$ is less important as an oxidant because of its low concentration and rapid photolysis (eq. 1.2).[11,12] The lifetime of $NO_3^\bullet$ in direct sunlight is not more than five seconds.[12]

$$NO_3^\bullet + h\nu \rightarrow NO^\bullet + O_2 \qquad \text{(eq. 1.2a)}$$

$$NO_3^\bullet + h\nu \rightarrow NO_2^\bullet + O(^3P) \qquad \text{(eq. 1.2b)}$$

However, at nighttime $NO_3^\bullet$ becomes the most dominant oxidant in the atmosphere.[11,13] As an example, Figure 1.1-4 shows that $NO_3^\bullet$ tropospheric concentrations were found to rise after sunset until early morning based on a long-term measurement in a semi-arid urban location of Jerusalem, Israel.[13c]

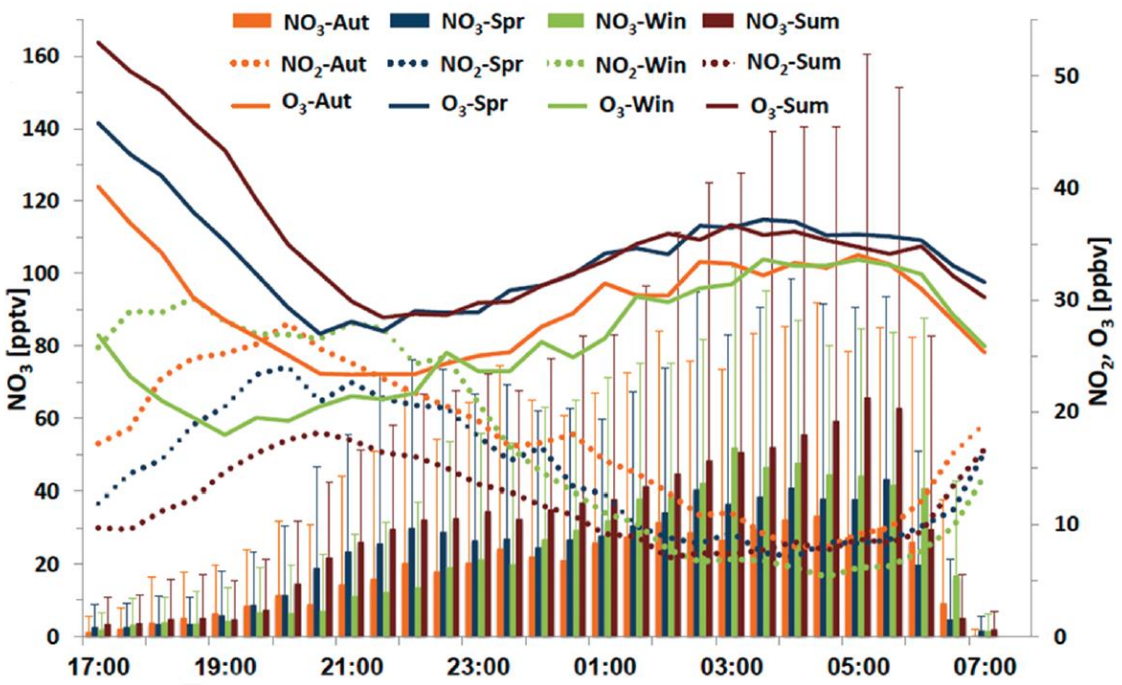


Figure 1.1-4. Seasonal half-hourly average concentrations of $NO_3^\bullet$, $NO_2^\bullet$ and O_3 in Jerusalem from July 2005 to September 2007, reproduced from Asaf et al.[13c]

While the $NO_3^\bullet$ concentration is still lower than the $NO_2^\bullet$ and O_3 concentrations, which are at ppb levels, $NO_3^\bullet$ is a far stronger oxidising agent [E^o ($NO_3^\bullet$/ NO_3^-) = 2.3–2.5 V vs normal hydrogen electrode (NHE)][14] than $NO_2^\bullet$ and O_3, and hence, its lower abundancy is outweighed by its high reactivity. In a polluted environment,

where $NO_2^{\bullet}$ and O_3 concentrations are quite high (above 100 ppb), $NO_3^{\bullet}$ can even persist during the day and significantly contribute to the oxidation of some reactive VOCs, such as α-pinene and isoprene.[15] It was shown that in the European troposphere, the contribution of $NO_3^{\bullet}$ to the daily VOC degradation reached 28%. This significantly exceeds the contribution by O_3 (only about 17%) even though the tropospheric concentration of O_3 is three orders of magnitude higher than that of $NO_3^{\bullet}$ (O_3 is at ppb level, while $NO_3^{\bullet}$ is at ppt level).[13a] Therefore, it is reasonable to suggest that $NO_3^{\bullet}$ plays an important role in the oxidative damage of biological systems that are exposed to the atmosphere and contributes to various respiratory health problems.

1.2 Oxidative damage and the human defence system

Inhalation is the major route of exposure to gaseous air pollutants, such as $NO_2^{\bullet}$ and O_3.[16] This exposure initiates a number of responses in the human body that might be particularly harmful to vulnerable members of the population. One of the responses is the development of inflammation in the respiratory tract.[17] When air pollutants are inhaled, the respiratory tract lining fluids (RTLF) are the first biological fluids that come into contact with the environmental oxidants. RTLF contain a range of substances, including low-molecular-mass antioxidants, metal-binding proteins, antioxidative enzymes and unsaturated lipids.[18] The antioxidants, most importantly uric acid, ascorbic acid (vitamin C) and glutathione, are the primary targets of the pollutants and act as the first line of defence to protect the underlying epithelial cells against any possible oxidative damage (Figure 1.2-1).[18b–d]

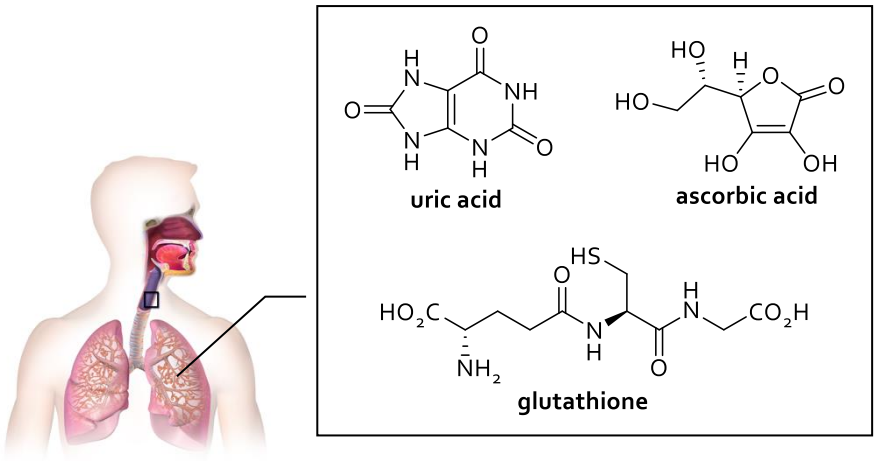


Figure 1.2-1. Antioxidants in RTLF act as the first line of defence against air pollution.

Antioxidants are continuously regenerated as a part of the body defence mechanism. However, following continuous exposure or short exposure to a relatively high-pollutant-content atmosphere, the non-antioxidant substrates can also become a target of attack, especially when antioxidants are already depleted.[17] When the oxidising pollutants react with these non-antioxidant substrates, including proteins and lipids on the cell surfaces or in the RTLF, highly reactive secondary oxidation products, such as free radicals, ozonides, peroxides and aldehydes, may be formed. This series of events can then damage the underlying epithelial cells, subsequently causing inflammation and cell injury.[17]

Several epidemiologic and exposure research studies have been carried out to investigate the oxidative damage of air pollution on human health.[9,10,19] Epidemiologic studies have been able to demonstrate strong correlations between certain population groups of people and hospital admissions or mortality rates, but cannot prove causation.[19d] For example, the concentration of RTLF protective antioxidants, such as ascorbic acid, was found to be lower in asthmatic patients than that in healthy individuals.[19b] The amount of glutathione, one of the antioxidants in the RTLF, in an individual decreases with age as the body ability to synthesise glutathione diminishes.[19c] Given the lowered antioxidant content in

asthmatic patients and the elderly,[19c] it is reasonable to suggest that the weakened antioxidant defence system is the reason for an increased sensitivity to air pollution and oxidative damage, which consequently increases the vulnerability to lung infections in these groups of people.

Meanwhile, exposure studies enabled identification of possible biomarkers for oxidative damage in the human body,[9c,10] but a detailed understanding of the individual mechanistic steps that lead from the initial oxidative damage to acute adverse health effects has not been fully elucidated. One of the main reasons for the lack of a clear reaction pathway is the complexity of human biological system. Typical human cells consist of various components, ranging from macromolecules such as carbohydrates, proteins, lipids and DNA to biomolecular complexes such as mitochondria, the golgi apparatus and ribosomes.[20] Each component possesses unique physical and chemical properties, distinct molecular compositions and different cellular functions. Therefore, when exposed to foreign molecules and particles from outside normal cell environments (in this case, air pollutants), these components exhibit diverse behaviours and reactivities, which then trigger a range of responses and may cause adverse health problems. Moreover, the presence of water in biological fluids and cells is thought to play an important role in biological processes due to the possible interactions with biological molecules.[21] This would add more complexity to the existing complex network of biological systems.

Considering that air pollutants in different places of the world have diverse chemical composition with different reaction properties, emission, stability and diffusion rates,[16] the interactions of air pollution with biological systems can trigger a large variety of complex reaction pathways that could result in health complications. In order to gain a better understanding of the health effects of air pollution, a number of studies have used different approaches to simplify and

model biological systems.[22] One study exposed laboratory rats to the real polluted environment for six months and studied the respiratory mechanics, mucociliary transport and morphological changes of their respiratory system in comparison to the control animals.[22a] It was found that the rats exposed to pollutants developed respiratory distress, showed an increased number of mucus-secreting cells in the airways and exhibited an unexpected growth of epithelial cells (epithelial hyperplasia). Another study focused on the effect of air pollution on DNA and discovered an elevated heritable DNA mutation rate in laboratory mice after exposure to pollutants from steel mills.[22b] The work by Devlin and co-workers modelled biological systems by using human bronchial epithelial cells and subjecting them to metal-containing air pollution particles.[22c] The results demonstrated that exposure to these pollutants stimulates rapid production of inflammatory mediators in the cultured human airway epithelial cells.

However, at best these studies serve as the answer to only some of the missing puzzle pieces of 'health effects of air pollution'. There is no one single experiment that can provide sufficient explanations or solutions to this major environmental issue humankind faces. Therefore, the work done in this thesis envisages to obtain another missing puzzle piece by using a different approach. The underlying mechanism of how air pollution causes adverse health effects will be explored at molecular levels by exposing one particular biological molecule to one particular air pollutant to really simplify the complex biological systems and eliminate any possible interference from the body defence system or interaction with biological fluids and water.

1.3 Scope of this thesis

This thesis will endeavour to address the following important question: "**What are the chemical pathways for the oxidative damage of biological molecules upon exposure to the mixtures of air pollutants $NO_2^{\bullet}$ and O_3?**" Without knowledge of this mechanism, effective strategies to overcome the environmental health risk of air pollution cannot be established. A considerable number of model compounds have been selected to represent the biological systems and aid in the elucidation of the reaction mechanism. Chapter 2 is dedicated to the design and synthetic methodology used to access the model compounds. The interactions of these model compounds with $NO_3^{\bullet}$, species that results from $NO_2^{\bullet}/O_3$ mixtures in the atmosphere, will be investigated and serve as the central theme of the thesis.

This work intends to use a simplified model system by stripping the biological system down to the basic reaction without any interference from its environment (such as biological fluids, salts or other molecules). It is a 'bottom-up' approach that enables the elucidation of fundamental chemical processes occurring between biological molecules and air pollutants. Once these processes are well understood in terms of reaction mechanisms and products, the system can then be extended to include more relevant biological conditions. Therefore, the work in this thesis does not intend to mimic the exact biological system. Below is a short summary of each chapter.

1.3.1 Chapter 3 outline: Reaction of nitrate radicals with aliphatic amino acids and peptides

This chapter studies the reactivity of aliphatic amino acids and peptides towards $NO_3^\bullet$ through a combination of kinetic and computational studies. Laser flash photolysis is employed to perform kinetic studies and rate coefficients for the reaction of aliphatic amino acid and peptide derivatives are obtained. Based on the rate data and through a process of isolation and elimination, the reactive sites of aliphatic substrates are identified. Electronic and steric effects on the reactivity are explored by varying the N-terminal protecting groups of the substrate. The contribution towards the reaction rate based on the obtained rate coefficients is then determined for each reactive site.

1.3.2 Chapter 4 outline: Reaction of nitrate radicals with aromatic amino acids and peptides

This chapter explores the reaction of $NO_3^\bullet$ with aromatic amino acids and peptides, especially phenylalanine. Under oxidative environments, aromatic ring is known to undergo an initial electron transfer (ET). This exisiting mechanism is further investigated mainly through kinetic studies, including kinetic isotope effect (KIE) studies. Phenylalanine is also incorporated in a peptide chain varying its position along the chain. Evaluation of the rate data reveals the influence of peptide environments on the susceptibility of the aromatic ring towards oxidative damage. Due to the high reactivity of tyrosine and tryptophan, they cannot be employed in the standard laser flash experiments. Thus, the side chains of both tyrosine and tryptophan are acetylated to 'tame' their high reactivity. Rate coefficients involving tyrosine and tryptophan derivatives are then determined.

1.3.3 Chapter 5 outline: Reaction of nitrate radicals with proline and proline-containing short peptides

This chapter investigates the reactivity of proline towards $NO_3^{\bullet}$. The distinct structure of proline provides an alternative reaction pathway due to the presence of a considerably more electron-rich amide. Rate data are evaluated for proline and proline-containing dipeptides. Product studies are carried out to provide a further mechanistic insight for reactions involving proline-containing dipeptides. Some model compounds that mimic proline structure are then employed in the laser flash experiments to investigate the reaction pathway, especially ET at the tertiary amide.

1.3.4 Chapter 6 outline: Reaction of nitrogen dioxide with biological molecules

This chapter examines the products resulting from the exposure of biological molecules, such as short peptides and cholesterol, to $NO_2^{\bullet}$. The work in this chapter is a part of collaborative works in the mechanistic investigation of $NO_2^{\bullet}$ reactions. A radical pathway is the most common pathway for $NO_2^{\bullet}$. However, an ionic pathway is also possible involving NO^+ and NO_3^-, which result from dissociation of N_2O_4, the dimer of $NO_2^{\bullet}$. Short peptides (up to tetrapeptides) have been found to undergo fragmentation-rearrangement reactions when exposed to $NO_2^{\bullet}$, shortening a peptide chain by expelling the least steric hindered residue (usually glycine). The side chain in the short peptides is varied in this work to investigate the effect of steric bulk on the fragmentation-rearrangement reaction rate. Furthermore, the products arising from the exposure of short peptides and cholesterol are used to indicate the reaction pathway adopted by $NO_2^{\bullet}$ in its reaction with these biological molecules.

Chapter Two: Synthesis of model compounds

One approach to investigate the oxidative damage of biological systems upon the exposure to air pollutants is by using model compounds that are present at biological interfaces. Proteins are the major component as they comprise more than 70% of the mass of the non-water components in various biological materials, such as liver, leukocytes and plasma.[23] Therefore, amino acids and short-chain peptides were chosen to be the simplest possible models for this abundant biomolecule. Cholesterol will also be used in the latter chapter as a minor supplement to this work. Any inhaled air pollutant would be expected to directly interact with cholesterol, considering it is the most abundant among other lipids in the epithelial lining fluid.[24]

This chapter presents synthetic methodologies of all model compounds used in this work, which are divided into several subchapters according to the appearance of the model compounds in the subsequent chapters. All amino acids and peptides had L-configuration, where relevant, for the simplicity of the spectroscopic data (e.g., NMR spectra and HPLC chromatograms). In most cases, zwitterionic structures of amino acids would not be drawn for the simplicity of the structure. All amino acids were obtained commercially (Sigma Aldrich, AK Scientific, Chem-Supply or Chem-Impex) and most of them are proteinogenic amino acids, except for 2-aminoisobutyric acid (Aib), L-pyroglutamic acid (Glp) and *N*-methylglycine (sarcosine). The N- and C-termini were protected to enhance their solubility in the organic solvent (acetonitrile) used in kinetic and product studies. Initially, the N-terminus was protected by an acetyl group to mimic the peptide bond, whereas the C-terminus was protected as a methyl ester for its lack of reactivity and ease of synthesis. Different N- and C-terminal protecting groups

were then explored to supplement the initial findings from the N-acetyl and C-terminal methyl ester group.

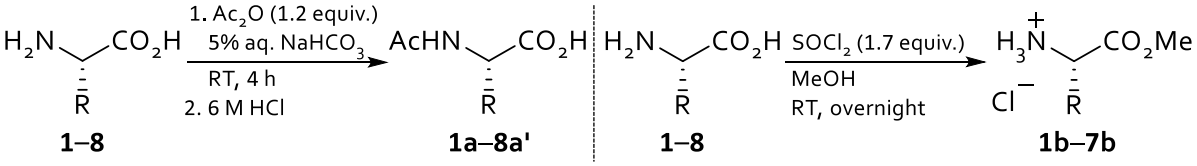
All model amino acids and short peptides obtained using the methods described in this chapter were purified by recrystallisation from hot solvents or column chromatography and characterised by ^{1}H and ^{13}C NMR spectroscopy and high-resolution mass spectrometry (HRMS) analysis. The purity was confirmed by analytical HPLC prior to all experiments with the oxidising radicals. Optimisation of the reaction yield was not attempted because synthesis was not the main focus of this work. Detailed reaction conditions and spectroscopic data are given in the experimental section Chapter 7 page 150.

2.1 Synthesis of N- and C-protected aliphatic amino acids and peptides

2.1.1 Synthesis of N-acetylated aliphatic amino acid and short peptide methyl esters

N-Acetylated amino acids and amino acid methyl esters were either obtained commercially or synthesised using established methods and employed as precursors to acquire N- and C-protected amino acids and short-chain peptides.[25] Table 2.1-1 shows the synthetic scheme of N-acetylation and C-terminal methylation including the yield of each reaction. The amino acid methyl esters were isolated as the hydrochloride salts. In some cases, the side chain of the amino acid was also acetylated or methylated depending on the reaction conditions and the nature of the side chain. Protection of functional groups in the side chain was required to exclude any unwanted reactions with the oxidising radicals occurring at this site.

Table 2.1-1. Synthesis of N-acetylated amino acids **1a–8a'** and amino acid methyl esters **1b–7b**.

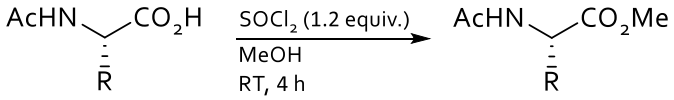


Amino acid	R	Acetylation product	Yield	Methylation product	Yield
Gly (**1**)	H	Ac-Gly (**1a**)	–[a]	Gly-OMe·HCl (**1b**)	97%
Ala (**2**)	Me	Ac-Ala (**2a**)	20%	Ala-OMe·HCl (**2b**)	99%
Val (**3**)	*i*-Pr	Ac-Val (**3a**)	–[a]	Val-OMe·HCl (**3b**)	99%
Leu (**4**)	*i*-Bu	Ac-Leu (**4a**)	89%	Leu-OMe·HCl (**4b**)	97%
Ile (**5**)	*sec*-Bu	Ac-Ile (**5a**)	52%	–[b]	–
Glu (**6**)	$(CH_2)_2CO_2H$	Ac-Glu (**6a**)	61%	Glu(OMe)-OMe·HCl (**6b'**)[c]	92%
Ser (**7**)	CH_2OH	–[b]	–	Ser-OMe·HCl (**7b**)	99%
Lys (**8**)	$(CH_2)_4NH_2$	Ac-Lys(NHAc) (**8a'**)[d]	24%	–[b]	–

[a] The material was obtained commercially (AK Scientific). [b] Not synthesised in this study. [c] The carboxylic acid moiety in the side chain was protected as a methyl ester. To achieve this side-chain methylation, 3.4 equiv. of $SOCl_2$ was used. [d] The amino group in the side chain was protected by an acetyl group. To achieve this side-chain acetylation, 2.4 equiv. of Ac_2O was used.

The N-acetylated amino acid methyl esters were generally obtained by C-terminal methylation of the respective N-acetylated amino acids according to the reaction scheme shown in Table 2.1-2 below.[26] Throughout this work, the N-acetylated amino acid methyl esters, when necessary, would be written as Ac-[**aa**]-OMe, where **aa** is amino acid.

Table 2.1-2. Synthesis of N-acetylated amino acid methyl esters **1c**–**6c**, **8c'**.

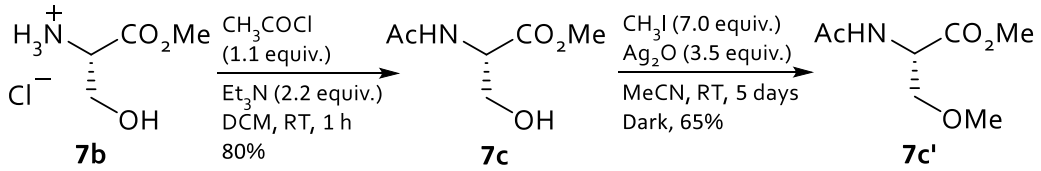


1a–6a, 8a' **1c–6c, 8c'**

Substrate	R	Product	Yield
Ac-Gly (**1a**)	H	Ac-Gly-OMe (**1c**)	37%
Ac-Ala (**2a**)	Me	Ac-Ala-OMe (**2c**)	74%
Ac-Val (**3a**)	*i*-Pr	Ac-Val-OMe (**3c**)	90%
Ac-Leu (**4a**)	*i*-Bu	Ac-Leu-OMe (**4c**)	88%
Ac-Ile (**5a**)	*sec*-Bu	Ac-Ile-OMe (**5c**)	82%
Ac-Glu (**6a**)	$(CH_2)_2CO_2H$	Ac-Glu(OMe)-OMe (**6c'**)[a]	84%
Ac-Lys(NHAc) (**8a'**)	$(CH_2)_4NHAc$	Ac-Lys(NHAc)-OMe (**8c'**)	17%

[a] The carboxylic acid moiety in the side chain was protected as a methyl ester. To achieve this side-chain methylation, 2.4 equiv. of $SOCl_2$ was used.

To obtain Ac-Ser(OMe)-OMe (**7c'**), Ser-OMe·HCl (**7b**), which was obtained according to the reaction scheme in Table 2.1-1, was first acetylated at the N-terminus using acetyl chloride, CH_3COCl, under basic conditions to afford Ac-Ser-OMe (**7c**). Side-chain methylation was then carried out through the reaction of **7c** with methyl iodide (CH_3I) and silver(I) oxide (Ag_2O) under the exclusion of light to give the serine derivative **7c'** (Scheme 2.1-1).[27]

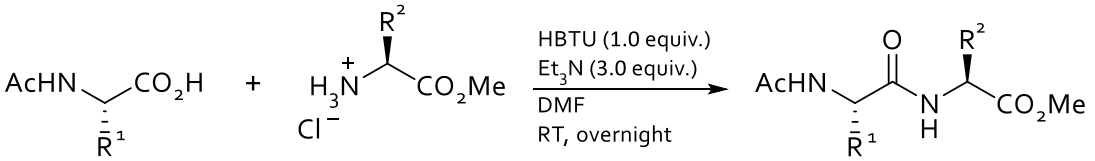


7b **7c** **7c'**

Scheme 2.1-1. Synthesis of Ac-Ser(OMe)-OMe (**7c'**) from Ser-OMe·HCl (**7b**).

The N-acetylated dipeptide methyl esters were obtained by coupling the N-acetylated amino acid to the hydrochloride salt of the amino acid methyl ester under basic conditions using 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) as the coupling agent (see the reaction scheme in Table 2.1-3).[25]

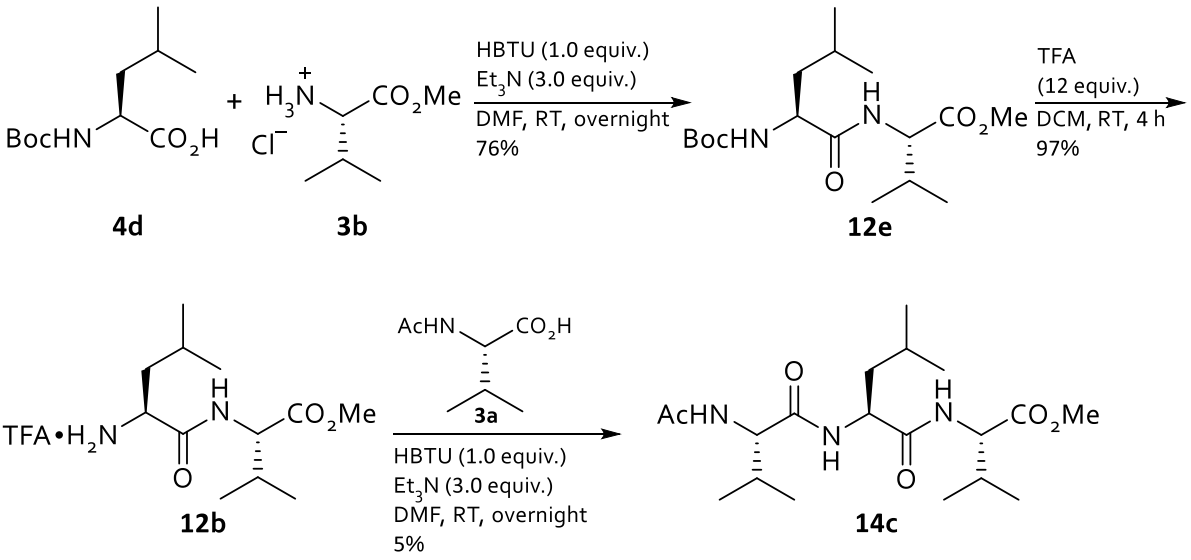
Table 2.1-3. Synthesis of aliphatic dipeptide derivatives **9c–13c**.



Substrate 1 (**3a, 4a**) **Substrate 2** (**2b–4b**) **Dipeptide** (**9c–13c**)

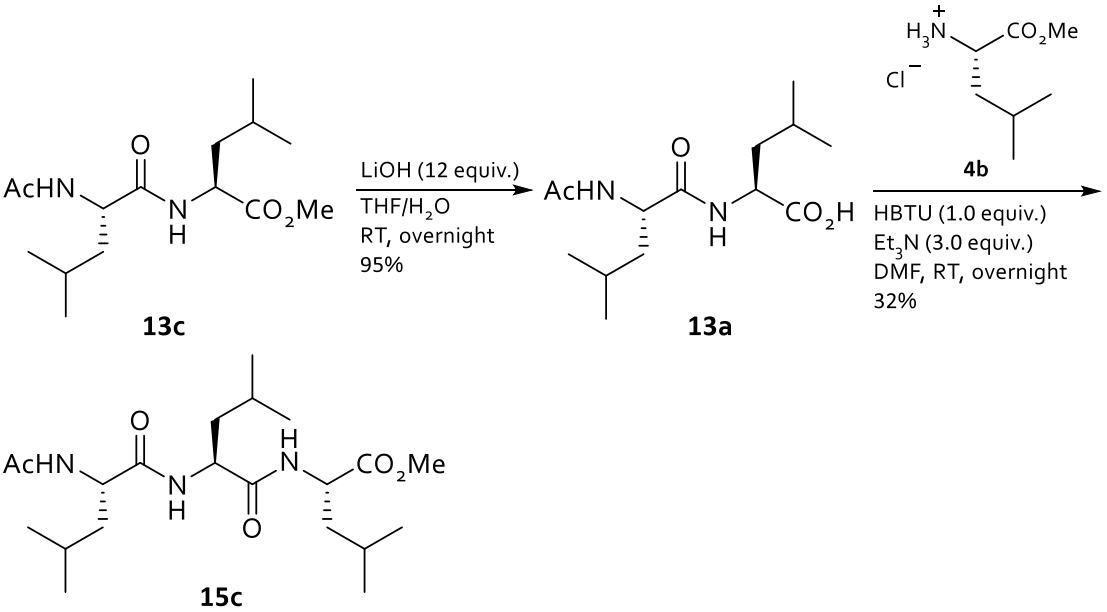
Substrate 1	**R[1]**	**Substrate 2**	**R[2]**	**Dipeptide**	**Yield**
Ac-Val (**3a**)	*i*-Pr	Val-OMe·HCl (**3b**)	*i*-Pr	Ac-Val-Val-OMe (**9c**)	77%
Ac-Val (**3a**)	*i*-Pr	Leu-OMe·HCl (**4b**)	*i*-Bu	Ac-Val-Leu-OMe (**10c**)	25%
Ac-Leu (**4a**)	*i*-Bu	Ala-OMe·HCl (**2b**)	Me	Ac-Leu-Ala-OMe (**11c**)	14%
Ac-Leu (**4a**)	*i*-Bu	Val-OMe·HCl (**3b**)	*i*-Pr	Ac-Leu-Val-OMe (**12c**)	21%
Ac-Leu (**4a**)	*i*-Bu	Leu-OMe·HCl (**4b**)	*i*-Bu	Ac-Leu-Leu-OMe (**13c**)	72%

Tripeptides were synthesised through a sequential procedure either from the C-terminus via N-Boc protected intermediates (Boc = *tert*-butoxycarbonyl) or from the N-terminus via N-acetylated dipeptides. Ac-Val-Leu-Val-OMe (**14c**) was prepared from the C-terminus (Scheme 2.1-2), starting from the coupling of Boc-Leu (**4d**; commercially obtained from AK Scientific) to Val-OMe·HCl (**3b**; prepared according to the reaction scheme in Table 2.1-1, page 15) to produce Boc-Leu-Val-OMe (**12e**). Treatment of dipeptide **12e** with trifluoroacetic acid (TFA) to give the trifluoroacetate salt of Leu-Val-OMe (**12b**),[28] followed by coupling to commercially obtained Ac-Val (**3a**; AK Scientific), would finally afford tripeptide **14c**.



Scheme 2.1-2. Sequential synthesis of tripeptide **14c** from the C-terminus.

Tripeptide Ac-Leu-Leu-Leu-OMe (**15c**) was accessed from the N-terminus through the coupling of Ac-Leu-Leu (**13a**) to Leu-OMe·HCl (**4b**) according to Scheme 2.1-3. The N-acetylated dipeptide intermediate **13a** was obtained from hydrolysis of the methyl ester in Ac-Leu-Leu-OMe (**13c**),[29] which was synthesised following the reaction in Table 2.1-3, page 17. Leu-OMe·HCl (**4b**) was obtained according to the reaction in Table 2.1-1, page 15.



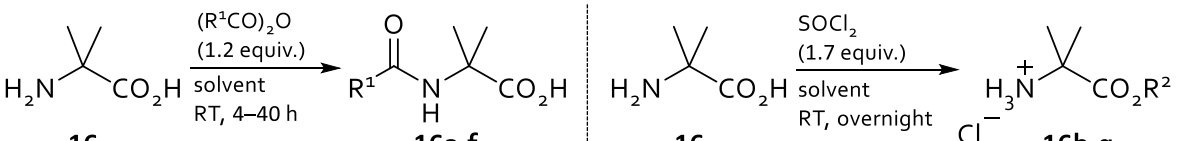
Scheme 2.1-3. Sequential synthesis of tripeptide **15c** from the N-terminus.

The N- and C-protected amino acids **1c–6c'** were employed in the laser flash experiments in Section 3.3.2, page 54, whereas amino acids **7c'** and **8c'** were used in Section 3.3.3.3, page 67. Dipeptides **9c–13c** and tripeptides **14c** and **15c** were used in the laser flash photolysis experiments in Section 3.3.2, page 58. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.1.2 Synthesis of 2-aminoisobutyric acid (Aib) derivatives

All proteinogenic amino acids have at least one hydrogen at the α-position. However, in the case of 2-aminoisobutyric acid (Aib), the α-position is substituted by two methyl groups. The N- and C-protected Aib derivatives were employed as a model compound to explore the reactivity of amino acids without α-hydrogen atoms. The synthesis was achieved through either N- or C-terminal protection of the commercially obtained Aib (**16**; Sigma-Aldrich), followed by the instalment of the other terminal protecting group. The synthesis of the intermediates is shown in the reaction schemes in Table 2.1-4.[25]

Table 2.1-4. Synthesis of N-protected or C-protected Aib intermediates **16a**,**b**,**f**,**g**.



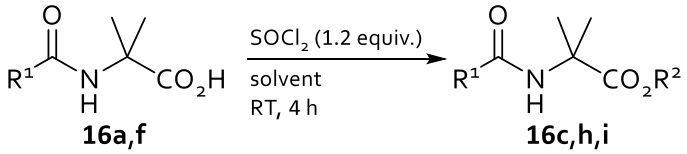
16 **16a,f** **16** **16b,g**

R^1	Solvent	Product	Yield	R^2	Solvent	Product	Yield
Me	5% aq. $NaHCO_3$	Ac-Aib (**16a**)	38%	Me	MeOH	Aib-OMe·HCl (**16b**)	96%
CF_3	CF_3CO_2H	Tfa-Aib (**16f**)[a]	72%	CD_3	CD_3OD	Aib-OCD_3·HCl (**16g**)	97%

[a] Tfa = trifluoroacetyl. The synthesis was based on Gordon-Wylie et al.[30]

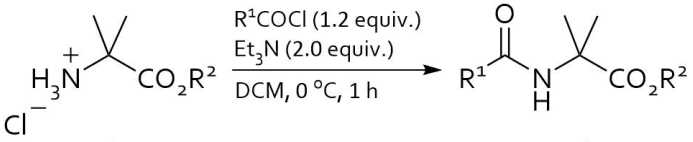
Reaction schemes in Table 2.1-5 show the final step to obtain the N- and C-protected Aib derivatives **16c**,**h–l**. Standard methylation procedure as described previously in Section 2.1.1 was used to obtain N- and C-protected Aib derivatives **16c**,**h**,**i** from **16a**,**f**.[26] Meanwhile, reaction of the hydrochloride salts of Aib methyl esters **16b**,**g** with an acyl chloride under basic conditions gave N- and C-protected Aib derivatives **16j–l**.[27]

Table 2.1-5. Synthesis of N- and C-protected Aib derivatives **16c**,**h–l**.



16a,f **16c,h,i**

Substrate	R^1	Solvent	Product	R^2	Yield
Ac-Aib (**16a**)	Me	MeOH	Ac-Aib-OMe (**16c**)	Me	47%
Ac-Aib (**16a**)	Me	CD_3OD	Ac-Aib-OCD_3 (**16h**)	CD_3	60%
Tfa-Aib (**16f**)	CF_3	MeOH	Tfa-Aib-OMe (**16i**)	Me	38%



16b,g **16j–l**

Substrate	R^2	R^1	Product	Yield
Aib-OMe·HCl (**16b**)	Me	CD_3	Ac(d_2)-Aib-OMe (**16j**)[a]	60%
Aib-OCD_3·HCl (**16g**)	CD_3	CD_3	Ac(d_3)-Aib-OCD_3 (**16k**)	81%
Aib-OMe·HCl (**16b**)	Me	*t*-Bu	Piv-Aib-OMe (**16l**)[b]	55%

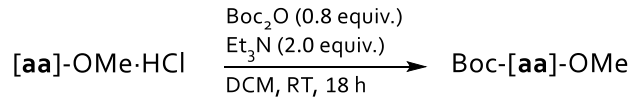
[a] Only about 20% is fully deuterated as Ac(d_3)-Aib-OMe based on NMR and HRMS.
[b] Piv = pivaloyl (trimethylacetyl).

Amino acids **16c**,**h**,**j–l** were used in the laser flash experiments in Section 3.3.2, page 56, while amino acid **16i** was used in Section 3.3.3.2, page 64. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.1.3 Synthesis of *N*-Boc amino acid and peptide methyl esters

N-Boc amino acid methyl esters were prepared starting from methylation of the C-terminus of amino acid according to the reaction scheme in Table 2.1-1, page 15 and Table 2.1-4, page 19. The obtained hydrochloride salts of the amino acid methyl esters were then reacted with Boc anhydride under basic conditions to give amino acids **1e**, **2e**, **4e**, **16e**.[31]

Table 2.1-6. Synthesis of *N*-Boc amino acid methyl esters **1e**, **2e**, **4e**, **16e**.[a]

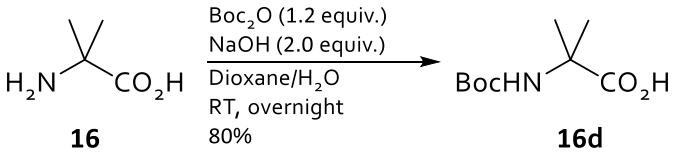


1b, 2b, 4b, 16b **1e, 2e, 4e, 16e**

Substrate	Product	Yield
Gly-OMe·HCl (**1b**)	Boc-Gly-OMe (**1e**)	77%
Ala-OMe·HCl (**2b**)	Boc-Ala-OMe (**2e**)	30%
Aib-OMe·HCl (**16b**)	Boc-Aib-OMe (**16e**)	77%
Leu-OMe·HCl (**4b**)	Boc-Leu-OMe (**4e**)	92%

[a] **aa** = amino acid. For detailed structures of the amino acids, refer to Table 2.1-1 and Table 2.1-4.

N-Boc dipeptide methyl esters were obtained through the coupling of the *N*-Boc amino acid to the hydrochloride salt of the amino acid methyl ester. Most of *N*-Boc amino acids, except Boc-Aib (**16d**), were obtained commercially (AK Scientific or Chem-Impex). The synthesis of **16d** was carried out according to Scheme 2.1-4.[32]

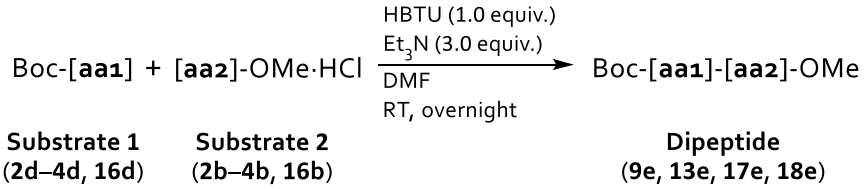


16 **16d**

Scheme 2.1-4. Synthesis of Boc-Aib (**16d**) from Aib (**16**).

The coupling was performed in the same fashion as the N-acetylated dipeptide methyl ester in Table 2.1-3, page 17 using HBTU as the coupling agent.

Table 2.1-7. Synthesis of *N*-Boc dipeptide methyl esters **9e**, **13e**, **17e**, **18e**.[a]

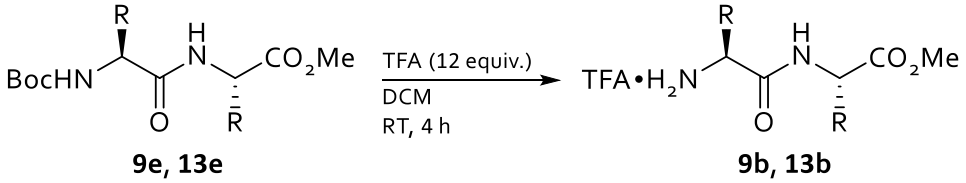


Substrate 1	Substrate 2	Dipeptide	Yield
Boc-Ala (**2d**)	Ala-OMe·HCl (**2b**)	Boc-Ala-Ala-OMe (**17e**)	28%
Boc-Aib (**16d**)	Aib-OMe·HCl (**16b**)	Boc-Aib-Aib-OMe (**18e**)	69%
Boc-Val (**3d**)	Val-OMe·HCl (**3b**)	Boc-Val-Val-OMe (**9e**)	89%
Boc-Leu (**4d**)	Leu-OMe·HCl (**4b**)	Boc-Leu-Leu-OMe (**13e**)	72%

[a] **aa1** = amino acid 1; **aa2** = amino acid 2. For detailed structures of the amino acids, refer to Table 2.1-1 and Table 2.1-4.

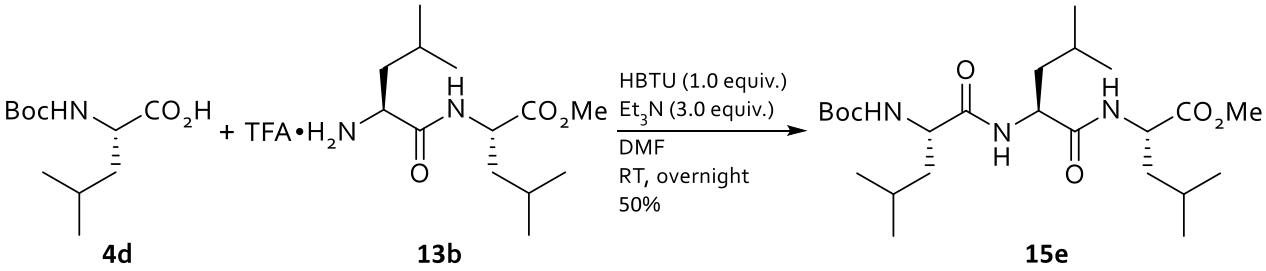
Some of the *N*-Boc dipeptide methyl esters were used as the intermediate for accessing tripeptides. The Boc-protecting group was removed using trifluoroacetic acid in dichloromethane to yield the trifluoroacetate salt of the dipeptide methyl ester (see the reaction scheme in Table 2.1-8), which can later be coupled to an N-protected amino acid to obtain a tripeptide.

Table 2.1-8. Synthesis of the dipeptide methyl ester intermediates **9b** and **13b**.



Substrate	R	Product	Yield
Boc-Val-Val-OMe (**9e**)	*i*-Pr	Val-Val-OMe·TFA (**9b**)	92%
Boc-Leu-Leu-OMe (**13e**)	*i*-Bu	Leu-Leu-OMe·TFA (**13b**)	98%

Tripeptide Boc-Leu-Leu-Leu-OMe (**15e**) was synthesised by coupling Boc-Leu (**4d**) to dipeptide **13b** using the standard coupling procedure (Scheme 2.1-5).

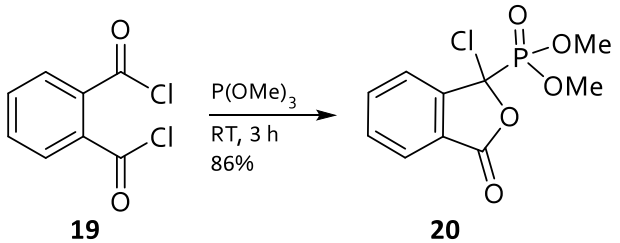


Scheme 2.1-5. Synthesis of Boc-Leu-Leu-Leu-OMe (**15e**).

The *N*-Boc amino acid and peptide methyl esters **1e**, **2e**, **4e**, **13e**, **15e–18e** were used in the laser flash experiments in Section 3.3.3.1, page 60. Meanwhile, dipeptide Val-Val-OMe·TFA (**9b**) was used as a precursor for a longer peptide in Section 2.2.4, page 34. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.1.4 Synthesis of N-phthaloylated amino acid and dipeptide methyl esters

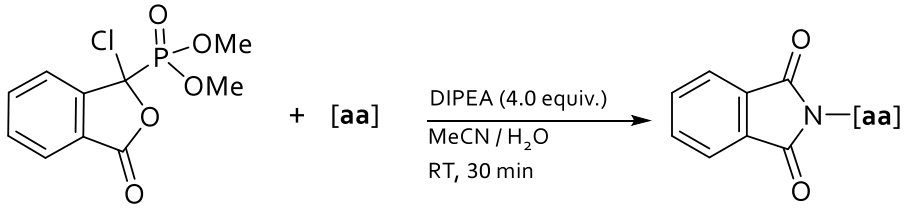
The phthaloyl group was introduced to the amino acid using 3-chloro-3-(dimethoxyphosphoryl)isobenzofuran-1(3*H*)-one (**20**) as the phthaloylating agent. Its synthesis is described in Scheme 2.1-6 below from the Arbuzov reaction of phthaloyl dichloride (**19**) with trimethyl phosphite.[33]



Scheme 2.1-6. Synthesis of the phthaloylating agent **20**.

This acylphosphonate derivative **20** reacts rapidly with primary amines in amino acids to yield phthalimides in the presence of diisopropylethylamine (DIPEA) according to the reaction scheme in Table 2.1-9.[34]

Table 2.1-9. Synthesis of N-phthaloylated amino acids **3m**, **4m**, **16m**.[a]



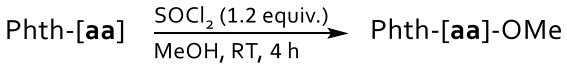
20 **3, 4, 16** (1.1 equiv.) Phth-[**aa**] (**3m, 4m, 16m**)

[**aa**]	**Product**	**Yield**
Aib (**16**)	Phth-Aib (**16m**)	47%
Val (**3**)	Phth-Val (**3m**)	42%
Leu (**4**)	Phth-Leu (**4m**)	63%

[a] **aa** = amino acid. For detailed structures of the amino acids, refer to Table 2.1-1 and Table 2.1-4.

C-Terminal methylation of the N-phthaloylated amino acids using thionyl chloride in methanol afforded the N-phthaloylated amino acid methyl esters according to the reaction scheme in Table 2.1-10, except for Phth-Aib-OMe (**16n**).

Table 2.1-10. Synthesis of N-phthaloylated amino acid methyl esters **1n**, **3n**, **4n**.[a]

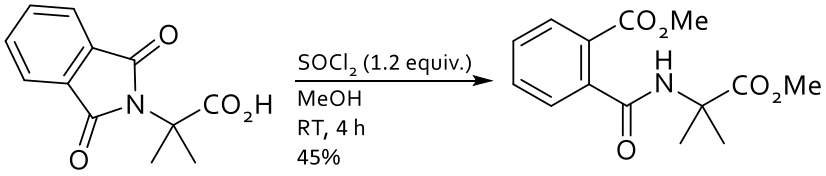


1m, 3m, 4m, 16m **1n, 3n, 4n**

Substrate	[**aa**]	**Product**	**Yield**
Phth-Gly (**1m**)[b]	Gly	Phth-Gly-OMe (**1n**)	91%
Phth-Aib (**16m**)	Aib	–[c]	–
Phth-Val (**3m**)	Val	Phth-Val-OMe (**3n**)	81%
Phth-Leu (**4m**)	Leu	Phth-Leu-OMe (**4n**)	70%

[a] **aa** = amino acid. For detailed structures of the amino acids, refer to Table 2.1-1 and Table 2.1-4. [b] Phth-Gly (**1m**) was provided by Dr. Luke Gamon. [c] Phth-Aib-OMe (**16n**) could not be obtained through this method, amino acid **16o** was isolated instead (see text below).

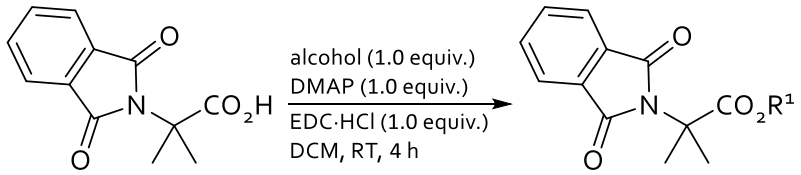
Under the conditions described in Table 2.1-10, partial deprotection of the phthalimide **16m** was observed. Instead of the expected Phth-Aib-OMe (**16n**), amino acid (*o*-CO_2Me)Bz-Aib-OMe (**16o**) was obtained as the major product (Scheme 2.1-7).



Scheme 2.1-7. Serendipitous synthesis of amino acid **16o** from Phth-Aib (**16m**).

Milder conditions were then employed to access Phth-Aib-OMe (**16n**) and Phth-Aib-OCD_3 (**16p**) through the coupling of Phth-Aib (**16m**) to methanol or deuterated methanol, using *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC·HCl) and 4-dimethylaminopyridine (DMAP) as the coupling agents (see the reaction scheme in Table 2.1-11).[35]

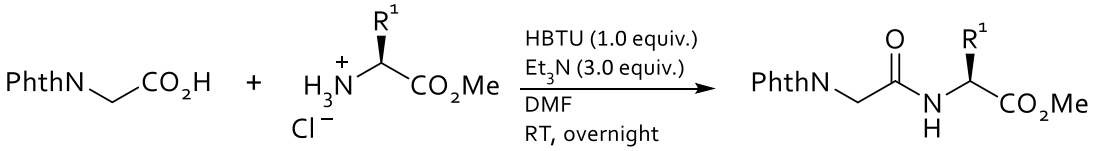
Table 2.1-11. Synthesis of N-phthaloylated Aib derivatives **16n** and **16p**.



Alcohol	R^1	Product	Yield
MeOH	Me	Phth-Aib-OMe (**16n**)	46%
CD_3OD	CD_3	Phth-Aib-OCD_3 (**16p**)	62%

N-Phthaloylated dipeptide methyl esters were obtained through the coupling of the N-phthaloylated amino acid to the hydrochloride salt of the amino acid methyl ester using HBTU as the coupling agent (see the reaction scheme in Table 2.1-12).

Table 2.1-12. Synthesis of N-phthaloylated dipeptide methyl esters **21n–23n**.[a]



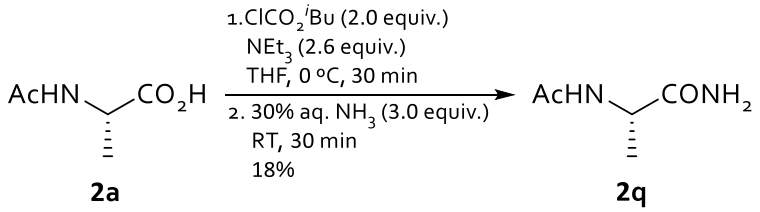
Substrate 1 (**1m**) | **Substrate 2** (**1b, 3b, 4b**) | **Dipeptide** (**21n–23n**)

Substrate 1	Substrate 2	R^1	Product	Yield
Phth-Gly (**1m**)	Gly-OMe·HCl (**1b**)	H	Phth-Gly-Gly-OMe (**21n**)	69%
Phth-Gly (**1m**)	Val-OMe·HCl (**3b**)	*i*-Pr	Phth-Gly-Val-OMe (**22n**)	89%
Phth-Gly (**1m**)	Leu-OMe·HCl (**4b**)	*i*-Bu	Phth-Gly-Leu-OMe (**23n**)	61%

The N-phthaloylated amino acid and dipeptide methyl esters **1n**, **3n**, **4n**, **16n–p**, and **21n–23n** were used in the laser flash experiments in Section 3.3.3.2, page 64. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.1.5 Synthesis of N-acetylated amino acid amide derivatives

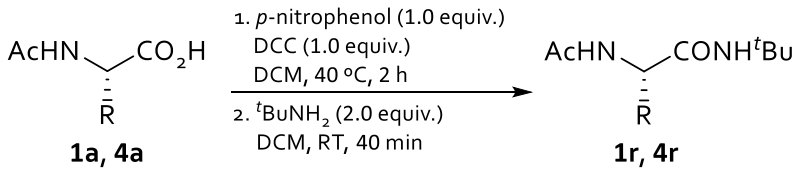
Apart from modifications at the N-terminus with various protecting groups, the C-terminus of aliphatic amino acids were also altered from methyl esters to amides. The N-acetylated amino acid amide was obtained through the coupling of the N-acetylated amino acid to ammonia (NH_3) using isobutyl chloroformate, $ClCO_2{}^iBu$, as the coupling agent.[36] Scheme 2.1-8 shows the synthesis of Ac-Ala-NH_2 (**2q**) from Ac-Ala (**2a**). The use of 30% aqueous ammonia solution limited the reaction yield due to the presence of water as a possible competing nucleophile.



Scheme 2.1-8. Synthesis of Ac-Ala-NH_2 (**2q**).

The synthesis of *t*-butyl amides was achieved through the coupling of the N-acetylated amino acid to *p*-nitrophenol ester using *N*,*N*'-dicyclohexylcarbodiimide (DCC) as the coupling agent to generate the *p*-nitrophenol ester. This intermediate was then reacted with *t*-butylamine to yield the N-acetylated amino acid *t*-butyl amides **1r** and **4r** (see the reaction scheme in Table 2.1-13).[37]

Table 2.1-13. Synthesis of *t*-butyl amides **1r** and **4r**.



Substrate	R	Product	Yield
Ac-Gly (**1a**)	H	Ac-Gly-NH^{t}Bu (**1r**)	2%
Ac-Leu (**4a**)	*i*-Bu	Ac-Leu-NH^{t}Bu (**4r**)	7%

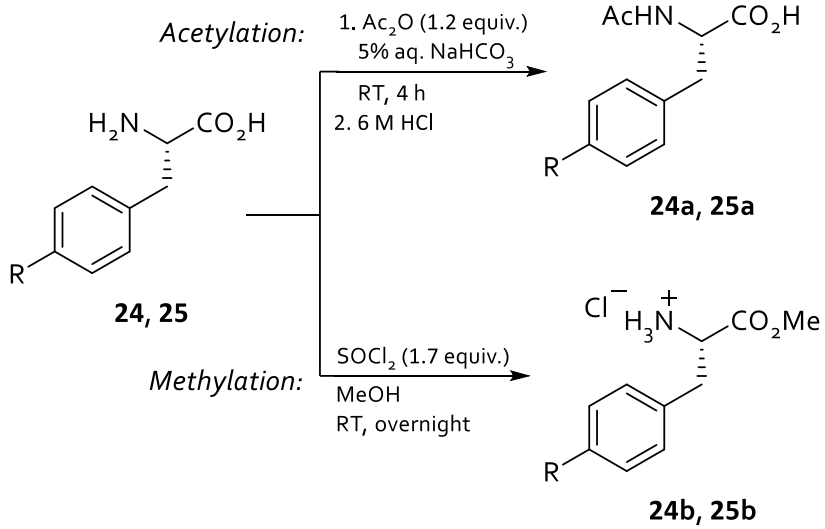
The yield was low due to multiple steps of purification (column chromatography) and recrystallisation to obtain pure products for the kinetic studies. These amino acid derivatives **2q**, **1r** and **4r** were used in the laser flash experiments in Section 3.3.2, page 59. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.2 Synthesis of N- and C-protected aromatic amino acids and peptides

2.2.1 Synthesis of N-acetylated aromatic amino acid methyl esters

Similar to the N- and C-protected aliphatic amino acids in Section 2.1, N-acetylated aromatic amino acids and amino acid methyl esters were either obtained commercially or synthesised using established methods and employed as precursors to acquire N- and C-protected amino acids and short peptides as the intended model compounds.[25] Table 2.2-1 shows the reaction scheme of N-acetylation and C-terminal methylation including the yield of each reaction.

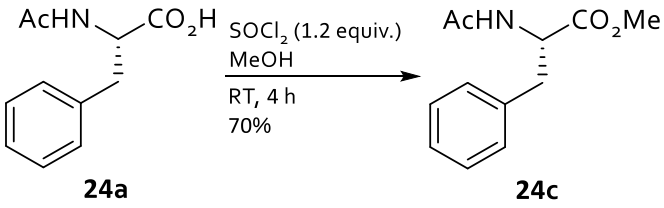
Table 2.2-1. Synthesis of N-acetylated aromatic amino acids **24a**, **25a** and hydrochloride salts of aromatic amino acid methyl esters **24b**, **25b**.



Amino acid	R	Acetylation product	Yield	Methylation product	Yield
Phe (**24**)	H	Ac-Phe (**24a**)	90%	Phe-OMe·HCl (**24b**)	99%
Tyr (**25**)	OH	Ac-Tyr (**25a**)	–[a]	Tyr-OMe·HCl (**25b**)	97%

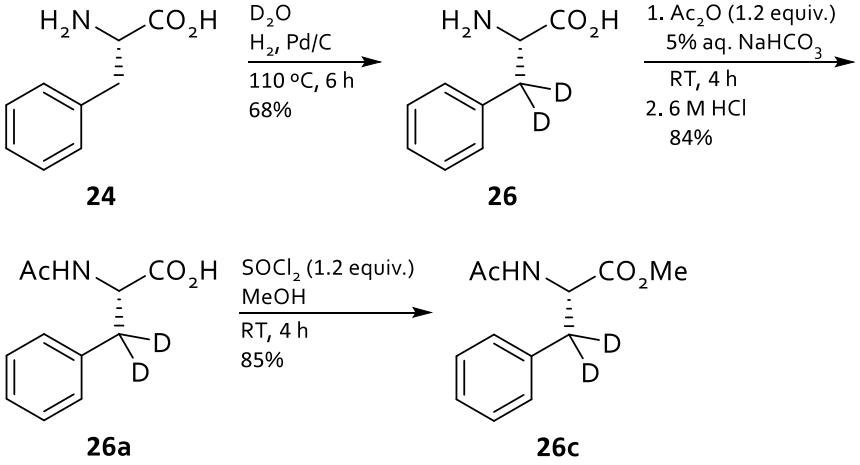
[a] The material was obtained commercially (AK Scientific).

Ac-Phe-OMe (**24c**) was synthesised through standard methylation of Ac-Phe (**24a**) (Scheme 2.2-1), while Ac-Tyr-OMe (**25c**) was provided by Dr. Luke Gamon and used as a precursor for another model compound in Section 2.2.5, page 36.



Scheme 2.2-1. Synthesis of Ac-Phe-OMe (**24c**).

Selective deuteration of Phe **24** at the β position for kinetic isotope studies (discussed in Section 4.3.1, page 89) was achieved using palladium on carbon (10% wt.) as the catalyst under hydrogen atmosphere in deuterated water at 110 °C to give Phe(β,β-d_2) **26**.[38] N-Acetylation of **26**, followed by C-terminal methylation, yielded Ac-Phe(β,β-d_2)-OMe (**26c**) as the final product (Scheme 2.2-2).[25,26]



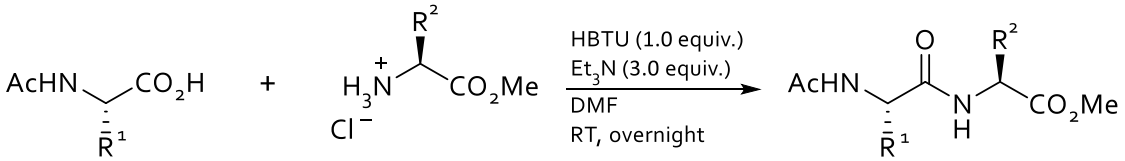
Scheme 2.2-2. Synthesis of Ac-Phe(β,β-d_2)-OMe (**26c**).

Intermediates **24a**,**b** and **25a**,**b** were used to access aromatic dipeptides and tripeptides in Section 2.2.2–2.2.4. Amino acids **24c** and **26c** were used in the laser flash experiment in Section 4.3.1, page 88. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.2.2 Synthesis of N-acetylated aromatic dipeptide methyl esters

Aromatic amino acids were coupled to various aliphatic and aromatic amino acids to obtain the N- and C-protected aromatic-containing dipeptides. The N-protected amino acids were coupled to the C-protected amino acids using HBTU as the coupling agent (see the reaction scheme in Table 2.2-2).[25] The precursors (Substrate 1 and 2) were synthesised according to the reaction schemes either in Table 2.1-1, page 15 or Table 2.2-1, page 28.

Table 2.2-2. Synthesis of N-acetylated aromatic dipeptide methyl esters **27c–38c**.



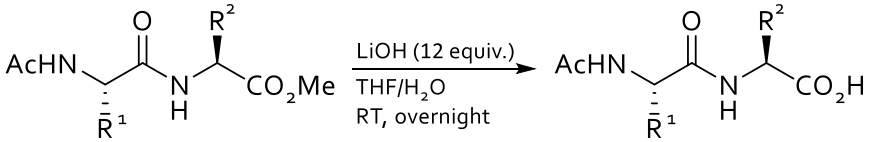
Substrate 1 (**1a, 3a, 4a, 24a, 25a**) **Substrate 2** (**1b, 3b, 4b, 6b', 24b, 25b**) **Dipeptide** (**27c–38c**)

Substrate 1	**R^1**	**Substrate 2**	**R^2**	**Dipeptide**	**Yield**
Ac-Gly (**1a**)	H	Phe-OMe·HCl (**24b**)	Bn	Ac-Gly-Phe-OMe (**27c**)	61%
Ac-Val (**3a**)	*i*-Pr	Phe-OMe·HCl (**24b**)	Bn	Ac-Val-Phe-OMe (**28c**)	45%
Ac-Leu (**4a**)	*i*-Bu	Phe-OMe·HCl (**24b**)	Bn	Ac-Leu-Phe-OMe (**29c**)	38%
Ac-Phe (**24a**)	Bn	Gly-OMe·HCl (**1b**)	H	Ac-Phe-Gly-OMe (**30c**)	39%
Ac-Phe (**24a**)	Bn	Val-OMe·HCl (**3b**)	*i*-Pr	Ac-Phe-Val-OMe (**31c**)	32%
Ac-Phe (**24a**)	Bn	Leu-OMe·HCl (**4b**)	*i*-Bu	Ac-Phe-Leu-OMe (**32c**)	81%
Ac-Phe (**24a**)	Bn	Glu(OMe)-OMe·HCl (**6b'**)	$(CH_2)_2CO_2Me$	Ac-Phe-Glu(OMe)-OMe (**33c'**)	53%
Ac-Phe (**24a**)	Bn	Phe-OMe·HCl (**24b**)	Bn	Ac-Phe-Phe-OMe (**34c**)	64%
Ac-Phe (**24a**)	Bn	Tyr-OMe·HCl (**25b**)	(*p*-OH)C_6H_4	Ac-Phe-Tyr-OMe (**35c**)	54%

Ac-Tyr (**25a**)	(*p*-OH)C_6H_4	Leu-OMe·HCl (**4b**)	*i*-Bu	Ac-Tyr-Leu-OMe (**36c**)	61%
Ac-Tyr (**25a**)	(*p*-OH)C_6H_4	Phe-OMe·HCl (**24b**)	Bn	Ac-Tyr-Phe-OMe (**37c**)	48%
Ac-Tyr (**25a**)	(*p*-OH)C_6H_4	Tyr-OMe·HCl (**25b**)	(*p*-OH)C_6H_4	Ac-Tyr-Tyr-OMe (**38c**)	60%

Some of these dipeptides were used as an intermediate for tripeptides (Section 2.2.4, page 34) or tyrosine-containing dipeptide derivatives (Section 2.2.5, page 36). In order to access the tripeptides from the N-terminus, the methyl ester group in dipeptides was deprotected according to the reaction scheme in Table 2.2-3, before the resulting dipeptide intermediate is coupled to a hydrochloride salt of a methyl ester.

Table 2.2-3. Synthesis of N-acetylated dipeptide intermediates **29a**, **32a**, **34a**.[a]



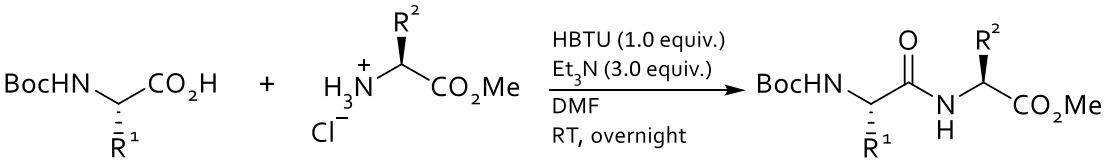
29c, 32c, 34c **29a, 32a, 34a**

Dipeptide	Product	Yield
Ac-Leu-Phe-OMe (**29c**)	Ac-Leu-Phe (**29a**)	79%
Ac-Phe-Leu-OMe (**32c**)	Ac-Phe-Leu (**32a**)	79%
Ac-Phe-Phe-OMe (**34c**)	Ac-Phe-Phe (**34a**)	100%[b]

[a] For detailed structures of the dipeptide, refer to Table 2.2-2. [b] Quantitative yield was obtained.

The N-Boc protected dipeptide methyl esters **27e–29e**, **31e** were synthesised and employed as intermediates to obtain some longer peptides in Section 2.2-4, page 34. HBTU was used to couple the commercially obtained *N*-Boc amino acid (AK Scientific or Chem-Impex) to the hydrochloride salt of the amino acid methyl ester according to the reaction scheme in Table 2.2-4.

Table 2.2-4. Synthesis of *N*-Boc dipeptide methyl esters **27e–29e**, **31e**.

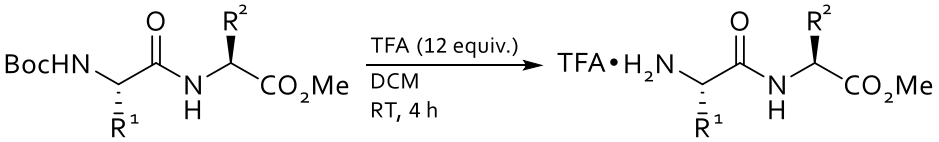


Substrate 1 (**1d, 3d, 4d, 24d**) | **Substrate 2** (**3b, 24b**) | **Dipeptide** (**27e–29e, 31e**)

Substrate 1	R^1	Substrate 2	R^2	Dipeptide	Yield
Boc-Gly (**1d**)	H	Phe-OMe·HCl (**24b**)	Bn	Boc-Gly-Phe-OMe (**27e**)	97%
Boc-Val (**3d**)	*i*-Pr	Phe-OMe·HCl (**24b**)	Bn	Boc-Val-Phe-OMe (**28e**)	83%
Boc-Leu (**4d**)	*i*-Bu	Phe-OMe·HCl (**24b**)	Bn	Boc-Leu-Phe-OMe (**29e**)	85%
Boc-Phe (**24d**)	Ph	Val-OMe·HCl (**3b**)	*i*-Pr	Boc-Phe-Val-OMe (**31e**)	87%

Deprotection of Boc group was carried out by treating dipeptides **27e–29e**, **31e** with trifluoroacetic acid in dichloromethane (see the reaction scheme in Table 2.2-5) to generate the trifluoroacetate salt of dipeptide methyl esters, which can then be coupled to an N-protected amino acid or peptide to obtain a longer peptide.

Table 2.2-5. Synthesis of dipeptide methyl ester intermediates **27b–29b**, **31b**.[a]



27e–29e, 31e | **27b–29b, 31b**

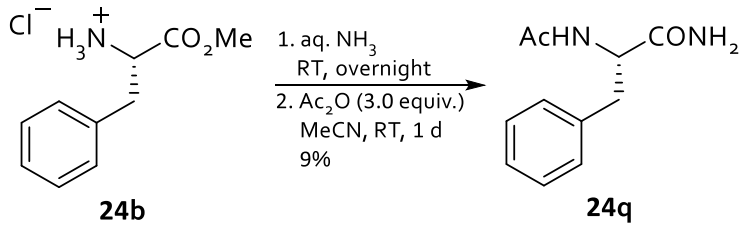
Dipeptide	Product	Yield
Boc-Gly-Phe-OMe (**27e**)	Gly-Phe-OMe·TFA (**27b**)	91%
Boc-Val-Phe-OMe (**28e**)	Val-Phe-OMe·TFA (**28b**)	93%
Boc-Leu-Phe-OMe (**29e**)	Leu-Phe-OMe·TFA (**29b**)	78%
Boc-Phe-Val-OMe (**31e**)	Phe-Val-OMe·TFA (**31b**)	88%

[a] For detailed structures of the dipeptide, refer to Table 2.2-4.

Dipeptides **27c–34c** were used in the laser flash experiments in Section 4.3.1, page 88, while dipeptides **35c–38c** were used as precursors to tyrosine-containing dipeptide derivatives in Section 2.2.5, page 36. Dipeptides **27c–29c**, **34c** were used in the product studies in Section 6.3.1, page 142. The dipeptide intermediates **29a**, **32a**, **34a** and **27b–29b**, **31b** were employed as precursors to build a longer peptide described in Section 2.2.4, page 34. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

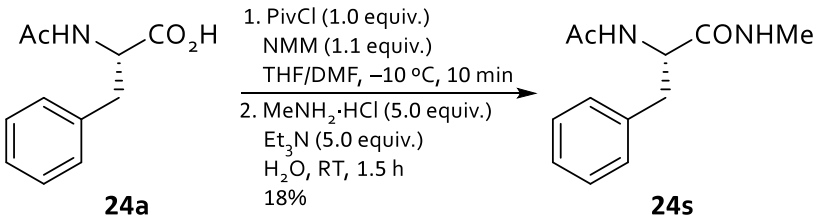
2.2.3 Synthesis of phenylalanine amide derivatives

A series of *N*-acetyl phenylalanine derivatives with the C-terminus protected as an amide were synthesised. The primary amide, Ac-Phe-NH_2 (**24q**), was synthesised from the hydrochloride salt of phenylalanine methyl ester (**24b**). Treating the salt in aqueous ammonia, followed by acetylation using acetic anhydride, yielded Ac-Phe-NH_2 (**24q**) according to Scheme 2.2-3.[39]



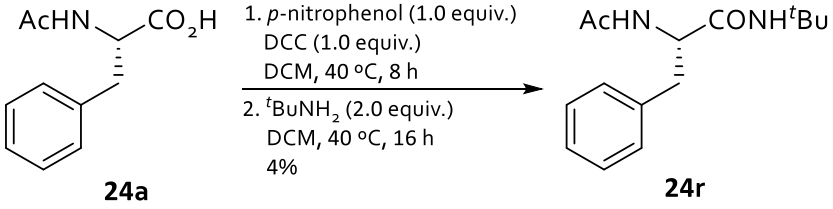
Scheme 2.2-3. Synthesis of Ac-Phe-NH_2 (**24q**).

The secondary amide, Ac-Phe-NHMe (**24s**), was obtained from *N*-acetyl phenylalanine (**24a**) through the in situ generation of a pivaloyl ester intermediate, followed by treatment with methylamine (Scheme 2.2-4).[40]



Scheme 2.2-4. Synthesis of Ac-Phe-NHMe (**24s**).

The other secondary amide derivative, Ac-Phe-NH^{t}Bu (**24r**), was synthesised through an activated *p*-nitrophenol ester intermediate, followed by its reaction with *t*-butylamine (Scheme 2.2-5).[37]



Scheme 2.2-5. Synthesis of Ac-Phe-NH^{t}Bu (**24r**).

In general, the yields for the synthesis of phenylalanine amide derivatives **24q–s** were found to be fairly low. Yield optimisation was not performed because sufficient material was obtained in all cases to carry out the kinetic studies in Section 4.3.2, page 91. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

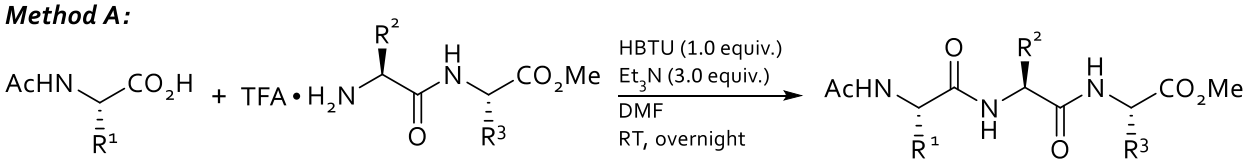
2.2.4 Synthesis of N-acetylated tri- and tetrapeptide methyl esters

The N-acetylated aromatic tripeptide methyl esters were obtained through synthesis starting at either the N- or the C-terminus. Synthesis from the C-terminus was achieved by coupling the N-acetylated amino acid to the trifluoroacetate salt of the dipeptide methyl ester (Method A), while synthesis from

the N-terminus was achieved by coupling the N-acetylated dipeptide to the hydrochloride salt of the amino acid methyl ester (Method B). Table 2.2-6 shows the reaction schemes for both methods and the substrates used for the coupling including the yield of each coupling reaction.

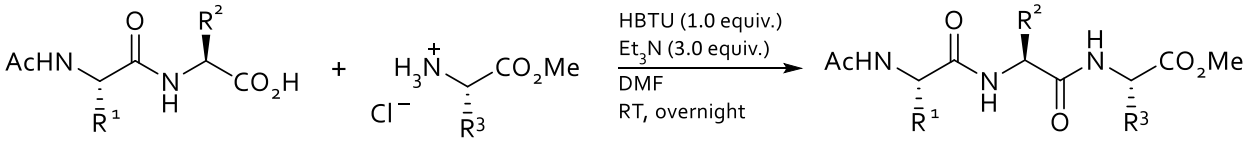
Table 2.2-6. Synthesis of N-acetylated aromatic tripeptide methyl esters **39c–48c**.

Method A:



Substrate 1 (**3a**, **4a**, **24a**) | **Substrate 2** (**9b**, **27b–29b**, **31b**) | **Tripeptide** (**39c–42c**, **45c–47c**)

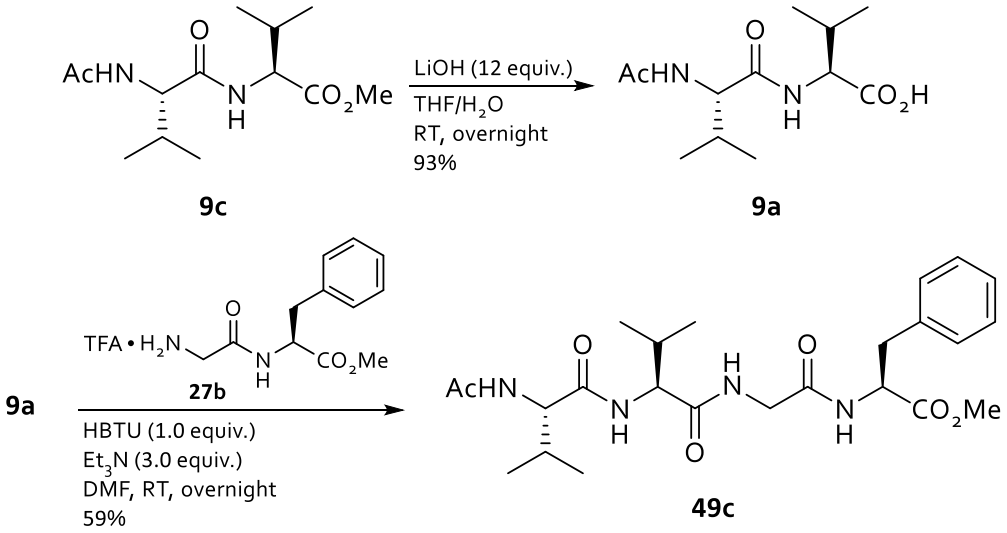
Method B:



Substrate 1 (**29a**, **32a**, **34a**) | **Substrate 2** (**4b**) | **Tripeptide** (**43c**, **44c**, **48c**)

Method	Substrate 1	Substrate 2	Tripeptide	Yield
A	Ac-Val (**3a**)	Val-Phe-OMe·TFA (**28b**)	Ac-Val-Val-Phe-OMe (**39c**)	7%
A	Ac-Val (**3a**)	Phe-Val-OMe·TFA (**31b**)	Ac-Val-Phe-Val-OMe (**40c**)	54%
A	Ac-Phe (**24a**)	Val-Val-OMe·TFA (**9b**)	Ac-Phe-Val-Val-OMe (**41c**)	69%
A	Ac-Leu (**4a**)	Leu-Phe-OMe·TFA (**29b**)	Ac-Leu-Leu-Phe-OMe (**42c**)	43%
B	Ac-Leu-Phe (**29a**)	Leu-OMe·HCl (**4b**)	Ac-Leu-Phe-Leu-OMe (**43c**)	43%
B	Ac-Phe-Leu (**32a**)	Leu-OMe·HCl (**4b**)	Ac-Phe-Leu-Leu-OMe (**44c**)	56%
A	Ac-Phe (**24a**)	Gly-Phe-OMe·TFA (**27b**)	Ac-Phe-Gly-Phe-OMe (**45c**)	43%
A	Ac-Phe (**24a**)	Leu-Phe-OMe·TFA (**29b**)	Ac-Phe-Leu-Phe-OMe (**46c**)	33%
A	Ac-Phe (**24a**)	Phe-Val-OMe·TFA (**31b**)	Ac-Phe-Phe-Val-OMe (**47c**)	65%
B	Ac-Phe-Phe (**34a**)	Leu-OMe·HCl (**4b**)	Ac-Phe-Phe-Leu-OMe (**48c**)	74%

Tetrapeptide Ac-Val-Val-Gly-Phe-OMe (**49c**) was obtained through the coupling of two dipeptides according to Scheme 2.2-6. Ac-Val-Val (**9a**) was prepared from dipeptide Ac-Val-Val-OMe (**9c**) following hydrolysis of the methyl ester, whereas Gly-Phe-OMe·TFA (**27b**) was obtained according to Table 2.2-5 page 32.

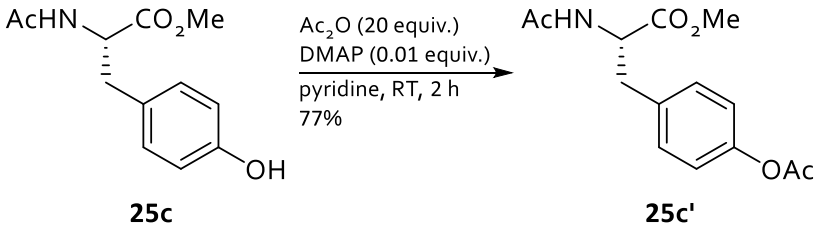


Scheme 2.2-6. Synthesis of tetrapeptide Ac-Val-Val-Gly-Phe-OMe (**49c**).

Tripeptides **39c**–**48c** were used in the laser flash experiments in Section 4.3.3, page 96. Tetrapeptide **49c** was used in the product studies in Section 6.3.2, page 142. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.2.5 Acetylation of tyrosine-containing residues

As previously stated, tyrosine-containing substrates were not used in the laser flash experiments due to their reaction with $NO_3^{\bullet}$ precursor (for more details, see Section 4.3.4, page 99). In order to prevent this reaction, the hydroxyl group in tyrosine residue was acetylated according to Scheme 2.2-7.[25] Ac-Tyr-OMe (**25c**) was provided by Dr. Luke Gamon.



25c **25c'**

Scheme 2.2-7. O-Acetylation of Ac-Tyr-OMe (**25c**).

Under the same reaction conditions, the hydroxyl groups in tyrosine-containing dipeptides **35c**–**38c** were acetylated to yield dipeptides **35c'**–**38c'** (Table 2.2-7).

Table 2.2-7. Synthesis of O-acetylated tyrosine dipeptide derivatives **35c'**–**38c'**.[a]

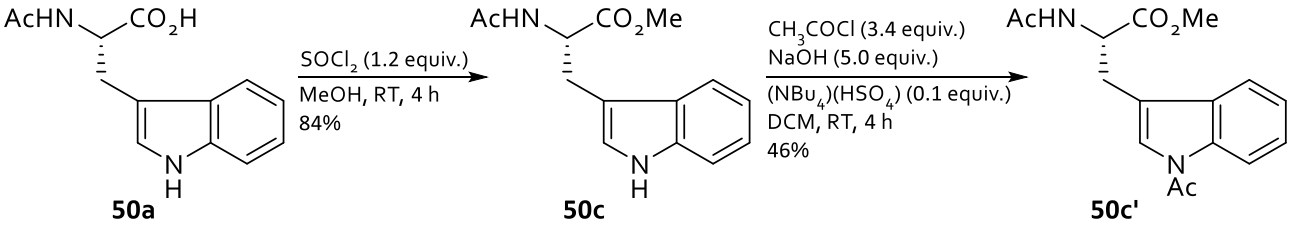
Substrate	Product	Yield
Ac-Phe-Tyr-OMe (**35c**)	Ac-Phe-Tyr(*O*-Ac)-OMe (**35c'**)	47%
Ac-Tyr-Leu-OMe (**36c**)	Ac-Tyr(*O*-Ac)-Leu-OMe (**36c'**)	92%
Ac-Tyr-Phe-OMe (**37c**)	Ac-Tyr(*O*-Ac)-Phe-OMe (**37c'**)	41%
Ac-Tyr-Tyr-OMe (**38c**)	Ac-Tyr(*O*-Ac)-Tyr(*O*-Ac)-OMe (**38c'**)	57%

[a] The reaction proceeds in the same fashion as Scheme 2.2-7. For detailed structures of the substrate, refer to Table 2.2-2. [b] To acetylate both hydroxy groups, 40 equiv. of Ac_2O and 0.02 equiv. of DMAP were used.

The O-acetylated tyrosine derivatives **25c'**, **35c'**–**38c'** were used in the laser flash experiments in Section 4.3.4, page 100. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.2.6 Synthesis of the N- and C-protected tryptophan derivative

Being as reactive as tyrosine, side-chain protection was also necessary for tryptophan residues. Following methylation of the commercially obtained Ac-Trp (**50a**; Sigma-Aldrich) to generate Ac-Trp-OMe (**50c**), acetylation of the side chain was performed using acetyl chloride, under basic conditions in the presence of tetrabutylammonium hydrogen sulfate as the phase-transfer agent, to yield Ac-Trp(*N*-Ac)-OMe (**50c'**) as the final product (Scheme 2.2-8).[41]



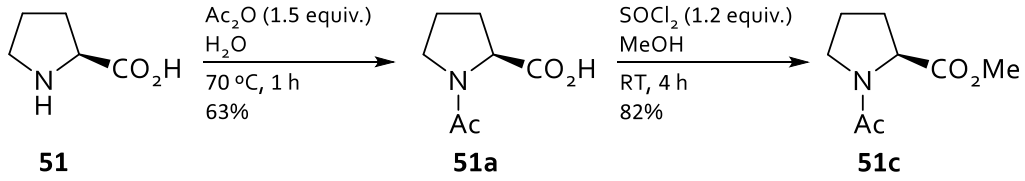
Scheme 2.2-8. Synthesis of Ac-Trp(*N*-Ac)-OMe (**50c**').

The N- and C-protected tryptophan derivative **50c**' was used in the laser flash experiment in Section 4.3.4, page 100. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

2.3 Synthesis of proline, proline-containing short peptides and proline-like molecules

2.3.1 Synthesis of N-acetylated proline and proline-containing peptide methyl esters

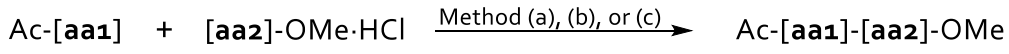
Proline was protected at the N- and C-terminus by an acetyl and methyl ester group, respectively. The synthesis was carried out according to Scheme 2.3-1, starting from acetylation of the commercially available proline (**51**; AK Scientific) at an elevated temperature,[42] followed by C-terminal methylation of *N*-acetyl proline (**51a**).[26]



Scheme 2.3-1. Synthesis of Ac-Pro-OMe (**51c**).

Proline-containing dipeptides were obtained through the coupling of an N-acetylated amino acid and a hydrochloride salt of an amino acid methyl ester using various coupling agents (Table 2.3-1).[43]

Table 2.3-1. Synthesis of proline-containing dipeptide derivatives **52c–56c**.[a]



Substrate 1 (**1a**, **3a**, **24a**, **51a**) | **Substrate 2** (**24b**, **51b**) | **Dipeptide** (**52c–56c**)

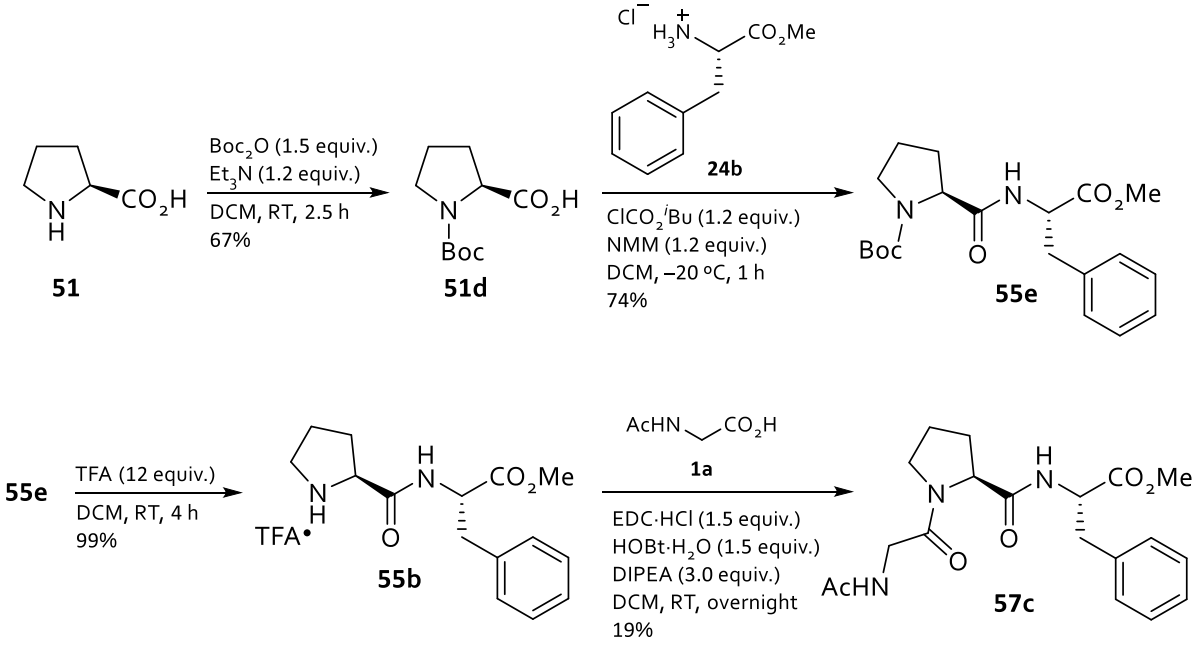
Method (a): EDC·HCl (1.5 equiv.), HOBt·H_2O (1.5 equiv.), DIPEA (3.0 equiv.), DCM, RT, overnight.[43a]
Method (b): Isobutylchloroformate (1.2 equiv.), *N*-methylmorpholine (2.3 equiv.), DCM, –20 °C, 1 h.[43b]
Method (c): HOBt·H_2O (1.0 equiv.), HBTU (1.0 equiv.), DIPEA (2.4 equiv.), DCM, 0 °C, 3 h.[43c]

Method	Substrate 1	Substrate 2	Dipeptide	Yield
(a)	Ac-Gly (**1a**)	Pro-OMe·HCl (**51b**)	Ac-Gly-Pro-OMe (**52c**)	43%
(a)	Ac-Val (**3a**)	Pro-OMe·HCl (**51b**)	Ac-Val-Pro-OMe (**53c**)	65%
(a)	Ac-Phe (**24a**)	Pro-OMe·HCl (**51b**)	Ac-Phe-Pro-OMe (**54c**)	21%
(b)	Ac-Pro (**51a**)	Phe-OMe·HCl (**24b**)	Ac-Pro-Phe-OMe (**55c**)	66%
(c)	Ac-Pro (**51a**)	Pro-OMe·HCl (**51b**)	Ac-Pro-Pro-OMe (**56c**)	19%

[a] **aa1** = amino acid 1; **aa2** = amino acid 2.

N-Acetylated amino acids **1a** and **3a** were obtained commercially (AK Scientific), while **24a** and **51a** were prepared from N-acetylation of Phe **24** and Pro **51** according to Table 2.2-1, page 28 and Scheme 2.3-1, respectively. Phe-OMe·HCl (**24b**) was synthesised according to Table 2.2-1, while Pro-OMe·HCl (**51b**) was obtained commercially (Chem-Impex).

Tripeptide Ac-Gly-Pro-Phe-OMe (**57c**) was synthesised from the C-terminus according to Scheme 2.3-2. Boc-Pro (**51d**) was obtained through the reaction of Pro **51** with Boc anhydride,[44] followed by coupling to Phe-OMe·HCl (**24b**) to produce the intermediate Boc-Pro-Phe-OMe (**55e**). Deprotection of **55e** gave the trifluoroacetate salt of Pro-Phe-OMe (**55b**), which was coupled to Ac-Gly (**1a**) to yield the tripeptide **57c**.

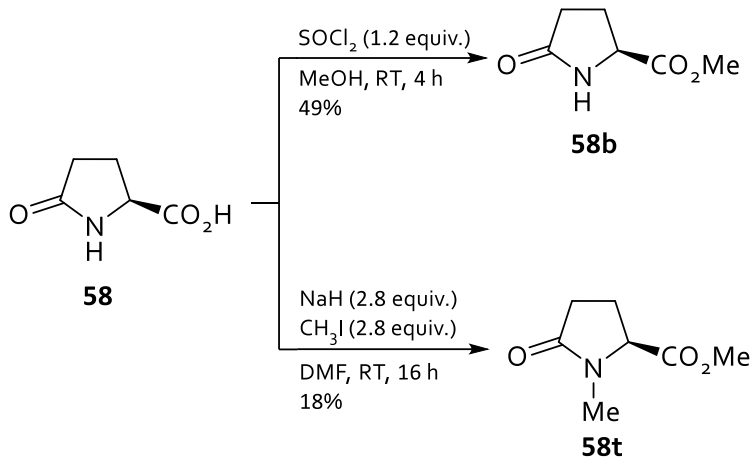


Scheme 2.3-2. Sequential synthesis of tripeptide **57c** from the C-terminus.

Ac-Pro-OMe (**51c**), dipeptides **52c–56c**, and tripeptide **57c** were used in the laser flash experiments in Section 5.3.1, page 111. Dipeptides **54c** and **55c** were also used in the product studies in Section 5.3.2, page 113. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

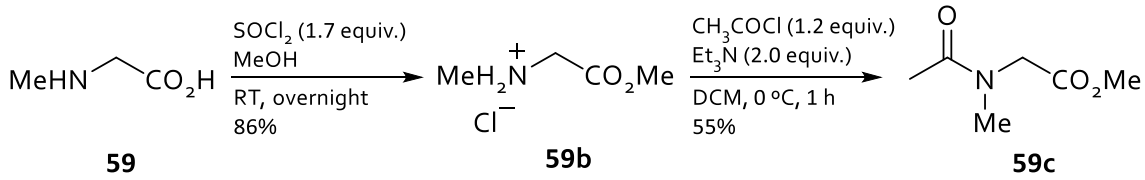
2.3.2 Synthesis of proline-like molecules

Other non-proteinogenic amino acids, pyroglutamic acid (Glp) and sarcosine (*N*-methylglycine) were employed as the model compounds that structurally mimic proline derivatives synthesised in Section 2.3.1. C-Terminal methylation of Glp **58** yielded methyl pyroglutamate, Glp-OMe (**58b**), which was used as a model compound, whereas methyl iodide (CH_3I) and sodium hydride (NaH) achieved total methylation at both the N- and C- termini of Glp **58** to access the other model compound, Me-Glp-OMe (**58t**) according to Scheme 2.3-3.[26,45]



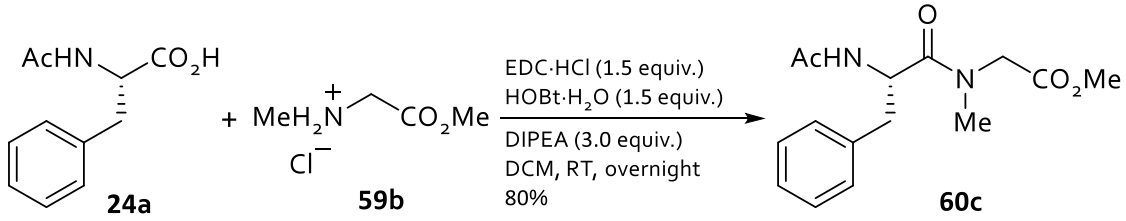
Scheme 2.3-3. Synthesis of Glp-OMe (**58b**) and Me-Glp-OMe (**58t**).

Commercially obtained *N*-methylglycine (**59**; Sigma-Aldrich) was protected at the N-terminus by an acetyl group and at the C-terminus as a methyl ester to yield Ac,Me-Gly-OMe (**59c**) according to Scheme 2.3-4.[25,27]



Scheme 2.3-4. Synthesis of Ac,Me-Gly-OMe (**59c**).

Me-Gly-OMe·HCl (**59b**) was used to obtain dipeptide Ac-Phe-(Me)Gly-OMe (**60c**) through coupling with Ac-Phe (**24a**) using 1-hydroxybenzotriazole monohydrate (HOBt·H_2O) and EDC·HCl as the coupling agents (Scheme 2.3-5).



Scheme 2.3-5. Synthesis of Ac-Phe-(Me)Gly-OMe (**60c**).

These pyroglutamic acid and sarcosine derivatives **58b**, **58t**, **59c** and **60c** were used in the laser flash experiments in Section 5.3.3, page 121. Details of reaction conditions and spectroscopic data are given in the experimental section Chapter 7, page 150.

Chapter Three: Reaction of nitrate radicals with aliphatic amino acids and peptides

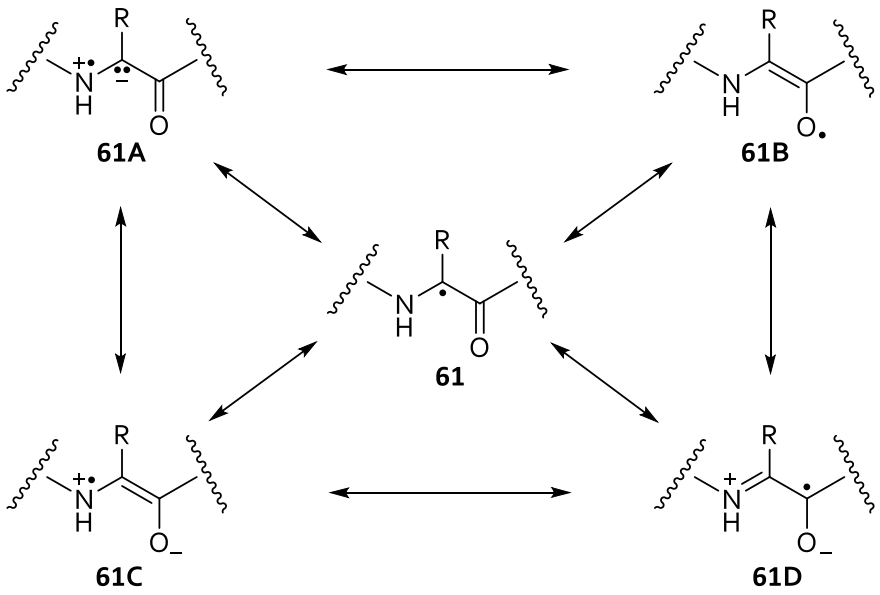
3.1 Introduction

3.1.1 The reactivity of aliphatic amino acid residues

Aliphatic amino acids are usually not considered as the primary target of oxidative damage in biological systems. This is due to the fact that the reactive aromatic and sulfur-containing residues would be attacked more rapidly by the oxidants.[23] However, in a protein where the 'less reactive' sites are exposed to the environment, the reaction of aliphatic residues with oxidants, including many radicals, becomes relevant. This oxidation can occur either at the protein backbone or the side chains of the amino acid. Generally, highly reactive radicals, such as hydroxyl radicals ($HO^{\bullet}$), have low selectivity and react through hydrogen atom transfer (HAT) from the backbone and the side chain as well as through electron transfer [$E°$ ($HO^{\bullet}$, H^+/H_2O) = 2.7 V vs NHE][46d] at the nitrogen atom.[23,46] Hawkins and Davies demonstrated that the regioselectivity of the $HO^{\bullet}$ attack depends on the nature of the amino acid residue.[46b] The attack at the backbone α-carbon only occurs when there is either no side chain (i.e., in glycine residues) or only a methyl group present (i.e., in alanine residues). In contrast to this, with longer side chains, such as in valine and leucine residues, the attack of $HO^{\bullet}$ takes place predominantly at the tertiary carbon of the side chain. Štefanić et al. found in amino acids with a methylated N-terminus that oxidation at the nitrogen atom is the major pathway, although HAT from the α-carbon also takes place to a lesser extent.[46c]

Less reactive radicals exhibit a stronger regioselectivity for HAT reactions.[47] For example, the electrophilic chlorine radicals ($Cl^{\bullet}$) react with aliphatic amino

acids through HAT from the side chains due to deactivating polar effects by the electron-withdrawing carbomethoxy group (–CO_2Me) at the α-position.[47a] The α-hydrogens are less susceptible towards electrophilic attack because their electron density has been reduced. In contrast to this, the reactions of the O-centred hydrogen phosphate radicals ($HPO_4^{\bullet-}$) and cumyloxyl radicals ($PhCMe_2O^{\bullet}$) with aliphatic amino acids proceed through HAT from the α-carbon.[47b,c] A similar reaction pattern was also reported for the reactions of the nucleophilic methyl radicals ($CH_3^{\bullet}$) and aminomethyl radicals ($H_2NCH_2^{\bullet}$) with glycine anions, $H_2NCH_2CO_2^{-}$, which abstract hydrogen from the α-position.[47d] In fact, the high preference for the α-attack is in agreement with various preceding product studies that reported the formation of products through α-carbon-centred radical intermediates **61**.[48] These intermediates are captodatively stabilised radicals due to an extensive delocalisation of the unpaired electron from the α-carbon onto the neighbouring electron-withdrawing (capto) and electron-donating (dative) groups, as shown in Scheme 3.1-1 (resonance structures **61A**–**D**).[48b,49]



Scheme 3.1-1. Resonance contributors of α-carbon-centred radicals **61**.[48b]

Among all α-carbon-centred radicals generated from amino acid and peptide derivatives, α-glycyl radicals (**62**) have been shown to form preferentially.[46b,50] Easton and co-workers reported that α-valyl radicals (**64**) have the lowest radical stabilisation energy (RSE) among radicals **62–64**, owing to their bulky isopropyl side chain (Figure 3.1-1).[50] This finding is counterintuitive, since α-alanyl radicals (**63**) and α-valyl radicals (**64**) are tertiary radicals, hence, they are expected to be more stable than the secondary α-glycyl radicals (**62**).[51] However, this finding was further confirmed through computational studies by Zipse et al., which showed that α-glycyl radicals (**62**) are the most thermochemically stable species compared to α-carbon-centred radicals arising from the other proteinogenic amino acids used in that work (e.g., alanine, tyrosine, cysteine, and proline).[52] In a peptide chain, the formation of an α-carbon-centred radical requires a planar arrangement of the bonds at the α-carbon in order to maximise overlapping between the singly occupied p orbital of the α-carbon and the π orbitals of the amido and carboxy substituents.[50] Therefore, the presence of a side chain in radicals **63** and **64** leads to destabilisation of the radicals due to steric repulsion between the side chain and the amide carbonyl (Figure 3.1-1).

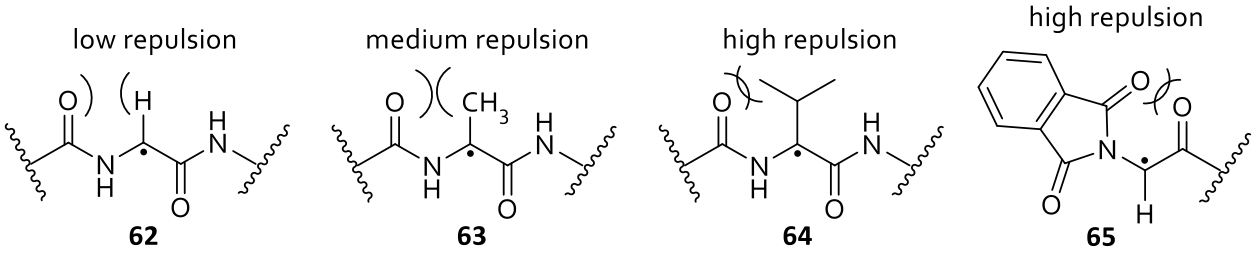


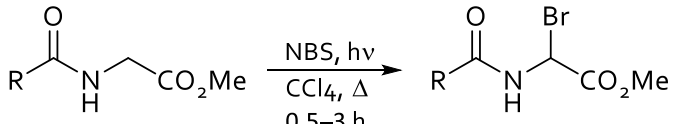
Figure 3.1-1. Repulsions from planar conformations of the α-carbon-centred radicals **61–65**.[50,55a]

Because of the stability of radical **62**, α-glycyl radicals have been well investigated in EPR and radiolysis studies,[46b,53] and have also found application as starting materials for the synthesis of branched α-amino acids,[48a] as well as precursors for α-brominated amino acid and peptide derivatives.[48c,54]

In addition to the electronic effects from the captodative radical stabilisation and the steric effects caused by the amino acid side chains, there is an additional stereo-electronic factor that affects the ease of formation of α-carbon-centred peptide radicals. Easton and co-workers have investigated the effect of different N-substituents on the stability of α-carbon-centred amino acid radicals, using the reactions of amino acid derivatives with *N*-bromosuccinimide (NBS) to determine the most reactive site towards hydrogen abstraction.[55] They found in product studies of amino acids where the N-terminus was protected as a phthalimide, formation of the α-carbon-centred radicals **65** is greatly disfavoured.[55a] In the reaction of *N*-phthaloyl phenylalanine and *N*-phthaloyl leucine, only the β- and γ-positions were brominated, respectively, whereas no reaction at the α-carbon was observed. This outcome was rationalised by the fact that the required planar geometry of the α-carbon-centred radicals **65** could not be achieved due to steric interactions between one of the carbonyl groups in the phthalimido substituent and the adjacent carbonyl group of the amide bond (Figure 3.1-1), rendering their formation energetically unfavourable.

Apart from the *N*-phthaloyl group it was shown through computational studies that other electron-withdrawing groups, such as a benzoyl or sulfonyl group, on the N-terminus also decreases the radical stabilisation energies (RSE) of the α-carbon-centred radicals.[55b] The ease of formation of the radicals was further examined by measuring the relative rates of α-bromination of glycine derivatives **1c**,**i**,**u**–**x** that possess different N-terminal protecting groups (Table 3.1-1).

Table 3.1-1. Relative rates, k_{rel}, of the bromination of glycine derivatives **1c**,**i**,**u**–**x**.[a]



1c, 1i, 1u–1x **66c, 66i, 66u–66x**

Entry	Substrate	R	Product	k_{rel}
1	**1c**	CH_3	**66c**	1.2
2	**1i**	CF_3	**66i**	0.05
3	**1u**	CF_3SO_2	**66u**	< 0.005
4	**1v**	C_6H_5	**66v**	1.0[b]
5	**1w**	*p*-FC_6H_4	**66w**	0.86
6	**1x**	C_6F_5	**66x**	0.25

[a] Selected data taken from Croft et al.[55b] [b] Assigned as unity.

It was found that the rate slowed down with increasing electron-withdrawing N-substituents. For example, the rate decreased when hydrogen atoms on the acetyl and benzoyl group were substituted with the more electronegative fluorine atoms (entry 1 vs 2 and entry 4 vs 5 and 6). *N*-Triflylglycine methyl ester (**1u**) was found to be the most inert substrate towards α-bromination (entry 3), which is due to the strong electron-withdrawing nature of the triflyl group (CF_3SO_2–). The extensive electron delocalisation and the presence of electronegative fluorine atoms lower the electron density on the nitrogen and decrease its electron-donating ability to the radical centre, which consequently hamper the formation of the product **66u**.

Overall, the product and kinetic studies with various radicals reported in the literature so far clearly show that HAT in aliphatic amino acid and peptide derivatives are influenced by a variety of different factors, that range from the reactivity of the attacking radicals to steric and electronic effects in the amino acid and peptide residues. The next section will discuss in particular what has been explored in the reaction of the air pollutant $NO_3^{\bullet}$ with aliphatic amino acids.

3.1.2 Kinetic studies of $NO_3^\bullet$ with aliphatic amino acids

In the only existing kinetic study of the reaction of $NO_3^\bullet$ with amino acids prior to this work, absolute rate coefficients were determined for a series of α- and β-amino acids.[56] Flash photolysis of cerium(IV) ammonium nitrate (CAN) in 6 M nitric acid (HNO_3) generated $NO_3^\bullet$ via a photoinduced electron transfer at $\lambda > 300$ nm according to eq 3.1 below.[56,57]

$$(NH_4)_2Ce(NO_3)_6 + h\nu\ (> 300\ nm) \rightarrow (NH_4)_2Ce(NO_3)_5 + NO_3^\bullet \qquad \text{(eq 3.1)}$$

The second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with a selection of aliphatic amino acids are provided in Table 3.1-2.

Table 3.1-2. Second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with some aliphatic amino acids in 6 M HNO_3 (298 ± 1 K).[a]

Amino acid	***k*** **($M^{-1}\ s^{-1}$)**
Leucine	2.1×10^5
Proline	1.6×10^5
Threonine	1.1×10^5
Serine	5.8×10^4
Valine	5.2×10^4
Hydroxyproline	4.1×10^4
Aspartic acid	3.5×10^3
Alanine	1.7×10^3
Glycine	1.5×10^3

[a] Selected data taken from Venkatachalapathy and Ramamurthy.[56] Experimental error ±10%.

Although the mechanism by which these aliphatic amino acids react with $NO_3^\bullet$ was not explored in detail, HAT was suggested as the preferred reaction pathway based on the magnitudes of the rate coefficients.[56] Interestingly, the range of the

obtained rate coefficients spans two orders of magnitude, which is quite significant given that HAT is the most likely reaction pathway and the similarity of the C–H bonds in these amino acids. It was reasoned that this variation might originate from the number of sites available for HAT, but the reactive site where the abstraction takes place was not further investigated. Therefore, the work presented in this chapter will endeavour to identify the site where the reaction with $NO_3^\bullet$ is most favourable.

3.2 Aim

In order to understand how air pollution, particularly $NO_3^\bullet$, reacts with aliphatic amino acids and peptides, laser flash photolysis studies in conjunction with computational studies, were performed as a tool to elucidate the reaction mechanism. The existing kinetic studies used 6 M HNO_3 as the solvent, which cannot be regarded as mimicking the biological environment, and the question of where exactly the reaction takes place remains open.[56] The work in this chapter will aid to fill the gap in the knowledge of $NO_3^\bullet$-site selectivity towards aliphatic amino acid residues. Acetonitrile (CH_3CN) will be used as a solvent in the laser flash experiments. Despite being not 'physiological conditions' either, this solvent has been shown previously to be highly suitable for studying the reaction of $NO_3^\bullet$ with a large number of substrates without any interfering with the reaction intermediates (for example through hydrogen bonding or hydrolysis).[25e,58] As outlined in Section 1.3, page 10, it is important to gain fundamental understanding of the chemical reactivity in the absence of interfering solvents, salts, ions or other molecules before extending the studies to more biologically relevant.

Computational studies presented in this chapter were performed by Prof. Uta Wille.

Specifically, the aims of this chapter were to:

- ✓ develop an experimental setup to study $NO_3^\bullet$ kinetics in acetonitrile according to literature precedent,
- ✓ gain insight of the reactivity of aliphatic amino acids and peptides towards $NO_3^\bullet$ by determining the absolute rate coefficients in acetonitrile,
- ✓ investigate how steric and electronic factors affect the reactivity of substrates towards $NO_3^\bullet$,
- ✓ calculate activation barriers ($\Delta E^\ddagger$) and reaction energies (ΔE) at various possible sites of $NO_3^\bullet$ attack on aliphatic amino acids to support the findings from the kinetic studies, and
- ✓ evaluate the contribution of the reactive sites in the aliphatic amino acid moieties towards $NO_3^\bullet$ attack based on kinetic and computational studies.

3.3 Results and Discussion

3.3.1 Generation of $NO_3^\bullet$ through irradiation of cerium(IV) ammonium nitrate in acetonitrile

All laser flash experiments were performed at 298 ± 1 K on an Edinburgh Instrument LP920 spectrometer using the third harmonic of a Quantel Brilliant B Nd:YAG laser (6 ns pulse, 10–30 mJ, λ = 355 nm) to generate the reaction transient. The detection system employed a Hamamatsu R2856 photomultiplier tube (PMT) interfaced with a Tektronix TDS 3012C Digital Phosphor Oscilloscope for transient absorption spectra.

CAN (Ajax Finechem, 98.5%) was purified by dissolving it in acetonitrile (Honeywell B&J, HPLC grade) and filtering through neutral alumina (Sigma-Aldrich). The solvent was then removed under reduced pressure and the resulting light orange solid was stored in the freezer, protected from ambient light, until used (can be stored indefinitely under the exclusion of moisture). The absorption spectrum of the transient species formed through irradiation of a 0.33 ± 0.03 mM CAN solution in acetonitrile showed three peaks in the visible region at λ = 595, 630, 670 nm with the maximum absorbance at 630 nm (see inset in Figure 3.3-1). This spectrum agrees well with the literature and demonstrates successful formation of $NO_3^\bullet$ through the reaction in eq 3.1 (page 47).[56,57] Using the extinction coefficient (ε) of 1350 M^{-1} cm^{-1} at 630 nm,[57c] the concentration of $NO_3^\bullet$ in this experiment was determined to be 80 ± 20 μM. All rate data in this work were obtained from the time-dependent decay of the signal at 630 nm. In the absence of O_2 by purging CAN solution in acetonitrile with argon in an ultrasound bath, it was found that $NO_3^\bullet$ was not generated following the irradiation of the solution for reasons currently not understood. Only without degassing the reaction

solution, $NO_3^\bullet$ was successfully generated as shown by its three characteristic peaks in its absorption spectrum (see inset in Figure 3.3-1). Since other studies also confirmed that the presence of O_2 has no effect on the rate coefficients for reactions involving $NO_3^\bullet$,[56] all kinetic experiments in this work were performed in the presence of O_2.

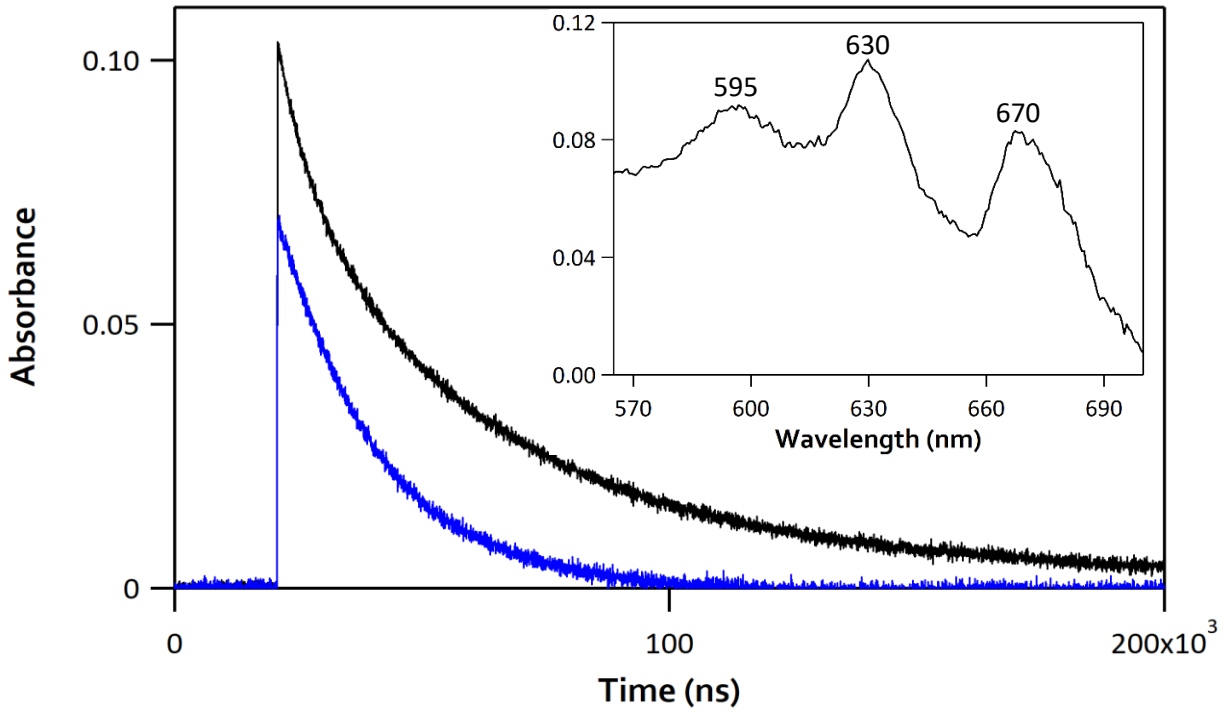


Figure 3.3-1. Time-dependent decay of the $NO_3^\bullet$ signal at 630 nm in acetonitrile ([CAN] = 0.33 mM), in the absence (black line) and presence of 7.7 mM Ac-Gly-OMe (blue line). Inset: $NO_3^\bullet$ absorption spectrum 10 μs after the laser flash.

It was found that the $NO_3^\bullet$ signal decayed over time even in the absence of any substrate (Figure 3.3-1), which was most likely due to a reaction with the solvent acetonitrile through HAT (eq 3.2). Addition of the electrophilic $NO_3^\bullet$ to the electron-poor π system of the nitrile (–CN) group is considered as an unlikely pathway.

$$NO_3^\bullet + CH_3CN \rightarrow HNO_3 + {}^\bullet CH_2CN \qquad \text{(eq 3.2)}$$

Using the concentration of pure acetonitrile of 19.1 M, the second-order rate coefficient, *k*, for this reaction was determined to be 0.9×10^3 M^{-1} s^{-1}, which agrees well with the literature value of 1.0×10^3 M^{-1} s^{-1}.[57b]

Various reaction conditions have been explored in order to optimise the generation of $NO_3^\bullet$ in acetonitrile. The concentration of CAN and the laser power were found to be the important factors for the quality of the $NO_3^\bullet$ decay signal. When low CAN concentrations (e.g., [CAN] = 0.20 mM) were used, increased photobleaching occurred, which was identified by a higher transmission of the reaction solution after a laser pulse, especially when a high laser power (30 mJ) was applied. However, using too high CAN concentrations may result in an increased probability of $NO_3^\bullet$ self-reaction, due to the higher amount of $NO_3^\bullet$ generated in the solution, producing $NO_2^\bullet$ and O_2 (eq 3.3).[59]

$$2NO_3^\bullet \rightarrow 2NO_2^\bullet + O_2 \qquad \text{(eq 3.3)}$$

Likewise, higher laser energies were also found to be quite problematic as they resulted in high concentrations of $NO_3^\bullet$. Thus, by employing [CAN] = 0.33 ± 0.03 mM with laser power of 10–20 mJ, it was found that self-reaction of $NO_3^\bullet$ during the laser flash experiments could be avoided (for more details on the self-reaction of $NO_3^\bullet$, refer to Section 3.3.3.2, page 63).

3.3.2 Kinetic studies of the reaction of $NO_3^\bullet$ with N-acetylated amino acid and peptide derivatives

For these experiments, stock solutions of 0.66 ± 0.06 mM of CAN and of the amino acids and peptides (in the range of 2–99 mM, depending on the substrate) were prepared. The amino acid/peptide solutions were diluted to obtain a series of five reactant solutions with different concentrations. Prior to the laser flash photolysis experiments, 5 mL of the CAN stock solution was mixed in an amber volumetric flask (for ambient light protection) with 5 mL of the substrate solution. The rate coefficients were obtained under pseudo-first-order conditions by measuring the rate of the $NO_3^\bullet$ signal decrease as a function of substrate

concentration, with [substrate]:[CAN] in the range of 2–150. Even at the lowest substrate concentration, the pseudo-first-order behaviour was maintained, because the concentration of CAN solution did not reflect the actual $NO_3^\bullet$ concentration generated through the laser flash. The pseudo-first-order rate coefficients, k_{obs}, were determined from decay profiles at each of five different substrate concentrations. Absolute second-order rate coefficients, k, were obtained from linear regression, plotting k_{obs} vs [substrate]. The slope of the line corresponds to the second-order rate coefficients, k. Typical plots of k_{obs} vs [substrate] are presented in Figure 3.3-2.

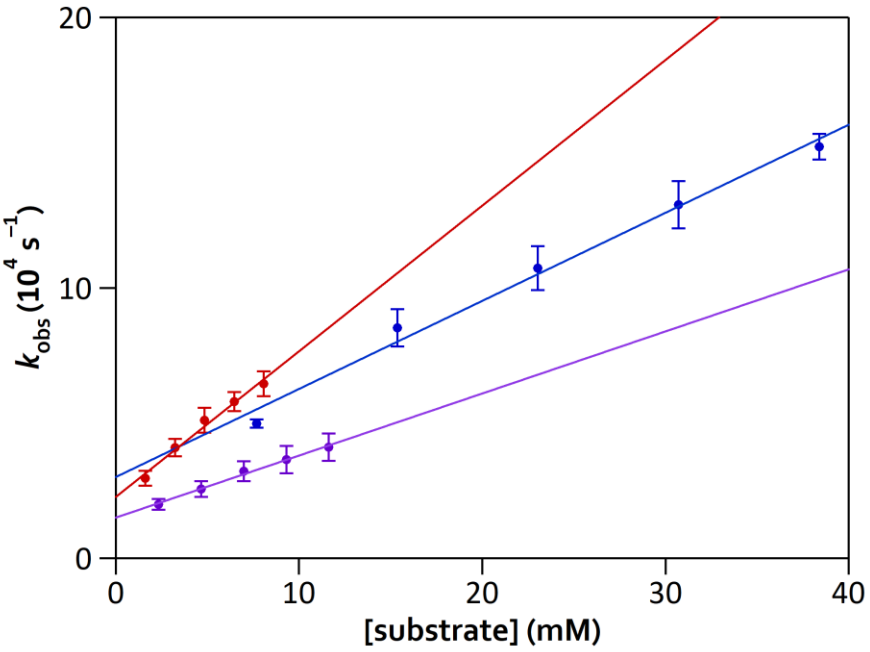


Figure 3.3-2. Plot of the pseudo-first-order rate coefficient (k_{obs}) vs [substrate]. Red: Ac-Val-Val-OMe, blue: Ac-Gly-OMe, purple: Ac-Val-OMe. Error bars shown are 2σ statistical uncertainties.

As shown in the example in Figure 3.3-1 (blue line, page 51) for Ac-Gly-OMe, the decay of $NO_3^\bullet$ signal sped up as a result of the bimolecular reaction of $NO_3^\bullet$ with the aliphatic amino acid. For all reactions studied in this work, k_{obs} increased linearly with the concentration of substrate. The non-zero y-intercept of about 10^4 s^{-1} in most cases (excluding the fast reactions), coincides approximately with the pseudo-first-order rate coefficient for the reaction of $NO_3^\bullet$ with acetonitrile. A

slight variation for the y-intercept was due to unavoidable errors in the kinetic measurements, which affect the linearity and the y-intercept of the slope. Using a solvent concentration ([CH_3CN]) of 19 M, the second-order rate coefficient for the reaction of $NO_3^\bullet$ with acetonitrile in Figure 3.3-2 and typical plots shown in the Appendices, page 226–228, was determined to be in the range of (1.0–2.6) x 10^3 M^{-1} s^{-1}. The difference of 1.6 x 10^3 M^{-1} s^{-1} between the upper and lower limit of the rate coefficient of this background decay is insignificant compared to the rate coefficients obtained from the kinetic plots for the reactions of $NO_3^\bullet$ with amino acids and peptides throughout this work (which are more than three orders of magnitude higher). The absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with aliphatic amino acids are compiled in Table 3.3-1.

Table 3.3-1. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with N- and C-protected aliphatic amino acids in acetonitrile (298 ± 1 K).[a]

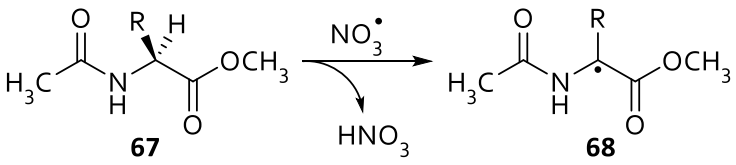
Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Ac-Gly-OMe (**1c**)	3.3
2	Ac-Ala-OMe (**2c**)	1.6
3	Ac-Val-OMe (**3c**)	2.3
4	Ac-Leu-OMe (**4c**)	2.0
5	Ac-Ile-OMe (**5c**)	2.4
6	Ac-Glu(OMe)-OMe (**6c'**)	1.4[b]

[a] Experimental error ±15%. [Substrate]:[CAN] = 6–200. [b]Experimental error ±20%.

The rate coefficients measured in acetonitrile are significantly higher than those in aqueous solutions (compare with Table 3.1-2, page 47).[56,57b] This finding could be attributed to the fact that there is no solvent interference in acetonitrile. Meanwhile, in aqueous solution, such as in 6 M HNO_3, significant hydrogen bonding with the solvent would require $NO_3^\bullet$ to be stripped off its solvation shell for the reaction to take place.[57b]

The side chain of the amino acids contains different types of carbon atoms (i.e., primary, secondary, and tertiary carbons), whose C–H bonds vary in strength. Alfassi and co-workers reported that HAT by $NO_3^\bullet$ in acetonitrile from a tertiary C–H bond proceeds about an order of magnitude faster than that from a secondary C–H bond (per H) and two orders of magnitude faster than that from a primary C–H bond (per H).[57b] The finding that the obtained rate coefficients for the reaction of $NO_3^\bullet$ with the aliphatic amino acids shown in Table 3.3-1 are very similar could suggest that the reaction of $NO_3^\bullet$ may not take place primarily at their side chains, but rather at the 'peptide' backbone.

Along this 'peptide' backbone moiety, hydrogen abstraction can occur from various sites. The hydrogen atom at the α-carbon of amino acid **67** is the most likely target, because a captodatively stabilised α-carbon-centred radical **68** would be formed upon the reaction (Scheme 3.3-1).



Scheme 3.3-1. HAT at the α-carbon in amino acid **67** to produce a stabilised radical **68**.

The higher rate coefficient obtained for the reaction of $NO_3^\bullet$ with Ac-Gly-OMe (**1c**) compared to that with the rest of the aliphatic amino acids (entry 1 vs 2–6, Table 3.3-1) could be accounted to the presence of two α-hydrogens, which means that there is a higher probability for successful collisions, in addition to the particular stability and planarity of the α-glycyl radical.[52] To further confirm this hypothesis, 2-amino isobutyric acid (Aib), with different N- and C-protecting groups, was employed as a model compound. In Aib, the α-position is substituted by two methyl groups, hence, removing any possibility of α-hydrogen abstraction. Thus, if $NO_3^\bullet$-induced HAT would occur exclusively at the α-carbon in aliphatic amino

acids, a significantly lower rate coefficient would be expected for the reaction of $NO_3^\bullet$ with Aib derivatives. Table 3.3-2 compiles the absolute second-order rate coefficients, *k*, determined for the reaction of $NO_3^\bullet$ with Aib derivatives **16c**,**h**,**j–l**.

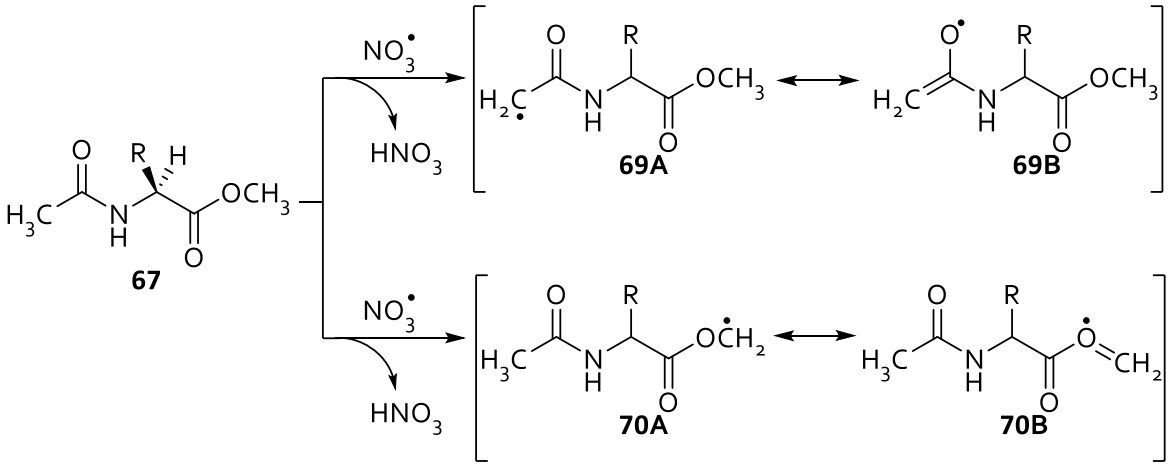
Table 3.3-2. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with the derivatives of 2-aminoisobutyric acid in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Ac-Aib-OMe (**16c**)	1.9
2	Ac-Aib-OCD_3 (**16h**)	1.8
3	Ac(d_2)-Aib-OMe (**16j**)	3.3
4	Ac(d_3)-Aib-OCD_3 (**16k**)	6.0
5	Piv-Aib-OMe (**16l**)	1.9

[a] Experimental error ±15%. [Substrate]:[CAN] = 8–75.

Surprisingly, the rate coefficient for Ac-Aib-OMe (**16c**) of 1.9 x 10^6 M^{-1} s^{-1} was similar to that for the reactions of amino acids with linear aliphatic side chains, such as Ac-Ala-OMe (**2c**) or Ac-Glu(OMe)-OMe (**6c'**; compare Table 3.3-2 entry 1 with Table 3.3-1 entries 2 and 6, page 54). This finding reveals that the reaction of $NO_3^\bullet$ with aliphatic amino acids is far more complex and does not occur only through HAT at the α-carbon of the amino acid derivatives.

The somewhat artificial modification of the amino acids through N- and C-terminus protecting groups to enable their studies in acetonitrile (i.e., acetyl and methyl ester groups) could possibly influence their reactivity towards $NO_3^\bullet$. Hydrogen abstraction at these sites results in radical intermediates **69** and **70** that are stabilised through resonance structure from either the electron-withdrawing substituent (carbonyl at the acetyl group: **69A** and **69B**) or the electron-donating substituent (methoxy at the methyl ester group: **70A** and **70B**), as shown in Scheme 3.3-2.



Scheme 3.3-2. Possible resonance stabilisation from the protecting group of amino acid **67** upon HAT.

Therefore, to examine if this is the case, kinetic isotope effect (KIE) studies were performed by replacing hydrogen atoms at the acetyl and methyl ester group with deuterium atoms. KIE studies have been widely used in the elucidation of reaction mechanisms, especially in HAT.[60] Primary KIE, where the rate of the deuterated substrate is considerably slower, would be observed if the transition state is symmetrical.[60a]

The rate coefficient for the reaction of $NO_3^\bullet$ with Ac-Aib-OCD_3 (**16h**) was determined to be 1.8 x 10^6 M^{-1} s^{-1}, indicating the absence of a primary H/D isotope effect (entry 2 vs 1). Interestingly, deuteration of the *N*-acetyl group exhibits a presumably inverse kinetic isotope effect in both Ac(d_2)-Aib-OMe (**16j**) and Ac(d_3)-Aib-OCD_3 (**16k**; entries 3 and 4 vs 1), which at current cannot be explained. It is likely that there are inseparable trace amounts of highly reactive impurities that dominate the reaction with $NO_3^\bullet$, and thus, increase the rate coefficients for the reactions involving Aib **16j** and **16k**. Another approach to investigate whether HAT from the *N*-acetyl group could be the major pathway is by removing all hydrogens on the *N*-acetyl group, for example by a pivaloyl (Piv) substituent, which lacks any abstractable C–H bonds adjacent to the carbonyl moiety. The rate coefficient obtained for the reaction of $NO_3^\bullet$ with Piv-Aib-OMe (**16l**) was found to be the same

as Ac-Aib-OMe (**16c**; entry 5 vs. 1). Therefore, the absence of a primary H/D isotope effect on the methyl ester group and the same rate coefficient obtained for Piv-Aib-OMe (**16l**) suggest that HAT at the methyl ester and acetyl group can be excluded.

Through the process of elimination of the possible reactive site so far, it appears that the reaction at the amide nitrogen contributes significantly to the reactivity of aliphatic amino acids towards $NO_3^\bullet$. This hypothesis would be further investigated in the latter section (see Section 3.3.3.2, page 63). Deuteration of the amide N–H moiety has been attempted but was unsuccessful (not shown).

In order to explore the rate of oxidative damage in aliphatic di- and tripeptides, kinetic data were obtained for the reaction of $NO_3^\bullet$ with N- and C-protected aliphatic di- and tripeptides **9c–15c**, which are compiled in Table 3.3-3.

Table 3.3-3. Absolute second-order rate coefficients, k, for the reaction of $NO_3^\bullet$ with aliphatic di- and tripeptides in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Ac-Val-Val-OMe (**9c**)	5.4
2	Ac-Val-Leu-OMe (**10c**)	5.7
3	Ac-Leu-Ala-OMe (**11c**)	5.8
4	Ac-Leu-Val-OMe (**12c**)	6.5
5	Ac-Leu-Leu-OMe (**13c**)	6.0
6	Ac-Val-Leu-Val-OMe (**14c**)	9.4
7	Ac-Leu-Leu-Leu-OMe (**15c**)	10.4[b]

[a] Experimental error ±15%. [Substrate]:[CAN] = 2–30. [b] Experimental error ±30%. Limited concentration range ([peptide]/[$NO_3^\bullet$] = 7–17) due to its low solubility in acetonitrile.

The rate coefficients are shown to increase with increasing number of aliphatic amino acid residues in the peptide chain. Dipeptides **9c–13c** were found to be 2–3 more reactive than amino acids **1c–6c**' towards oxidative damage by $NO_3^\bullet$, whereas tripeptides **14c** and **15c** react 3–5 times faster with $NO_3^\bullet$ than amino acids **4a–f**

(compare Table 3.3-3 with Table 3.3-1, page 54). These findings confirm the earlier proposition that reactions at N- or C-terminal protecting group can be excluded. If the reaction would occur only at the acetyl group or the methyl ester group, the rate coefficients for the reactions of amino acids, di- and tripeptides are expected to be the same as the N- and C-protected amino acids.

N-Acetylated aliphatic amino acid amide derivatives were employed as a model system for simplified dipeptides. Table 3.3-4 compiles the absolute second-order rate coefficients for the reaction of $NO_3^\bullet$ with the amides **2q**, **4q**, **1r**, **4r**.

Table 3.3-4. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with aliphatic di- and tripeptides in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Ac-Ala-NH_2 (**2q**)	7.5
2	Ac-Leu-NH_2 (**4q**)[b]	10.3
3	Ac-Gly-NH^{t}Bu (**1r**)	54.6
4	Ac-Leu-NH^{t}Bu (**4r**)	28.1

[a] Experimental error ±10%. [Substrate]:[CAN] = 2–25. [b] Prepared by Julius Kerstien.

For reasons not yet fully understood, aliphatic amino acids with the C-terminal amide were found to be significantly more reactive than the corresponding methyl ester protected amino acids and dipeptides (compare Table 3.3-4 with Table 3.3-1, page 54 and Table 3.3-3, page 58). Thus, it is reasonable to suggest, as proposed earlier, that the reaction at the amide moiety needs to be considered when studying the oxidative damage of amino acids and peptides. This proposed pathway will further be explored later in computational studies in Section 3.3.4, page 69. In the next section, other N-terminal protecting groups would be employed to examine if electronic effects might affect the reactivity of these substrates towards $NO_3^\bullet$.

3.3.3 The effect of different substituents and functional groups on the kinetics of $NO_3^•$ with aliphatic amino acids and peptides

3.3.3.1 Kinetic studies of the reaction of $NO_3^•$ with *N-tert*-butoxycarbonyl (*N*-Boc) amino acid and peptide derivatives

The *tert*-butoxycarbonyl (Boc) group is known to be more electron-donating than the standard acetyl group used throughout this work.[61] In view of the more electron-donating nature of Boc group, introduction of this protecting group at the N-terminus would affect the reaction rates, especially if the reaction takes place at this moiety.

Laser flash experiments were performed under the same conditions as described in Section 3.3.2, page 52. The kinetic data obtained for the reaction of $NO_3^•$ with N-Boc protected amino acid and peptide derivatives are compiled in Table 3.3-5.

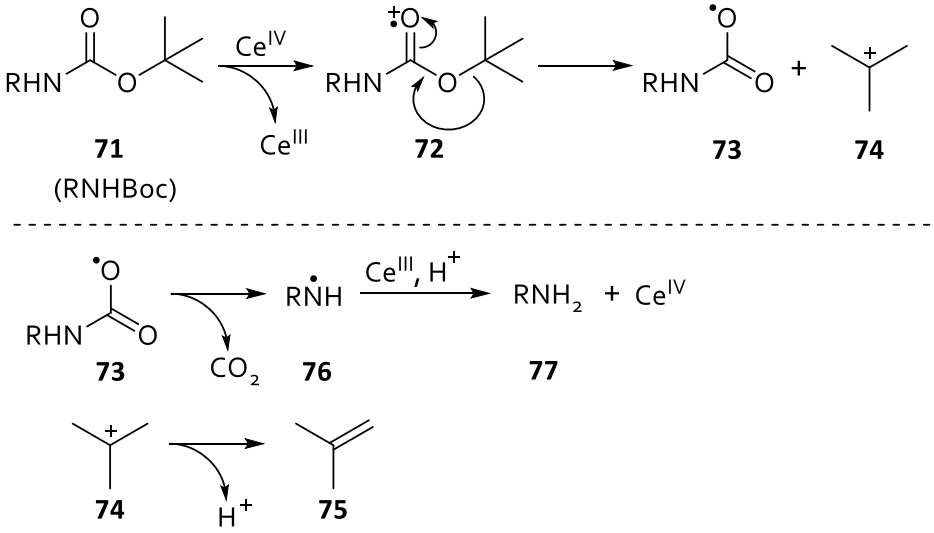
Table 3.3-5. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^•$ with N-Boc protected amino acid, di- and tripeptide methyl esters in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Boc-Gly-OMe (**1e**)	8.7
2	Boc-Ala-OMe (**2e**)	8.4
3	Boc-Aib-OMe (**16e**)	6.9
4	Boc-Leu-OMe (**4e**)	9.2
5	Boc-Ala-Ala-OMe (**17e**)	22.0
6	Boc-Aib-Aib-OMe (**18e**)	23.2
7	Boc-Leu-Leu-OMe (**13e**)	17.6
8	Boc-Leu-Leu-Leu-OMe (**15e**)	22.1

[a] Experimental error ±10%. [Substrate]:[CAN] = 4–50.

There is an increase of the rate coefficient by a factor of about 4 when Boc groups were used at the N-terminus compared with acetyl groups (compare Table 3.3-5 entries 1–4 with Table 3.3-1 entries 1–6, page 54). The higher reactivity of N-Boc protected derivatives highlights the contribution from amide N–H moiety to the overall reactivity of the substrates. Interestingly, the reactions of $NO_3^{\bullet}$ with dipeptide derivatives **13e**, **17e**, **18e** were found to be much faster than would be expected (more than twice as fast as the amino acid derivatives; compare entries 5–7 with entries 1–4), which unfortunately could not be explained at this stage. However, an additional residue on the peptide chain in tripeptide **15e** has no noticeable effect on the reactivity (entry 8 vs 5–7). There is also no significant effect observed in the absence of α-hydrogen (entries 3 and 6 vs 2 and 5), which suggests that the reactive site in such N-Boc protected amino acid and peptide derivatives **1e**, **2e**, **4e**, **13e**, **15e–18e** is located at the N-terminus.

In this regard, it is important to note that Hwu et al. observed deprotection of the Boc group in the presence of a catalytic amount of CAN (0.2 equiv.) in acetonitrile (Scheme 3.3-2).[62]



Scheme 3.3-3. Deprotection of Boc-group catalysed by CAN.[62]

The reaction is believed to proceed through initial oxidation of the carbamate carbonyl group in **71** to generate intermediate **72**, followed by the release of the *t*-butyl cation **74** affording radical **73** (upon deprotonation, **74** will be released as gaseous isobutene **75**). The resulting carbamoyl radical **73** can then undergo decarboxylation to yield the N-radical **76**. The radical **76** is reduced by Ce(III) to afford the amine **77** upon protonation and regenerate Ce(IV). While the reaction might not occur to the full extent (as it requires about 2 days at room temperature),[62] it is possible that partial deprotection of the N-Boc protected amino acid and peptide derivatives during sample preparation for kinetic studies would yield the free amine. Therefore, the obtained rate coefficients from the kinetic studies in Table 3.3-5, page 60 are possibly derived from a mixture of the N-Boc protected substrates and the free amines. It is also likely that during the laser flash experiments, $NO_3^{\bullet}$ as a stronger oxidant than CAN adopts the same reaction pathway as shown in Scheme 3.3-3, where Boc group is deprotected through an initial oxidation at the carbamate carbonyl group.

Considering this possible side reaction, kinetic data for N-Boc protected substrates display an increase of the reactivity of the amino acid and peptide derivatives in a way that does not reflect the actual reactivity of amino acid and peptide residues towards $NO_3^{\bullet}$. Therefore, this Boc-protecting group is considered unsuitable for the purpose of the kinetic studies in this work.

3.3.3.2 Kinetic studies of the reaction of $NO_3^\bullet$ with aliphatic amino acid and peptide derivatives possessing electron-withdrawing substituents at the N-terminus

The kinetic data for the reaction of $NO_3^\bullet$ with aliphatic amino acids shown in Section 3.3.2, page 52–58 suggest that the amide moiety contributes to the reactivity of aliphatic amino acids towards $NO_3^\bullet$. ET at the amide nitrogen and HAT from the N–H bond are the two possible pathways for the reaction involving $NO_3^\bullet$ at this amide moiety. To further explore the mechanism of this reaction, chemical modification was required that changes the reaction at this site, for example by attaching an electron-withdrawing group, which should deactivate the amide nitrogen from ET. Product studies by Easton et al. have shown that N-phthaloylation of the amino acid leads to destabilisation of the α-carbon radical (see page 45).[55a] Therefore, N-phthaloyl protecting group can be used to eliminate two possible reactive sites, which are the amide nitrogen (due to its poor electron density and the absence of N–H bond) and the α-carbon (due to destabilisation of the radical). If the reaction took place at one of these sites, a decrease in the rate coefficients for the reaction of $NO_3^\bullet$ with N-phthaloylated derivatives would be observed. To complement the results, different electron-withdrawing N-protecting groups, for example trifluoroacetyl (CF_3CO–) group, which reduces the electron density at the nitrogen and thus, prevents ET, but allows HAT to still take place, were also employed in the kinetic studies.

Laser flash experiments were performed under the same conditions as described in Section 3.3.2, page 52. The absolute second-order rate coefficients, *k*, obtained for the reaction of $NO_3^\bullet$ with amino acids and peptides with N-terminal electron-withdrawing protecting groups are compiled in Table 3.3-6.

Table 3.3-6. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with amino acids, di- and tripeptides possessing an N-terminal electron-withdrawing protecting group in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^6 M^{-1} s^{-1})
1	Phth-Gly-OMe (**1n**)	0.6
2	Phth-Aib-OMe (**16n**)	0.9
3	Phth-Aib-OCD_3 (**16p**)	0.6
4	Phth-Val-OMe (**3n**)	0.8
5	Phth-Leu-OMe (**4n**)	1.2
6	Phth-Gly-Gly-OMe (**21n**)	1.5
7	Phth-Gly-Val-OMe (**22n**)	5.0
8	Phth-Gly-Leu-OMe (**23n**)	3.9
9	(*o*-CO_2Me)Bz-Aib-OMe (**16o**)	2.3
10	Tfa-Aib-OMe (**16i**)[b]	0.8

[a] Experimental error ±15%. [Substrate]:[CAN] = 4–75. [b] Tfa = trifluoroacetyl.

The reactivity of the N-phthaloylated derivatives **1n**, **3n**, **4n**, **16n**, **16p** towards $NO_3^\bullet$ was found to be lower than that of the N-acetylated derivatives **1c**–**6c**' (compare Table 3.3-6 entries 1–5 with Table 3.3-1 entries 1–6, page 54). Due to this low reactivity, $NO_3^\bullet$ can undergo self-reaction (see eq 3.3, page 52) which slightly accelerates its decay as seen in Figure 3.3-3 (black line shows the decay of $NO_3^\bullet$).[59] The decay of $NO_3^\bullet$ signal appears to follow a bi-exponential decay more than a first-order exponential decay (compare right image with left image in Figure 3.3-3) especially at the start of the reaction, where $NO_3^\bullet$ exists at a relatively high concentration (red line shows the fitting using the exponential fit for the left image and the bi-exponential fit for the right image).

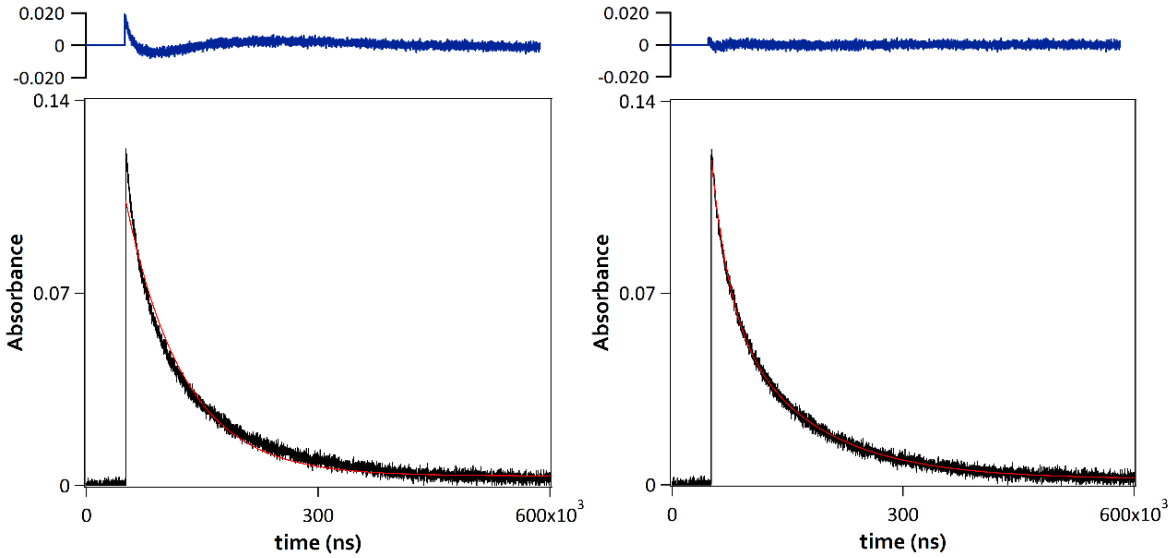


Figure 3.3-3. Time-dependent decay of the $NO_3^\bullet$ signal at 630 nm in acetonitrile ([CAN] = 0.33 mM) in the presence of 3.2 mM Phth-Gly-OMe (**1n**). In red: the exponential decay fitting (left image) and the bi-exponential decay fitting (right image). In blue: the residuals from the exponential fitting (left image) and the residuals from the bi-exponential fitting (right image).

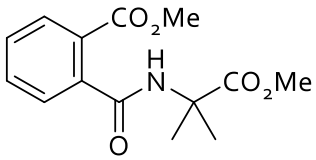
The quality of the fit can be seen from the residuals displayed on top of each decay curve. The residuals from the exponential fit show a trend (left image in blue), which indicates that the exponential fit is not a good fit for the data, whereas, the residuals from the bi-exponential fit are random (right image in blue) and fits the raw data better. Evaluation of rate coefficients at various Phth-Gly-OMe (**1n**) concentrations reveals that the absolute second-order rate coefficients for the reaction of $NO_3^\bullet$ with Phth-Gly-OMe (**1n**) obtained using either decay fitting do not differ much (k = 0.6 x 10^6 M^{-1} s^{-1} using the exponential decay fitting vs k = 0.5 x 10^6 M^{-1} s^{-1} using the bi-exponential decay fitting). However, this finding shows that the rate of the reaction is at the lower limit where the values need to be considered carefully because of the competing self-reactions of $NO_3^\bullet$.

The drop of the rate coefficient by about a factor of 5 in the reaction of $NO_3^\bullet$ with Phth-Gly-OMe (**1n**) compared to that with Ac-Gly-OMe (**1c**) clearly shows that the low reactivity is due to the phthalimide moiety (compare Table 3.3-6 entry 1 with Table 3.3-1 entry 1, page 54). An increase by about 50% was found in the reaction of $NO_3^\bullet$ with Phth-Aib-OMe (**16n**), highlighting the contribution of the

six primary β-hydrogen atoms to the reactivity of amino acid **16n** towards $NO_3^\bullet$ (entry 2 vs 1). Moreover, KIE studies for the reaction of $NO_3^\bullet$ with Phth-Aib-OMe (**16n**) and Phth-Aib-OCD_3 (**16p**) reveal a value of $k_H/k_D = 1.5$ (entry 2 vs 3). This low KIE indicates that HAT from the ester moiety is only a minor contributor to the reactivity in addition to HAT from the β-position. Through isolation of the side chain, such as that in Phth-Val-OMe (**3n**) and Phth-Leu-OMe (**4n**), the rate coefficients were found to be (0.8–1.2) x 10^6 M^{-1} s^{-1}, suggesting that there is also some minor contribution to the reactivity from the side chain (entries 4 and 5).

The rate coefficients determined for the reaction of $NO_3^\bullet$ with Phth-Gly-Gly-OMe (**21n**), Phth-Gly-Val-OMe (**22n**), and Phth-Gly-Leu-OMe (**23n**) need to be interpreted carefully (entries 6–8). While a relatively low rate coefficient was obtained for Phth-Gly-Gly-OMe (**21n**) since it is only the C-terminal glycine that reacts with $NO_3^\bullet$, the rate was too fast for dipeptides **22n** and **23n** (entry 6 vs 7 and 8). It is possible that highly reactive impurities are present in these dipeptides **22n** and **23n** (which could originate from the peptide coupling), which then dominate the reaction with $NO_3^\bullet$. Therefore, these rate data need to be regarded with care. Nevertheless, N-phthaloylated substrates react much slower than the N-acetylated counterparts (compare Table 3.3-6 with Table 3.3-1, page 54 and Table 3.3-3, page 58).

When a serendipitously prepared (*o*-CO_2Me)Bz-Aib-OMe (**16o**) was employed in the kinetic studies, the rate coefficient was found to increase by a factor of more than 2 compared to the N-Phth protected Aib **16n** (entry 9 vs 2). Amino acid **16o** possesses a similar functional group with Phth-Aib-OMe (**16n**), however, with the nitrogen atom being protected as an amide instead of an imide (Figure 3.3-4).



16o

Figure 3.3-4. The structure of (*o*-CO_2Me)Bz-Aib-OMe (**16o**).

The rate of 2.3 x 10^6 M^{-1} s^{-1} was found to be in the range of the N-acetylated amino acids (see Table 3.3-1 entries 1–5, page 54), which further supports the argument that there is a significant contribution from the amide nitrogen to the reactivity of aliphatic amino acids towards $NO_3^{\bullet}$.

More evidence of amide contribution to the reactivity can be seen in the reaction of $NO_3^{\bullet}$ with Tfa-Aib-OMe (**16i**). The strong electron-withdrawing properties of a trifluoroacetyl group would greatly reduce the electron density at the nitrogen, making the moiety less susceptible to the electrophilic $NO_3^{\bullet}$. Indeed, the rate coefficient for the reaction of $NO_3^{\bullet}$ with Tfa-Aib-OMe (**16i**) significantly dropped to 0.8 x 10^6 M^{-1} s^{-1}, which is in the range of the N-Phth protected amino acids (entry 10 vs 2). Even though the amide N–H moiety still exists in amino acid **16i**, the decreased rate to the same value as those with no amide N–H moiety clearly confirms that the reaction of $NO_3^{\bullet}$ with aliphatic amino acids takes place at the amide moiety to a significant extent.

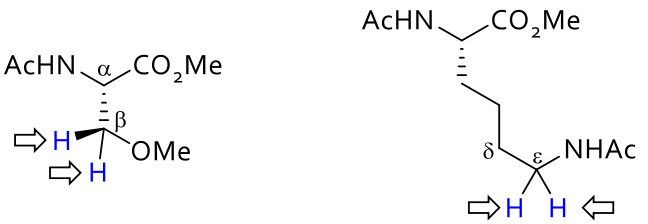
3.3.3.3 Kinetic studies of the reaction of $NO_3^{\bullet}$ with aliphatic amino acids possessing different functional groups on the side chain

The presence of a functional group in the aliphatic amino acid residue can increase its susceptibility towards oxidative damage. Previous studies revealed that over 75% of aliphatic residues are damaged by $HO^{\bullet}$ in the γ-radiolysis of oxygenated solutions of heme α-globin, soluble collagen, and thymus histone.[53,63] The most affected aliphatic residues in those biological molecules are those with

various functional groups, including serine, threonine, lysine and arginine, which prompted the study of their reactivity towards $NO_3^\bullet$. How much faster would they react with $NO_3^\bullet$ compared to the other aliphatic amino acids studied in the previous sections?

To investigate the reactivity of different functional groups towards $NO_3^\bullet$, serine and lysine were employed as model compounds for an oxygen- and nitrogen-containing side chain in an aliphatic residue. The hydroxyl group in the serine side chain was protected as a methyl ether and the amino group in the lysine side chain was protected as an acetamide.

The rate coefficients for the reaction of $NO_3^\bullet$ with Ac-Ser(OMe)-OMe (**7c'**) and Ac-Lys(NHAc)-OMe (**8c'**) were determined to be $(4.6 \pm 0.5) \times 10^6$ M^{-1} s^{-1} and $(9.9 \pm 0.1) \times 10^7$ M^{-1} s^{-1}, respectively. The higher rate coefficients obtained for these amino acids, compared with those for the other aliphatic amino acids, suggest that the reactions mainly occur at the side chain, presumably through HAT from methylene groups adjacent to the functional groups (Figure 3.3-5).



Ac-Ser(OMe)-OMe (**7c'**) Ac-Lys(NHAc)-OMe (**8c'**)

Figure 3.3-5. Primary targets for HAT in **7c'** and **8c'** by $NO_3^\bullet$, indicated by the arrows.

A slight increase of reactivity for Ac-Ser(OMe)-OMe (**7c'**) was due to the activation of β-hydrogens. The presence of the adjacent methoxy group provided additional resonance stabilisation to the radical intermediate formed following HAT at the β-carbon. Surprisingly, Ac-Lys(NHAc)-OMe (**8c'**) exhibited a much higher reactivity by a factor of about 500 towards $NO_3^\bullet$, compared to the other

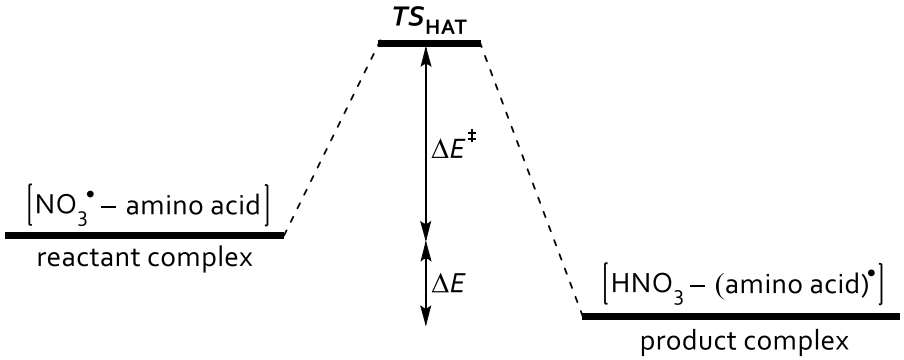
aliphatic amino acids (see Table 3.3-1 entries 1–5, page 54). Similar to **7c**', the hydrogen atoms at the ε-position of amino acid **8c**' are activated. However, the lower electronegativity of nitrogen compared with oxygen makes ε-hydrogens more electron-rich than ε-hydrogens in amino acid **7c**', and thus, more susceptible towards electrophilic attack by $NO_3^{\bullet}$. Another possible pathway is ET at the amide nitrogen of the side chain, which is more electron-rich in comparison to the backbone nitrogen. The backbone nitrogen is in the proximity of the electron-withdrawing carbomethoxy group ($-CO_2Me$).

3.3.4 Computational studies of the reaction of $NO_3^{\bullet}$ with various aliphatic amino acid derivatives

In order to consolidate all the hypotheses and results from the kinetic studies, Prof. Uta Wille performed computational studies to gain a further mechanistic insight for the reaction of $NO_3^{\bullet}$ with aliphatic amino acids. The computations were carried out with the Gaussian 09 program[64] using the M06-2X method, which has been employed previously to investigate α-hydrogen abstraction from amino acids by $HO^{\bullet}$ and $HS^{\bullet}$.[65] All calculations were performed in acetonitrile using the Conductor-like Polarizable Continuum Model (CPCM).[66] The M06-2X method was used in combination with the less expensive 6-31+G* basis set to investigate the reaction mechanism, because of its broad applicability and its moderate average mean absolute errors.[66,67] All ground and transition structures were verified by vibrational frequencies analysis at the same level of theory, and all identified transition structures had only one imaginary frequency. The spin expectation value, $< s^2 >$, was very close to 0.75 after spin annihilation. The Gaussian archive entries for all optimised ground and transition structures, including the imaginary frequencies of the transition structures are given in the Appendices, page 229–253.

The potential energy surface for HAT by $NO_3^\bullet$ in acetonitrile was calculated for all possible sites in a range of N- and C-protected aliphatic amino acids to identify the most likely reaction pathways. The activation barriers ($\Delta E^\ddagger$) and energies of the product association complex [HNO_3 – (amino acid)$^\bullet$] (ΔE) for various reaction pathways, which are given relative to the energy of the respective reactant association complex [$NO_3^\bullet$ – amino acid], are compiled in Table 3.3-7. Formation of association complexes between reactants prior to the actual reaction is common in radical reactions.[68]

Table 3.3-7. Calculated activation barriers ($\Delta E^\ddagger$) and reaction energies (ΔE) for $NO_3^\bullet$-induced HAT at all possible sites in N- and C-protected aliphatic amino acids in acetonitrile (M062X/6-31+G* level of theory, free energies in kJ mol^{-1}).[a]



Amino acid	Entry	Reaction site[b]	$\Delta E^\ddagger$	ΔE
Ac-Gly-OMe	**1**	C***H***$_3$C(O)NH-CH$_2$-CO$_2$CH$_3$	41.7	–75.0
	2	AcN***H***-CH$_2$-CO$_2$CH$_3$	–5.9[c]	–20.3
	3	AcNH-C***H***$_2$-CO$_2$CH$_3$	13.9	–126.0
	4	AcNH-CH$_2$-CO$_2$C***H***$_3$	31.7	–29.7
Phth-Gly-OMe	**5**	(CHC(O))$_2$N-C***H***$_2$-CO$_2$CH$_3$[d]	24.5	–75.8
Ac-Ala-OMe	**6**	AcNH-C***H***(CH$_3$)-CO$_2$CH$_3$	13.2	–105.5
	7	AcNH-CH(C***H***$_3$)-CO$_2$CH$_3$	34.9	–34.7
Ac-Aib-OMe	**8**	AcN***H***-C(CH$_3$)$_2$-CO$_2$CH$_3$	–18.3[c]	–26.7
	9	AcNH-C(C***H***$_3$)$_2$-CO$_2$CH$_3$	26.9	–42.4
Phth-Aib-OMe	**10**	(CHC(O))$_2$N-C***(***C***H***$_3$)$_2$-CO$_2$CH$_3$[d]	32.9	–40.5
Tfa-Aib-OMe	**11**	Ac(F$_3$)N***H***-C(CH$_3$)$_2$-CO$_2$CH$_3$	37.8	–9.5

Ac-Val-OMe	**12**	AcNH-C***H***(CH(CH_3)$_2$)-CO_2CH_3	16.3	–122.1
	13	AcNH-CH(C***H***(CH_3)$_2$)-CO_2CH_3	11.5	–105.9
	14	AcNH-CH(CH(C***H***$_3$)$_2$)-CO_2CH_3	25.8	–34.6
Ac-Leu-OMe	**15**	AcNH-CH(CH_2C***H***(CH_3)$_2$)-CO_2CH_3	9.1	–110.5
Ac-Ser(OMe)-OMe	**16**	AcNH-CH(C***H***$_2$$OCH_3$)-$CO_2CH_3$	7.2	–73.3
Ac-Lys(NHAc)-OMe	**17**	$CH_3C(O)N$***H***-CH_2CH_3[e]	–40.4 [c]	–39.3
	18	$CH_3C(O)NH$-C***H***$_2$$CH_3$[e]	–20.1 [c]	–122.9

[a] Given energies relative to the respective reactant association complex [$NO_3^\bullet$ – amino acid]. [b] Abstracted hydrogens are indicated in bold italics. [c] Reaction proceeds through an intermediate charge-transfer (CT) complex, see text. [d] Maleimide was used as a simplified model, see text. [e] *N*-Ethylacetamide was used as a simplified model, see text.

In general, HAT by $NO_3^\bullet$ from various sites of the amino acid is found to be thermodynamically favourable and associated with low to moderate activation barriers, highlighting the high reactivity of $NO_3^\bullet$ in the laser flash experiments. Quantum mechanical tunnelling effects were not taken into account, and thus, the calculated activation barriers ($\Delta E^\ddagger$) were not corrected accordingly because the computations were aimed to identify the most favourable reaction pathways and not to simulate rate coefficients.

From the computational studies, it is clear that α-hydrogen abstraction is the most exothermic pathway with a low $\Delta E^\ddagger$ of about 15 kJ mol^{-1} (entries 3, 6, 12), and thus, kinetically and thermodynamically more favourable than HAT from most of other C–H moieties. On the other hand, HAT from the *N*-acetyl group has the highest $\Delta E^\ddagger$ of 42 kJ mol^{-1} (entry 1), indicating that this pathway is not competitive, in agreement with the kinetic studies in Section 3.3.2, page 57–58. The relatively high barrier can be rationalised by polar effects from the electron-withdrawing carbonyl group, which deactivate C–H bonds from the attack by the electrophilic $NO_3^\bullet$. HAT from the methyl groups at the ester terminus or the β-position,

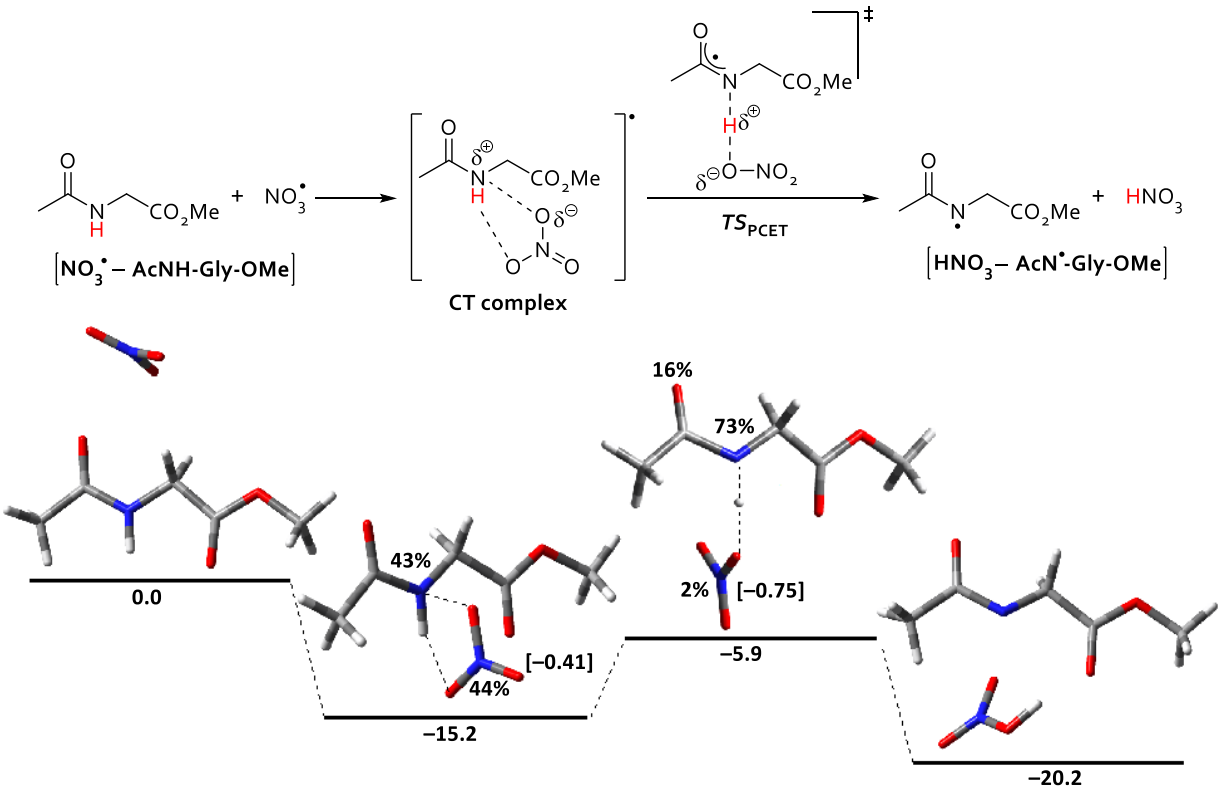
respectively, is kinetically less favourable by about 16 kJ mol^{-1} compared to that from the α-position (entries 4, 7, 9 vs 3, 6, 12).

When the *N*-phthaloyl group was employed in place of the *N*-acetyl group, $\Delta E^{\ddagger}$ for α-HAT was found to increase by about 11 kJ mol^{-1} (entry 3 vs 5), confirming the drop of the reaction rate found in the kinetic studies in Section 3.3.3.2, page 63–67 (computations were performed using maleimide as simplified model system for phthalimide).[55b] Moreover, it was found that the impact of the *N*-phthaloyl group on $\Delta E^{\ddagger}$ for HAT from the β-position seems to be less significant (only about 6 kJ mol^{-1}; entry 9 vs 10).[55a]

HAT from a tertiary C–H moiety in the side chain was found to be kinetically and thermodynamically similar to α-HAT (entries 13, 15 vs 3, 6, 12). In contrast to this, HAT from a primary C–H moiety in the side chain has a considerably higher $\Delta E^{\ddagger}$ by about 14 kJ mol^{-1} (entry 14 vs 13), showing the significant difference in the bond strength of a primary and tertiary C–H bond. This result agrees with previous kinetic studies reported in the literature, where the rate of HAT by $NO_3^{\bullet}$ from a tertiary C–H bond was found to be two orders of magnitude faster than that from a primary C–H bond in acetonitrile.[57b]

From previously performed kinetic studies, a process of elimination of reactive moieties showed that the N–H group plays an important role in the reaction of $NO_3^{\bullet}$ with aliphatic amino acids. This proposition is confirmed by the computational studies for the reaction pathway at the N–H moiety. The transition state for this HAT was found to lie energetically below the reactant association complex [$NO_3^{\bullet}$ – amino acid] (entries 2 and 8). It was revealed that after closer inspection, [$NO_3^{\bullet}$ – amino acid] is connected with the transition state through barrier-less formation of a charge-transfer (CT) complex, which is characterised by a coordinating interaction between $NO_3^{\bullet}$ and the amide N–H moiety (shown in

Scheme 3.3-4 for the reaction of $NO_3^\bullet$ with Ac-Gly-OMe). The change of the Mulliken charges in the CT complex relative to [$NO_3^\bullet$ – amino acid] shows significant development of negative charge on the NO_3 fragment (–0.41e), indicating oxidation of the amide by $NO_3^\bullet$, which is further supported by the approximately equal spin distribution between the amide and NO_3 moieties.



Scheme 3.3-4. Potential energy diagram for the reaction of $NO_3^\bullet$ with the amide group in Ac-Gly-OMe. M062X/6-31+G* optimised geometries and free energies (in kJ mol^{-1} in acetonitrile) relative to [$NO_3^\bullet$ – Ac-Gly-OMe]. Percent numbers are spin densities. Numbers in square brackets are changes of the Mulliken charge on the NO_3 fragment relative to [$NO_3^\bullet$ – Ac-Gly-OMe].

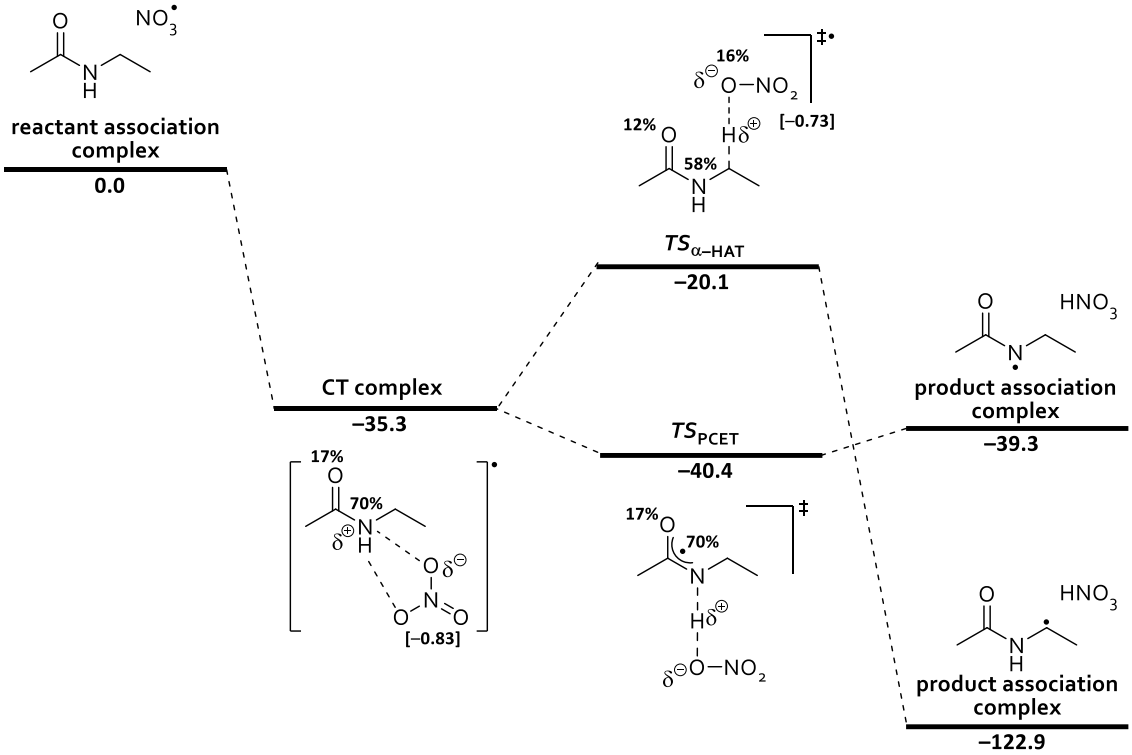
Following the formation of this CT complex, the reaction proceeds to the products through a transition state (***TS***$_{PCET}$), which shows that only about 2% of spin density is located on the NO_3 fragment. There is significant accumulation of negative charge (–0.75e relative to [$NO_3^\bullet$ – Ac-Gly-OMe]), indicating that a proton transfer takes place during transition state from the oxidised amide to nitrate (NO_3^-). Therefore, the reaction of $NO_3^\bullet$ with the amide N–H group appears to proceed through a proton-coupled ET (PCET) involving formation of a CT

complex, which is immediately followed by a proton transfer via $\boldsymbol{TS}_{\mathbf{PCET}}$.[69] This type of mechanism has previously been reported to explain the selective homolysis of strong E–H bonds (such as N–H bonds) in preference to weaker C–H bonds.[69b] The reaction requires a Lewis base which can coordinate to the proton and a one-electron oxidant to accept the electron. In the reaction of $NO_3^\bullet$ with aliphatic amino acids, $NO_3^\bullet$ possesses both the Lewis base and one-electron oxidant properties, making it a unique oxidising/HAT agent in PCET.

To confirm this proposed mechanism, further computations were performed for the reaction of $NO_3^\bullet$ with Tfa-Aib-OMe and revealed that formation of the CT complex is no longer barrier-less, but rather 7.5 kJ mol^{-1} endothermic (not shown). Moreover, $\Delta E^\ddagger$ for this reaction was found to be higher by about 56 kJ mol^{-1}, compared with Ac-Aib-OMe (entry 11 vs 8), clearly confirming the PCET mechanism at the N–H moiety, where the reaction is much less favourable when electron density at that moiety is lowered. This finding is important because the formation of N-centred radicals through HAT from the peptide bond has not received much attention. In the reactions of hypochlorous aicd (HOCl) with peptides, N-centred radicals are obtained as secondary products from rapid decomposition of the intermediate chloroamides.[70] Only recently, computational studies showed that amidyl radicals can be formed through hydrogen bonding and/or partial deprotonation by electrophilic radicals (such as $HO^\bullet$), which tautomerise to the α-carbon radicals, leading to peptide bond cleavage.[71]

When the side chain consists of different functional groups, the obtained rate coefficients were observed to increase (Setion 3.3.3.3, page 67). This higher rate coefficient for Ac-Ser(OMe)-OMe can be rationalised by lone pair donation from the ether oxygen in the side chain that lowers $\Delta E^\ddagger$ for HAT from the adjacent methylene C–H bonds (entry 16). However, Ac-Lys(NHAc)-OMe shows much

higher reactivity with the rate coefficient of 9.9 x 10^7 M^{-1} s^{-1}. While the similar lone pair donation can take place in this side chain from the amide nitrogen, it does not explain why the rate escalates dramatically. Computations using *N*-ethylacetamide as a simplified model system explored two possible pathways (Scheme 3.3-5).



Scheme 3.3-5. Potential energy diagram for the reaction of $NO_3^\bullet$ with *N*-ethylacetamide as the model system for Ac-Lys(NHAc)-OMe. M062X/6-31+G* optimised geometries and free energies (in kJ mol^{-1} in acetonitrile) relative to the reactant association complex. Percent numbers are spin densities. Numbers in square brackets are changes of the Mulliken charge on the NO_3 fragment relative to the reactant association complex.

These two possible pathways proceed through the barrier-less formation of a CT complex where there is no spin density on the NO_3 fragment, but accumulation of the negative charge (–0.83e relative to the reactant association complex). The spin density lies predominantly at the amide nitrogen with some delocalisation to the neighbouring groups. The CT complex can then undergo either an α–HAT with a relatively low $\Delta E^\ddagger$ of about 15 kJ mol^{-1} with respect to the complex, or a barrier-less proton transfer (most likely diffusion-controlled) which then produces the N-centred radical in a similar fashion as the previously discussed

PCET. The α–HAT involves HAT at the methylene position in *N*-ethylacetamide, which is analogous to the ε-position found in Ac-Lys(NHAc)-OMe. Meanwhile, the transition state for PCET was found to lie below the CT complex, which may be due to the level of theory used in this work. Calculations using methods that give more precise energies might find the transition state that lies above the complex. It is also possible that the transition state for PCET (***TS***$_{\mathbf{PCET}}$) has a starting different point that has yet to be identified. Future calculations using theoretical methods that provide more precise energies will be required to gain a more fundamental understanding on these pathways. However, the relatively low energy barriers of these two pathways may contribute to the high reactivity of Ac-Lys(NHAc)-OMe towards $NO_3^{\bullet}$.

3.3.5 Determining the contribution from various reaction sites to the reactivity of aliphatic amino acids towards $NO_3^{\bullet}$

Through a series of kinetic experiments performed in Section 3.3.2 and 3.3.3 and computational studies in Section 3.3.4, contribution from various reaction sites can be evaluated. The rate coefficient for PCET at the amide N–H bond is estimated based on the values obtained for Aib derivatives, where α-HAT is excluded. PCET would likely happen in the reaction of $NO_3^{\bullet}$ with Ac-Aib-OMe (**16c**) but not with Tfa-Aib-OMe (**16i**) or Phth-Aib-OMe (**16n**) due to either the lack of the N–H moiety or the poor electron density at the nitrogen atom. Therefore, the rate coefficient for PCET can be approximated through the difference between their rate coefficients (k = 1.9 x 10^6 M^{-1} s^{-1} for **16c**, k = 0.8 x 10^6 M^{-1} s^{-1} for **16i** and k = 0.9 x 10^6 M^{-1} s^{-1} for **16n**), which leads to the value of 1.0 x 10^6 M^{-1} s^{-1}. The rate coefficient found for Aib **16i** and **16n** can be attributed to the slow HAT from the methyl substituents in both the side chain and the C-terminal methyl ester. The value of 0.6 x 10^6 M^{-1} s^{-1} was obtained for C-terminal methyl ester HAT according

to the rate coefficient determined for Phth-Gly-OMe (**1n**). Thus, HAT from the β-carbons in the side chain of Aib **16i** and **16n** can be determined to be 0.3 x 10^6 M^{-1} s^{-1} (for all six hydrogens). Meanwhile, the rate of HAT from the N-terminal acetyl group is negligible (< 10^5 M^{-1} s^{-1}), considering replacing its all three hydrogens with three methyl groups shows no rate difference (k = 1.9 x 10^6 M^{-1} s^{-1} for both Ac-Aib-OMe, **16c** and Piv-Aib-OMe, **16l**).

The rate coefficients for the reactions of $NO_3^\bullet$ with aliphatic amino acids containing unbranched side chains are quite similar (about 1.5 x 10^6 M^{-1} s^{-1} for both Ac-Ala-OMe, **2c** and Ac-Glu(OMe)-OMe, **6c'**), suggesting that there are two pathways involved in the reactions, which are PCET at the amide moiety and HAT from the α-carbon. With amino acids containing branched side chains, such as those in Ac-Val-OMe (**3c**), Ac-Leu-OMe (**4c**), and Ac-Ile-OMe (**5c**), the rate coefficients are slightly higher (about 2.2 x 10^6 M^{-1} s^{-1}), indicating involvement of the tertiary C–H bond. This contribution from the tertiary bond was explored in isolation through kinetic studies of the reaction of $NO_3^\bullet$ with Phth-Val-OMe (**3n**) and Phth-Leu-OMe (**4n**). In this system, PCET at the N-terminal amide and HAT from the α-carbon are both suppressed, which results in a drop of the rate coefficient by about half to (0.8–1.2) x 10^6 M^{-1} s^{-1}. Considering that $\Delta E^\ddagger$ for HAT from the α-carbon and tertiary carbon in the side chain are relatively low (see Table 3.3-7 entries 3, 6, 13, 15, page 70–71), it is reasonable to suggest that the reaction of $NO_3^\bullet$ with branched aliphatic amino acids proceeds through PCET at the amide moiety and HAT from both the α-carbon and the tertiary carbon in the side chain (when available). The upper rate coefficient for each of these HAT processes is estimated to be about 1.0 x 10^6 M^{-1} s^{-1}, similar to PCET at the amide bond. From these values for each isolated site, the higher rate coefficient determined in the reaction involving Ac-Gly-OMe (**1c**; k = 3.3 x 10^6 M^{-1} s^{-1}) can be rationalised by the

presence of two α-hydrogens and the resulting probability for a successful HAT, and also PCET at the amide moiety.

In the reaction of $NO_3^\bullet$ with Ac-Ser(OMe)-OMe (**7c'**), where the side-chain hydroxyl group is protected as a methyl ether, the higher rate coefficient of 4.6 x 10^6 M^{-1} s^{-1} can be attributed to the faster HAT at methylene hydrogens of the side chain (–C***H***$_2$OCH$_3$), in addition to PCET at the amide moiety and HAT from the α-carbon. Taking into account the contribution of the methoxy group (–CH$_2$OC***H***$_3$) of about 0.6 x 10^6 M^{-1} s^{-1}, the rate coefficient for HAT from the methylene hydrogens can be estimated to be about 1.0 x 10^6 M^{-1} s^{-1} (per H).

Figure 3.3-6 summarises the contribution from each site to the rate coefficients for the reaction of $NO_3^\bullet$ with aliphatic amino acids of type Ac-[**aa**]-OMe. The overall rate consists of the reaction at the most reactive sites: (i) PCET at the amide moiety, (ii) HAT from the α-carbon, and (iii) HAT from activated positions in the side chain (when available), such as tertiary C–H bonds or hydrogens adjacent to an ether oxygen. Abstraction from non-activated methyl groups such as the methoxy group is slower by about one order of magnitude and does not significantly contribute to the reactivity of single amino acids towards $NO_3^\bullet$.

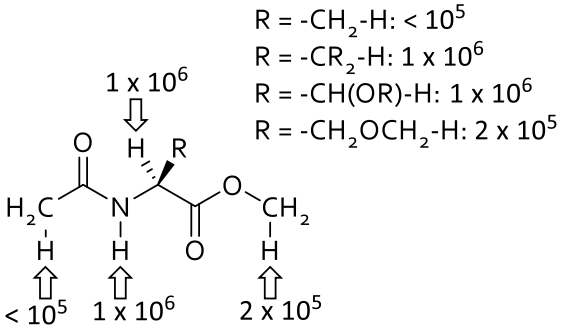


Figure 3.3-6. Contributions to the rate coefficients per hydrogen atom (in M^{-1} s^{-1}) for the reaction of $NO_3^\bullet$ at various sites in aliphatic amino acids in acetonitrile.

Based on this, the rate coefficients obtained for dipeptides and tripeptides in Table 3.3-3, page 58 can be attributed as the sum of all reactive sites possible in the peptide moiety. The reactions involving dipeptides **9c–13c** have the rate coefficients about (5.9 ± 0.6) x 10^6 M^{-1} s^{-1} (see Table 3.3-3 entries 1–5, page 58), which are made up from the rate increments for reactions at the amide bonds through PCET, the α-carbons and the tertiary C–H bonds in the side chain (where available) through HAT. Similarly, the rate coefficients of about (9.9 ± 0.6) x 10^6 M^{-1} s^{-1} (see Table 3.3-3 entries 6 and 7, page 58) for the reactions involving tripeptides **14c** and **15c** are comprised of reactions at those reactive sites (i.e., three amide bonds, α-carbons and tertiary bonds). The kinetics of longer peptide chains cannot be studied because of their low solubility in acetonitrile.

3.4 Summary of findings

Both kinetic and computational studies have shown that there are multiple reactive sites for $NO_3^\bullet$ attack and the reactions with aliphatic amino acids and peptides are kinetically controlled. Through process of isolation and elimination, the reaction at the amide N–H moiety was found to be an important contributor through a proton-coupled electron transfer (PCET) mechanism. In addition to this, HAT also takes place to a certain extent at various sites, including α-carbons, tertiary carbons and 'activated' C–H bonds.

The key findings of this chapter are:

- ✓ The reaction of $NO_3^\bullet$ with the amide N–H moiety proceeds through PCET via the formation of a charge-transfer complex with the rate coefficient of approximately 1.0 x 10^6 M^{-1} s^{-1}.
- ✓ The electron density at the amide nitrogen greatly affects the rate for PCET as demonstrated by a drop of the rate coefficient by a factor of about two in the reaction of $NO_3^\bullet$ with Tfa-Aib-OMe (**16i**).
- ✓ When possible, $NO_3^\bullet$ abstracts hydrogens from the α-carbons, tertiary C–H bonds, and/or 'activated' C–H bonds (e.g., those that are adjacent to an electron-donating group), with a similar rate coefficient of 1.0 x 10^6 M^{-1} s^{-1}. The weak C–H bonds at these sites were revealed through a relatively low $\Delta E^\ddagger$ obtained from the computational studies.
- ✓ The rate coefficients obtained for reactions involving aliphatic dipeptides and tripeptides can be ascribed to the sum of all possible reactive sites within the molecule through PCET and/or HAT mechanism.

- ✓ Electronic effects play an important role in determining the reactivity of a substrate towards $NO_3^{\bullet}$. For example, N-phthaloylated amino acids exhibit low reactivity in general because of its strong electron-withdrawing properties, which shut down any possible HAT at the α-carbon adjacent to that group.
- ✓ Rate amplification was obtained in the reaction of $NO_3^{\bullet}$ with Ac-Lys(NHAc)-OMe (**8c'**). The rate increment reached a factor of 500 compared to the other aliphatic amino acids used in this work. It was suspected that the acetamide group in the side chain plays a role in this dramatic rate amplification. Two major pathways have been proposed, which are PCET at the side chain amide and HAT from the ε-carbon. The rates for these pathways are faster because of the higher electron density at these moieties, compared to the other aliphatic amino acids.

Some work in this chapter was reported in the following publication:

"Reaction of Amino Acids, Di- and Tripeptides with the Environmental Oxidant $NO_3^{\bullet}$: A Laser Flash Photolysis and Computational Study"

Nathanael, J. G.; Hancock, A. N.; Wille, U. *Chem. Asian J.* **2016**, *11*, 3188–3195.

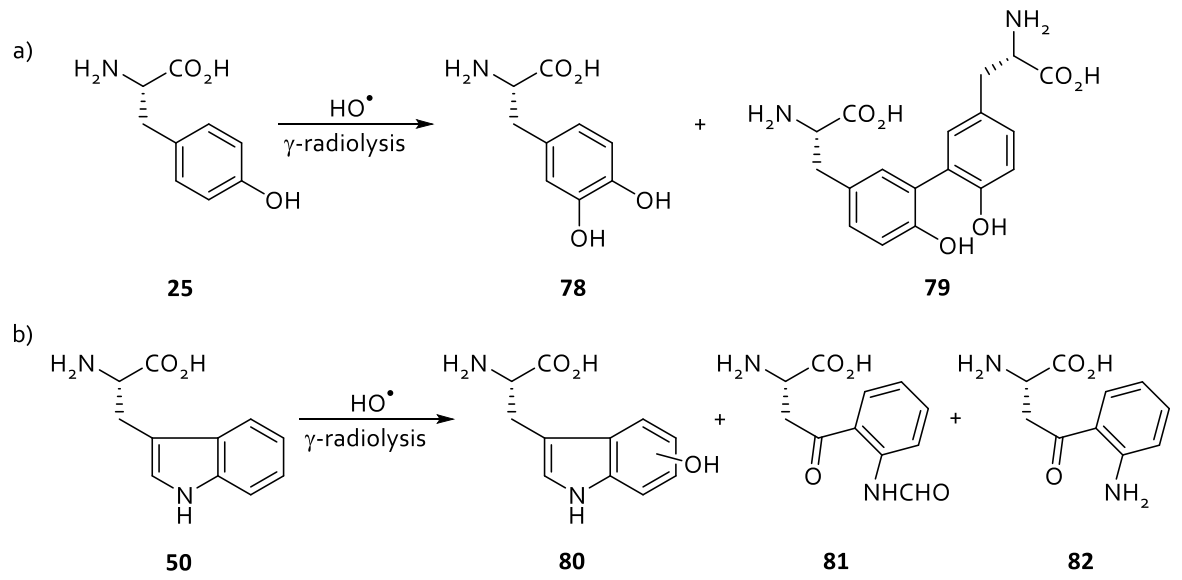
"Oxidative Damage in Aliphatic Amino Acids, Di- and Tripeptides by the Environmental Free Radical Oxidant $NO_3^{\bullet}$: The Role of the Amide Bond Revealed by Kinetic and Computational Studies"

Nathanael, J. G.; Wille, U. *J. Org. Chem.* **2019**, *84,* 3405–3418.

Chapter Four: Reaction of nitrate radicals with aromatic amino acids and peptides

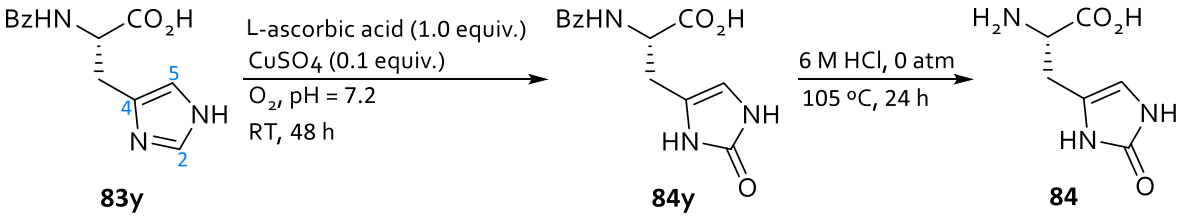
4.1 Introduction

There are far fewer aromatic amino acids than aliphatic ones with only 4 out of 22 proteinogenic amino acids being aromatic. However, aromatic amino acids are generally considered to be more susceptible to oxidative damage because of their low redox potentials [for tyrosine and tryptophan, $E^0 \approx 1$ V vs NHE at pH = 7].[23,72] The oxidative modifications often take place exclusively at the aromatic ring and are usually irreversible because the initially formed intermediate rapidly undergoes subsequent reactions.[73] It has been shown that reactive oxygen species (ROS), such as HO$^\bullet$, which can be obtained through γ-hydrolysis of water, react with tyrosine **25** to give the hydroxy derivative **78** and the dityrosine cross-linked derivative **79** (Scheme 4.1-1a).[73b,74] A similar reaction to give various hydroxylated derivatives **80** was observed in the reaction of HO$^\bullet$ with tryptophan **50**, however, HO$^\bullet$ also readily targets tryptophan **50** to form N_ε-formylkynurenine (**81**) and kynurenine (**82**) through oxidation of the more electron-rich pyrrol ring (Scheme 4.1-1b).[73b,75]



Scheme 4.1-1. Reactions of HO• with: a) tyrosine (**25**) and b) tryptophan (**50**).[73b,74,75]

Oxidation of histidine occurs at the most electron-rich carbon C-2 to afford 2-oxohistidine (Scheme 4.1-2).[73b] Under radical conditions (Cu/ascorbate), *N*-benzoyl histidine (**83y**) is converted to 2-oxohistidine (**84**), following the acid hydrolysis of the intermediate *N*-benzoyl-2-oxohistidine (**84y**).[76]



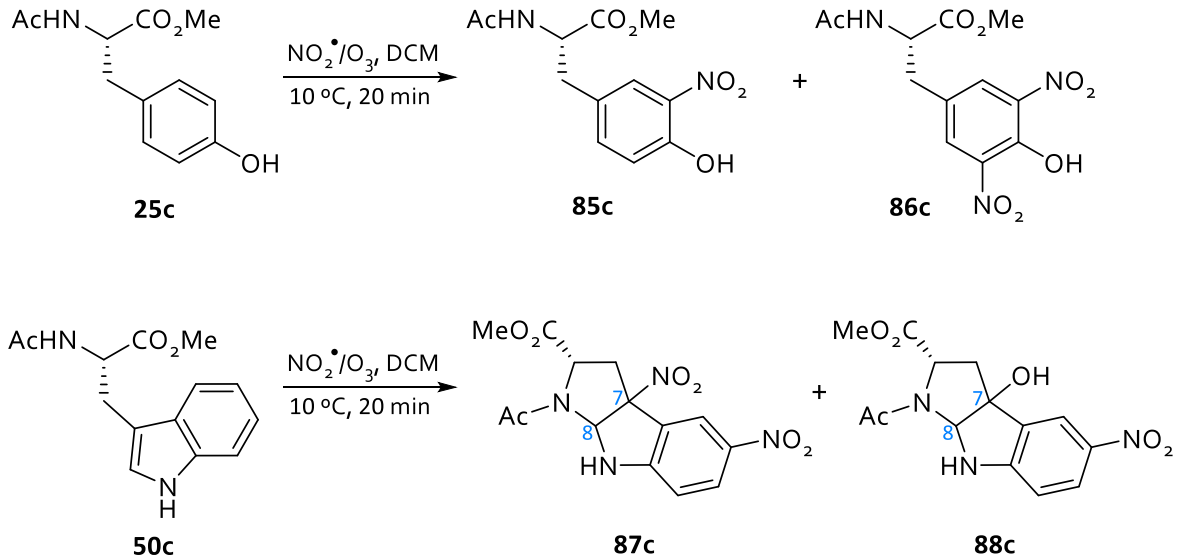
Bz = PhCO–

Scheme 4.1-2. Oxidation of histidine **83y** by the copper/ascorbate system.[76]

The oxidation product **84** has been proposed as a biological marker for oxidative damage as it was shown that after treating bovine serum albumin (BSA) with the copper/ascorbate system and hydrolysing the mixtures, 2-oxohistidine was found to be the major modification product.[76] The result suggests that detection of 2-oxohistidine in mammalian cells can be useful to assess the accumulation of modified proteins under oxidative stress.

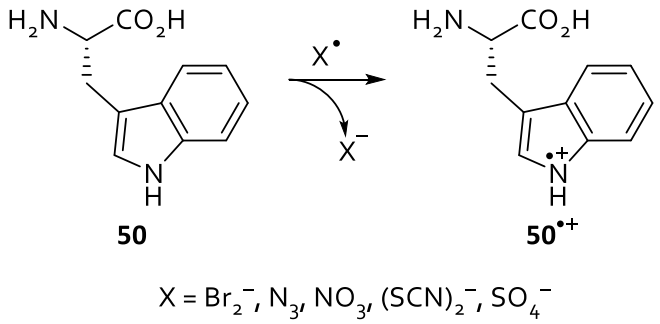
Histidine susceptibility towards oxidation seems to be dependent on the pH of the environment. For example, in the photooxidation of histidine residues in papain, the rate of oxidation drops at pH < 6.[77a] This is due to the protonation of the imidazole side chain of histidine (pK_a ≈ 6.6 ± 1 depending on the protein environment),[77b] which significantly reduces its electron density and ultimately its susceptibility to oxidation.

Another irreversible oxidative damage has been reported for the reaction of $NO_3^\bullet$ (generated from $NO_2^\bullet/O_3$ mixtures)[11] with aromatic amino acids.[78] Aromatic ring of tyrosine **25c** was nitrated to give nitrotyrosine derivatives **85c** and **86c**, while oxidative cyclisation of tryptophan **50c** was observed to afford diastereomeric pyrroloindolines **87c** and **88c** (chiral centre at both C-7 and C-8) following a short exposure time (Scheme 4.1-3).[78a]



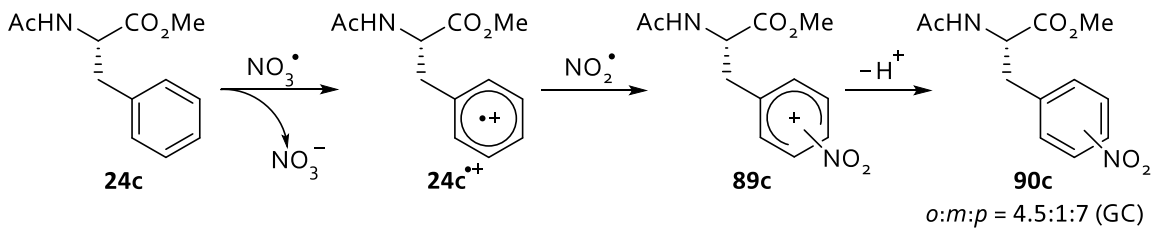
Scheme 4.1-3. Reactions of $NO_2^\bullet/O_3$ mixtures with tyrosine **25c** and tryptophan **50c**.[78a]

In most cases, electron transfer (ET) at the aromatic ring has been proposed as the initial step of the oxidative damage. For example, the reaction of tryptophan **50** with most of electrophilic radicals (e.g., $Br_2^{\bullet-}$, $N_3^\bullet$, $NO_3^\bullet$, $(SCN)_2^{\bullet-}$, $SO_4^{\bullet-}$) proceeds through a one-electron charge transfer at the nitrogen of the side chain to give the radical cation $\mathbf{50}^{\bullet+}$ (Scheme 4.1-4).[47b,75b,78a]



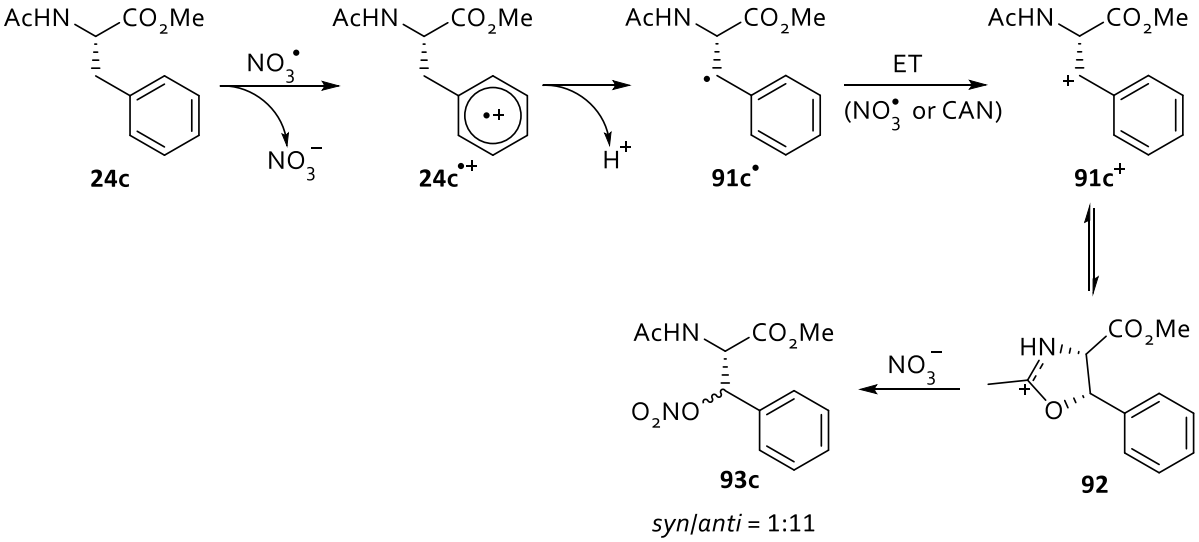
Scheme 4.1-4. Initial ET for the reaction of tryptophan (**50**) with electrophilic radicals.[47b,75b,78a]

Phenylalanine, on the other hand, is unlike the other three aromatic amino acids. Its aromatic ring is considered as more inert to oxidation under biochemical conditions because the redox potential of an alkylated phenyl group is about 2 V versus NHE, which is higher than the rest.[78a,79] Wang and co-workers reported that phenylalanine, compared to methionine, histidine and tryptophan, is the least vulnerable in terms of the degree of oxidation.[80] Thus, phenylalanine has not been considered as readily oxidisable. However, it surprisingly undergoes significant chemical changes in its reaction with $NO_3^\bullet$.[58c,78] The outcome of the reaction varies depending on the source of $NO_3^\bullet$ used in the reaction. Under simulated atmospheric conditions ($NO_3^\bullet$ obtained from the mixture of $NO_2^\bullet/O_3$ according to eq. 1-1, page 4),[11] the reaction of $NO_3^\bullet$ with phenylalanine **24c** results in mononitration of the aromatic ring at various position (*o*:*m*:*p* = 4.5:1:7 based on GC analysis).[78a] Scheme 4.1-5 shows that the reaction presumably proceeds through an initial ET of the aromatic ring by $NO_3^\bullet$ to generate the radical cation intermediate **24c**$^{\bullet+}$,[56,57c,81] which can then be trapped by $NO_2^\bullet$ to give isomeric nitrophenylalanines **90c** after deprotonation of the adduct **89c**.[78b]



Scheme 4.1-5. Reaction of $NO_3^\bullet$ with Phe **24c** in the presence of the radical cation trap $NO_2^\bullet$.[78a]

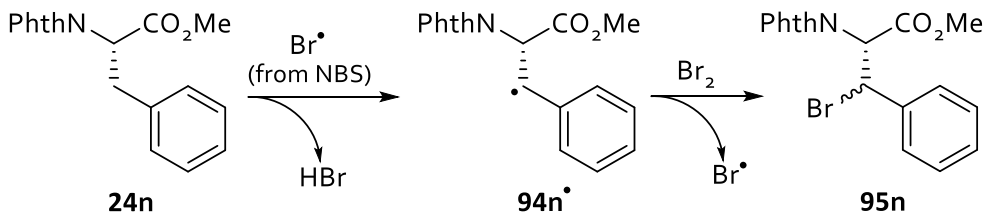
In the absence of $NO_2^{\bullet}$, when $NO_3^{\bullet}$ is generated from CAN photolysis under preparative conditions (according to eq 3.1, page 47),[56,57] benzylic deprotonation in **24c**$^{\bullet+}$ takes place following the initial ET to generate benzyl radical **91c**$^{\bullet}$ (Scheme 4.1-6).[58c] Under the oxidative environment of the experiment (due to the presence of $NO_3^{\bullet}$ and CAN), the benzyl cation **91c**$^{+}$ can be formed and its recombination with NO_3^{-} would afford the diastereomeric β-nitrated Phe derivative **93c**.



Scheme 4.1-6. Reaction of $NO_3^{\bullet}$ with Phe **24c** in the absence of the radical cation trap $NO_2^{\bullet}$.[58c]

The high diastereoselectivity for the *anti* isomer in the formation of **93c** indicates stabilisation of the intermediate benzyl cation **91c**$^{+}$ through a cyclic oxazolinium ion **92** that directs the attack of $NO_3^{\bullet}$ from the opposite side.

Due to the relatively lower susceptibility of the aromatic ring, phenylalanine reacts with less oxidising radicals, such as $Br^{\bullet}$, through a different mechanism.[82] Benzylic hydrogen abstraction was found to be the major pathway in the reaction of the phenylalanine derivative **24n** with *N*-bromosuccinimide (NBS) to give the bromide **95n** via the benzylic radical **94n**$^{\bullet}$ (Scheme 4.1-7).



Scheme 4.1-7. Benzylic hydrogen abstraction of Phe **24n** by $Br^{\bullet}$.[82]

This chapter will focus mainly on the reactivity of phenylalanine as an amino acid and a part of small peptides towards $NO_3^{\bullet}$, utilising kinetic studies as the main tool of investigation.

4.2 Aim

As per the discussion above, there are few possible pathways for radical reactions of phenylalanine. Following from the previous findings of product studies, the work presented in this chapter intends to further consolidate the mechanism for the reactions of $NO_3^{\bullet}$ with aromatic amino acids, especially phenylalanine, through kinetic and computational studies. Computational studies were performed by Prof. Paul R. Rablen (Swarthmore College, PA, USA).

Specifically, the aims of this chapter were to:

- ✓ determine the absolute rate coefficients for the reactions of $NO_3^{\bullet}$ with aromatic amino acids and peptides, especially phenylalanine, in acetonitrile,
- ✓ incorporate aromatic amino acids into a dipeptide to investigate the influence of peptide environments on the aromatic ring oxidation by $NO_3^{\bullet}$, and
- ✓ vary the position of phenylalanine residue in a short peptide chain to explore its reactivity towards $NO_3^{\bullet}$.

4.3 Results and Discussion

4.3.1 Reaction of $NO_3^\bullet$ with N- and C-protected phenylalanine and phenylalanine-containing dipeptides

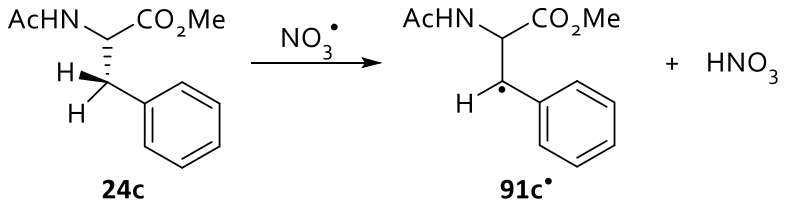
Table 4.3.1 compiles the absolute second-order rate coefficients, *k*, for the reactions of $NO_3^\bullet$ with phenylalanine and phenylalanine-containing dipeptides. Under the same reaction conditions (see Section 3.3.2, page 52), the rate coefficient for the reaction of $NO_3^\bullet$ with Ac-Phe-OMe (**24c**) was found to be 1.1 x 10^7 M^{-1} s^{-1} (entry 1). This higher rate coefficient compared to those determined for aliphatic amino acids [*k* = (1.4–3.3) x 10^6 M^{-1} s^{-1}, Table 3.3-1, page 54] can be taken as an indication that the reaction occurs at the aromatic ring. As previously found from the product studies, ET at the aromatic ring is likely to be the initial step.

Table 4.3-1. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with N- and C-protected phenylalanine and phenylalanine-containing dipeptides in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^7 M^{-1} s^{-1})
1	Ac-Phe-OMe (**24c**)	1.1
2	Ac-Phe(β,β-d_2)-OMe (**26c**)	1.0
3	Ac-Gly-Phe-OMe (**27c**)	2.0[b]
4	Ac-Val-Phe-OMe (**28c**)	1.1
5	Ac-Leu-Phe-OMe (**29c**)	1.1
6	Ac-Phe-Gly-OMe (**30c**)	3.6
7	Ac-Phe-Val-OMe (**31c**)	3.6
8	Ac-Phe-Leu-OMe (**32c**)	3.6[b]
9	Ac-Phe-Glu(OMe)-OMe (**33c**)	3.5
10	Ac-Phe-Phe-OMe (**34c**)	6.9

[a] Experimental error ±10%. [Substrate]:[CAN] = 10–100. [b] Experimental error ±15%.

However, in order to exclude that HAT from the β-carbon of Ac-Phe-OMe (**24c**) occurs to produce a resonance-stabilised benzylic radical intermediate **91c**• (Scheme 4.3-1), as has been reported previously in the reaction of Br• with phenylalanine,[82] KIE studies were carried out with a β,β-bis-deuterated Phe **26c**.



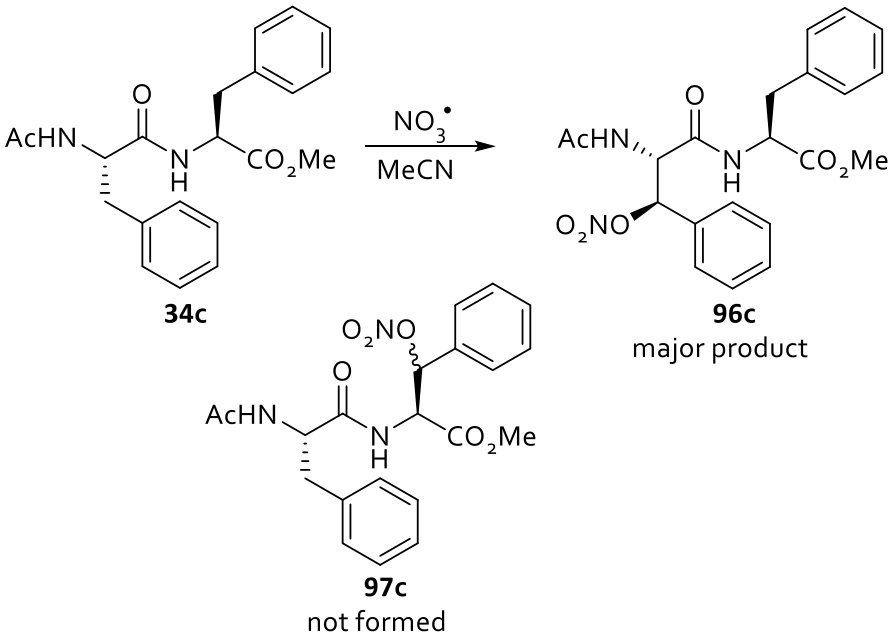
Scheme 4.3-1. HAT from the β-carbon of **24c** produced a stable benzylic radical **91c**•.

If the reaction proceeded through direct benzylic hydrogen abstraction, the rate coefficient for the reaction of $NO_3^•$ with Ac-Phe(β,β-d_2)-OMe (**26c**) would be expected to decrease. It was found that the rate coefficient did not change dramatically (k = 1.0 x 10^7 M^{-1} s^{-1}; entry 2). A typical primary KIE for hydrogen abstraction would exhibit rate retardation by a factor of 6 or more.[60b,83] Therefore, the absence of an H/D isotope effect (in this case k_H/k_D = 1.1) clearly supports the suggestion that the reaction does not occur through HAT but proceeds through an initial ET at the aromatic ring.

As shown previously, $NO_3^•$ with aliphatic amino acids reacts through PCET and/or HAT with a rate that is about 5–6 times slower in acetonitrile than the ET reaction with Phe **24c** (compare Table 3.3-1 entries 1–6, page 54 with Table 4.3-1 entry 1). Hence, incorporating an aliphatic amino acid residue into a phenylalanine-containing dipeptide should not affect the reactivity of phenylalanine residue towards $NO_3^•$. Indeed, the rate coefficients obtained for the reaction of $NO_3^•$ with dipeptides Ac-Gly-Phe-OMe (**27c**), Ac-Val-Phe-OMe (**28c**) and Ac-Leu-Phe-OMe (**29c**) are largely determined by the reactivity of phenylalanine residue alone (entries 3–5). A slight but significant increase of the

rate coefficients by a factor of 3 was surprisingly found for the reaction of $NO_3^\bullet$ with dipeptides Ac-Phe-Gly-OMe (**30c**), Ac-Phe-Val-OMe (**31c**), Ac-Phe-Leu-OMe (**32c**) and Ac-Phe-Glu(OMe)-OMe (**33c**), compared to those dipeptides with phenylalanine at the C-terminus (entries 6–9 vs 3–5). The rate coefficient was further increased by a factor of at least 6 for the reaction of $NO_3^\bullet$ with Ac-Phe-Phe-OMe (**34c**; entry 10). This observed accelerating effect implies that aromatic ring at the N-terminus is oxidised 3 times more rapidly than that at the C-terminus.

The finding from kinetic studies is further supported by product studies. Scheme 4.3-2 shows product studies carried out by Dr. Luke F. Gamon on this system where Ac-Phe-Phe-OMe (**34c**) was reacted with an excess of $NO_3^\bullet$ obtained from photolysis of CAN (according to eq 3.1, page 47). The major oxidation product was dipeptide **96c**, which was obtained as a single diastereomer possessing a β-nitrate substituent at the N-terminal Phe residue.



Scheme 4.3-2. Regioselective oxidation of the N-terminal Phe residue in dipeptide **34c** by $NO_3^\bullet$.

The observed diastereoselectivity in this reaction is governed by stabilisation of the intermediate benzyl cation by the adjacent acetamide group (through formation of a cyclic oxazolinium ion intermediate as described previously in Scheme 4.1-6,

page 86). The stabilisation directs the attack by NO_3^- from the opposite site, which consequently produces the *anti* diastereomer **96c**.

The fundamental difference between the two aromatic rings in dipeptide **34c** is the type of functional groups in the vicinity of the ring. The aromatic ring of the N-terminal Phe residue is flanked by two amide moieties in both the N- and C-direction, while that of the C-terminal Phe residue is flanked by one amide group and one ester group. The absence of isomeric dipeptide **97c** with a β-nitrate ester at the C-terminal Phe residue in the reaction product confirms the different susceptibility of the two aromatic rings towards oxidation. To investigate the origin of this different degree of oxidation, phenylalanine amide derivatives would be employed in the next section to serve as a simplified model system of phenylalanine-containing dipeptides.

4.3.2 Reaction of $NO_3^\bullet$ with phenylalanine amide derivatives

Table 4.3-2 compiles the absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with phenylalanine amide derivatives. The higher rate coefficients obtained for phenylalanine **24q–s** (entries 1–3) by a factor of about 7 to 15 revealed a higher reactivity of amide derivatives compared to their ester counterparts (k = 1.1 x 10^7 M^{-1} s^{-1} for Ac-Phe-OMe).

Table 4.3-2. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with phenylalanine amide derivatives in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^7 M^{-1} s^{-1})
1	Ac-Phe-NHMe (**24s**)	7.6
2	Ac-Phe-NH*t*Bu (**24r**)	8.4[b]
3	Ac-Phe-NH_2 (**24q**)	15.8

[a] Experimental error ±5%. [Substrate]:[CAN] = 2–30. [b] Experimental error ±20%.

This accelerating effect by a neighbouring amide group has been previously reported in the literature.[82,84] The reaction of $Br^{\bullet}$ with phenylalanine *t*-butyl amide through benzylic HAT was found to be 4 times faster than that with phenylalanine methyl ester.[82] This enhanced reactivity was rationalised by the higher nucleophilicity of an amido substituent compared to an ester group, which enables a stabilising interaction with an electron-deficient carbon centre in the HAT transition state (Figure 4.3-1a). Another example is given in the work by Glass and Schöneich, who showed in pulse radiolysis experiments, that the redox potential of dialkyl thioethers was reduced by over 0.5 V when an adjacent amide group could stabilise a positive charge on the sulfur atom through formation of a two-centre three-electron S–O bond (Figure 4.3-1b).[84]

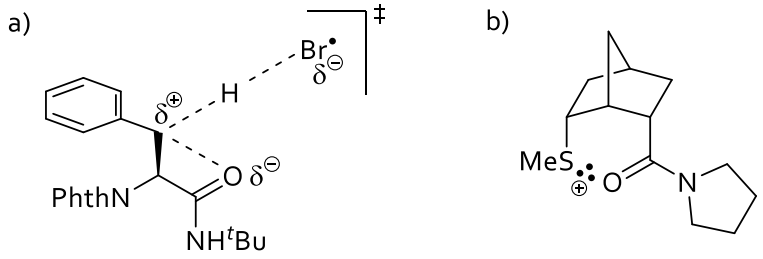


Figure 4.3-1. a) Neighbouring amide group stabilisation of the transition state for $Br^{\bullet}$-induced HAT from the benzylic position in Phe side chain.[82] b) Neighbouring amide group stabilisation following ET at the sulfur atom through formation of a two-centre three-electron (2c-3e) S–O bond.[84]

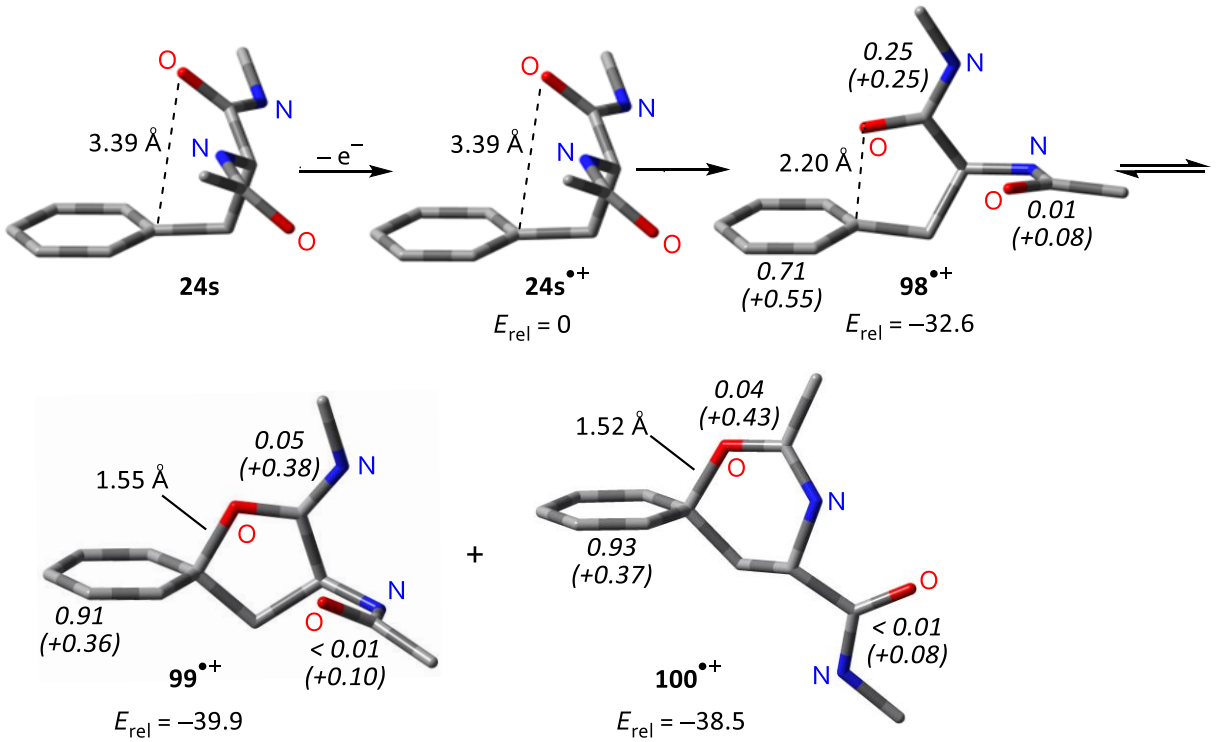
These findings give a clear message that an adjacent amide group affects the reactivity of a nearby site, such as benzyl or thioether groups. Therefore, it is reasonable to suggest that the presence of an amide functional group at the C-terminus may accelerate the oxidation rate of phenylalanine by reducing the redox potential of the aromatic ring.

Direct $NO_3^{\bullet}$ oxidation at the amide nitrogen, as found with aliphatic amino acids, can be excluded because a faster reaction with the more electron-rich secondary amides **24r** and **24s** than with the primary amide **24q** would be expected if this were the case. Moreover, the magnitude of the rate coefficients for such

reaction (in the form of PCET) is in the range of 10^6 M^{-1} s^{-1} (Section 3.3.5, page 76), which is lower than the obtained rate coefficients for amides **24q–s**. HAT from the 'activated' methyl amide in **24s**, which has been shown as one of the major pathways for the reaction of $NO_3^\bullet$ with Ac-Lys(NHAc)-OMe (see Section 3.3.3.3, page 67 and Section 3.3.4, page 69), can also be excluded because the absence of the 'activated' methyl amide in **24r** did not affect the reaction rate (Table 4.3-2 entry 1 vs 2).

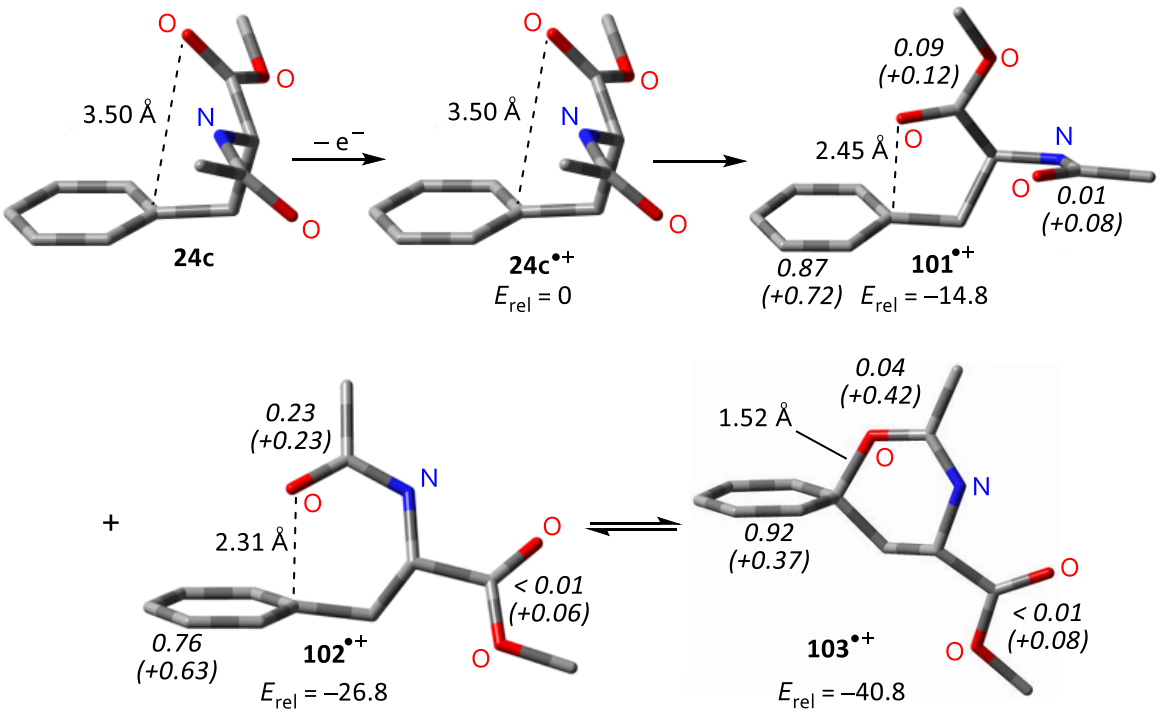
In order to further consolidate the results from the kinetic experiments, quantum chemical CBS-QB3//M062X/6-31* calculations were performed by Prof. Paul R. Rablen for the radical cations of Ac-Phe-NHMe (**24s**) and Ac-Phe-OMe (**24c**), generated following ET at the aromatic ring by $NO_3^\bullet$.

The radical cation **24s**$^{\bullet+}$, generated from Ac-Phe-NHMe (**24s**) can be stabilised by the two neighbouring amide groups through the formation of a covalent bond ("σ complex") in **99**$^{\bullet+}$ and **100**$^{\bullet+}$ (Scheme 4.3-3). These σ complexes (more stable by about 40 kJ mol^{-1}) were formed from a π complex (only one π complex, **98**$^{\bullet+}$, is shown here, which leads to **99**$^{\bullet+}$). The π complex **98**$^{\bullet+}$ represents the transition state for ET, where amide neighbouring group participation enables delocalisation of charge and spin (of about 25%) from the radical cation of the aromatic ring to this amide group. The delocalisation of the positive charge in **98**$^{\bullet+}$ onto the amide moiety can be rationalised as the source of extra stabilisation of the radical cation **24s**$^{\bullet+}$.



Scheme 4.3-3. Stabilisation of the radical cation of amide **24s** through neighbouring-group participation. Energies E_{rel} [kJ mol^{-1}] (in acetonitrile) relative to **24s**$^{\bullet+}$, which retains the conformation of neutral **24s**. Italic numbers: spin densities and increases in Mulliken atomic charges from **24s** (in brackets) summed up for the phenyl ring and the amide moieties.

In the case of Ac-Phe-OMe (**24c**), the radical cation **24c**$^{\bullet+}$ is stabilised by one amide and one ester group. The amide group stabilises the aromatic radical cation in a similar fashion described for Ac-Phe-NHMe (**24s**) above, which is through the formation of σ complex **103**$^{\bullet+}$ via a π complex **102**$^{\bullet+}$ (Scheme 4.3-4). On the other hand, stabilisation by an ester group through a π complex **101**$^{\bullet+}$ reduces the energy by only about 15 kJ mol^{-1} relative to that of a radical cation **24c**$^{\bullet+}$, which is less than stabilisation by the amide group. There is also less change in Mulliken charges at the ester moiety (only about 12%), which shows that an ester group is not as involved as an amide group in stabilising the developing positive charge during the ET.



Scheme 4.3-4. Stabilisation of the radical cation of ester **24c** through neighbouring-group participation. Energies E_{rel} [kJ mol^{-1}] (in acetonitrile) relative to **24c**$^{\bullet+}$, which retains the conformation of neutral **24c**. Italic numbers: spin densities and increases in Mulliken atomic charges from **24c** (in brackets) summed up for the phenyl ring and the amide/ester moieties.

The fast $NO_3^{\bullet}$-induced oxidation of phenylalanine suggests an early transition state that is energetically close to the ground state of the neutral amino acid. In the amide **24s** and ester **24c**, the oxygen atoms of the C-terminal carbonyl groups are about 3.4–3.5 Å away from the *ipso* carbon atom of the phenyl group and, thus, well positioned to stabilise the developing positive charge on the aromatic ring following ET. However, due to the higher nucleophilic properties of an amide group, a neighbouring amide group can stabilise the developing radical cation at the aromatic ring more efficiently than a neighbouring ester group. This effect causes a faster rate of oxidation of **24s**, relative to **24c**. The same effect also governs oxidation reactions of longer peptides, which will be discussed in the next section.

4.3.3 Reaction of $NO_3^•$ with phenylalanine-containing tripeptides

Tripeptides were also employed to explore the role of amide neighbouring group effects in larger systems. Unfortunately, all tripeptides have low solubility in acetonitrile, hence, lower concentration range was employed in the laser flash experiments. The absolute second-order rate coefficients for the reaction of $NO_3^•$ with phenylalanine-containing tripeptides are compiled in Table 4.3-3.

Table 4.3-3. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^•$ with phenylalanine-containing tripeptides in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^7 M^{-1} s^{-1})
1	Ac-Val-Val-Phe-OMe (**39c**)	2.3[b]
2	Ac-Val-Phe-Val-OMe (**40c**)	6.3
3	Ac-Phe-Val-Val-OMe (**41c**)	7.6
4	Ac-Leu-Leu-Phe-OMe (**42c**)	1.9[b]
5	Ac-Leu-Phe-Leu-OMe (**43c**)	6.3
6	Ac-Phe-Leu-Leu-OMe (**44c**)	6.9
7	Ac-Phe-Gly-Phe-OMe (**45c**)	8.8
8	Ac-Phe-Leu-Phe-OMe (**46c**)	7.6
9	Ac-Phe-Phe-Val-OMe (**47c**)	9.1
10	Ac-Phe-Phe-Leu-OMe (**48c**)	10.1

[a] Experimental error ±10%. [Substrate]:[CAN] = 1–3. All tripeptides have low solubility in acetonitrile. [b] Experimental error ±30%.

Tripeptides **40c**, **41c**, **43c** and **44c**, in which Phe has a C-terminal amide bond, are also oxidised about 3 times faster than tripeptides **39c** and **42c**, in which Phe is located at the ester-protected C-terminus (entries 2, 3, 5 and 6 vs 1 and 4). A further rate increment was observed when two phenylalanine residues are incorporated into the peptide chain. The increment was not dramatic when one Phe is still at the C-terminus while the other Phe is at the N-terminus (entries 7 and 8).

However, the rate coefficients increased by a factor of up to 10, when both phenylalanine residues are flanked by amide functional groups in both the N- and C-direction (entries 9 and 10). This increase of rate coefficients shows that phenylalanine embedded in a peptide environment is considerably more susceptible towards oxidation by $NO_3^{\bullet}$ than the single amino acid, due to the assistance from neighbouring amide moieties that stabilise the developing positive charge on the aromatic ring.

The greater oxidisability of phenylalanine in a peptide environment enables its aromatic ring to act as a relay amino acid. Long-distance ET has been known to be broken down into a multistep process, utilising easily oxidisable amino acid side chains as the stepping stones to relay the electron, since the superexchange process over 20 Å is a slow process (Figure 4.3-2).[85] This multistep process, called a hopping mechanism, breaks up a long ET process into several short ones and the rate was found to increase dramatically.[86]

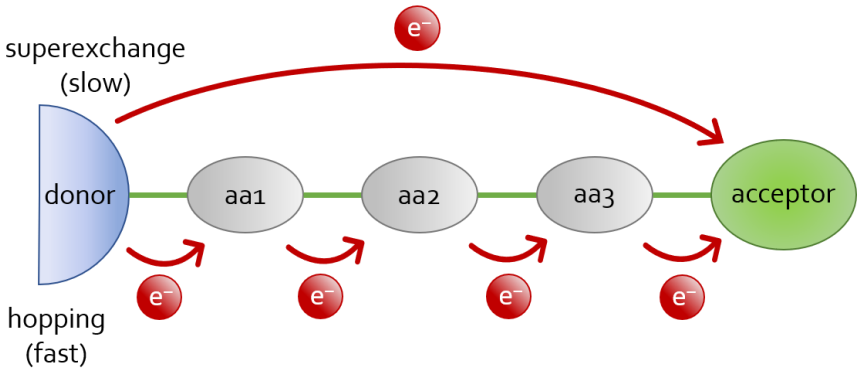
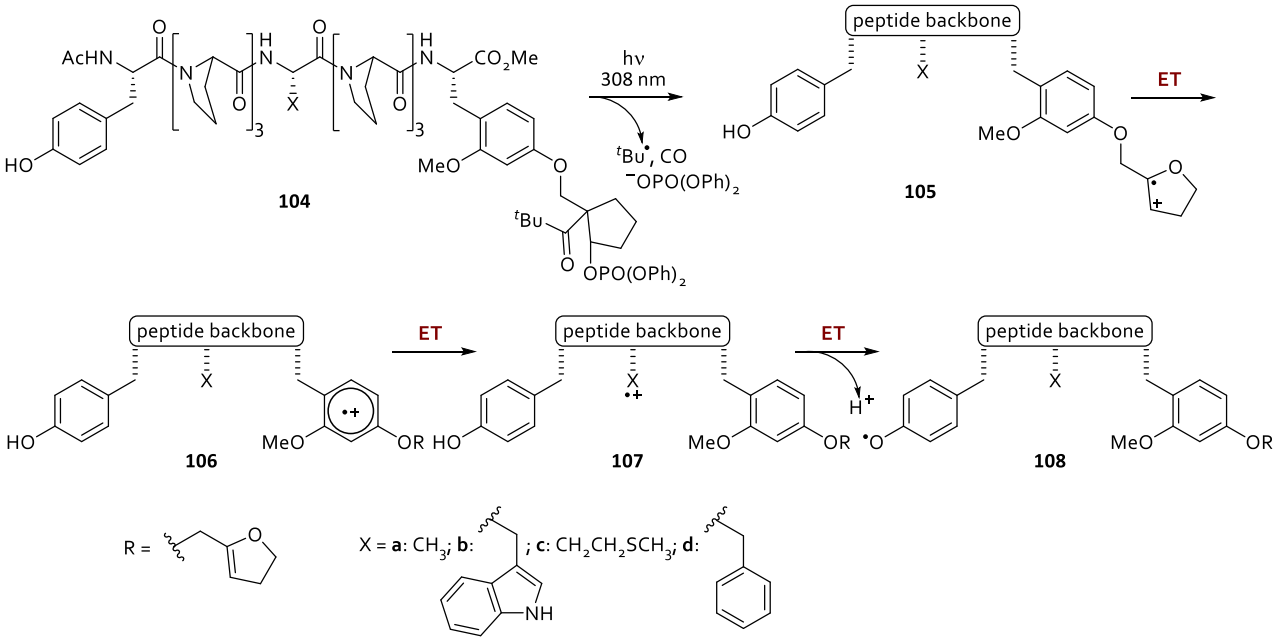


Figure 4.3-2. ET through peptides can proceed through a superexchange process or a faster hopping mechanism using side chains of amino acids (**aa1–3**) as stepping stones.[85]

Giese and co-workers have previously developed an assay to measure ET efficiencies in peptides, using nonapeptide **104** based on Scheme 4.3-5.[87]



Scheme 4.3-5. Peptide assay used to study long-distance ET in peptides.[87b]

Aryl radical cation **106**, generated from laser flash photolysis of **104** at λ = 308 nm via intermediate **105**, acts as the electron acceptor, whereas the N-terminal tyrosine of the nonapeptide **104** acts as the electron donor. It was found that the tyrosyl radical **108** was formed when tryptophan, methionine or phenylalanine residue is present along the peptide backbone to act as a relay for the long-distance ET process.[87] In contrast, intermediates **107a** and **108a** were not detected because the methyl side chain in alanine (X = CH_3) is very difficult to oxidise, and therefore, cannot act as a relay amino acid. The result was not surprising for tryptophan due to its low oxidation potential (E^o ≈ 1 V vs NHE at pH = 7).[23] Interestingly, methionine can also function as a relay amino acid because its oxidation potential is lowered by about 0.5 V through stabilisation from a neighbouring pyrrolidine amide, as outlined previously (see Figure 4.3-1b, page 92).[84,87a] It is then reasonable to suggest that even though phenylalanine itself is not readily oxidisable, amide neighbouring group effects facilitate formation of phenylalanine radical cation, so that phenylalanine can act as a stepping stone for the long-distance ET.

In fact, the role of phenylalanine as a relay amino acid has been proposed by Reguera and co-workers based on the structural evaluation of pili which revealed that some phenylalanine residues are located at the prime position for ET (Figure 4.3-3).[88]

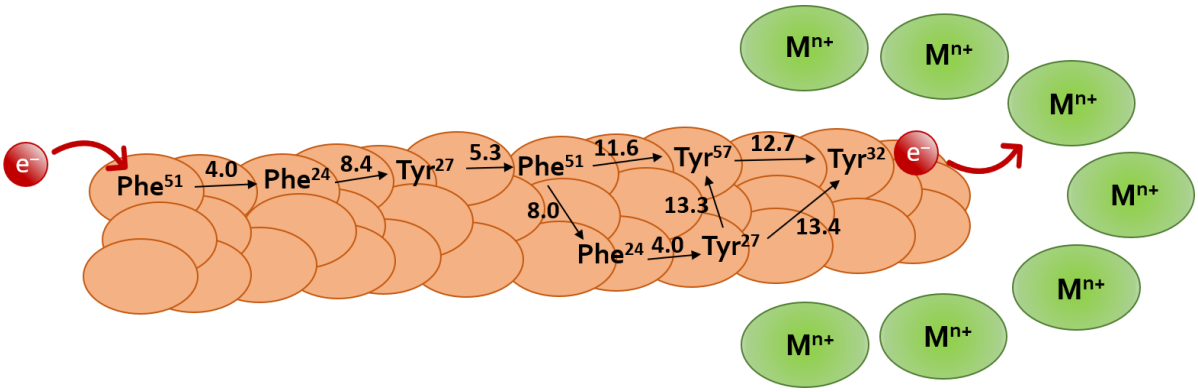
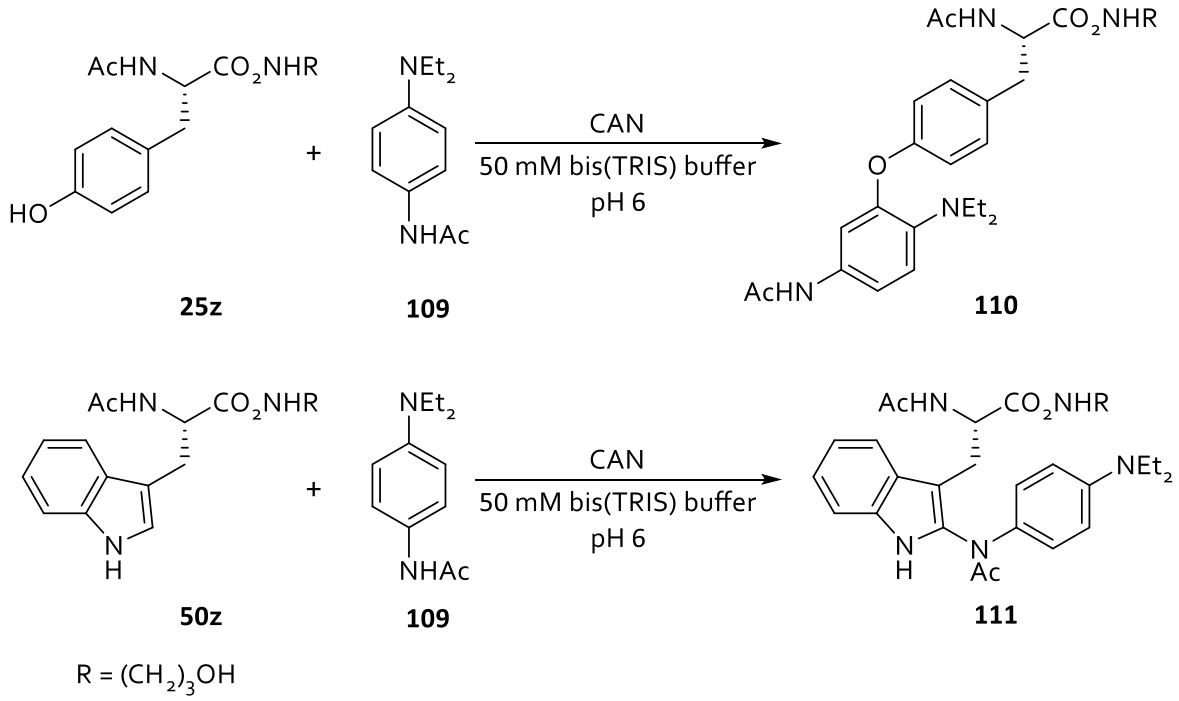


Figure 4.3-3. Pili transport electrons using Tyr and Phe residues as relay amino acids. Orange ovals are peptides that aggregates to the pili. Numbers above the arrows are interaromatic distances (in Å) for the Asp2 pilus (cell constituents are omitted for clarity).[88]

The distance between the aromatic residues (Tyr and Phe) ranges from 4.0 to 13 Å, which are all less than 20 Å, and makes them the ideal stepping stones to support multistep hopping ET. This structural evaluation can be seen as a proof that in nature, phenylalanine indeed relays electrons and aids in a long-distance ET in a peptide chain.

4.3.4 Reaction of $NO_3^{\bullet}$ with other aromatic amino acids

As previously outlined, tyrosine and tryptophan are considerably more reactive in oxidation reactions than phenylalanine. In fact, both amino acids readily react with CAN, the $NO_3^{\bullet}$ precursor used in the laser flash experiments. It has been shown that in the presence of CAN, tyrosine and tryptophan derivatives readily couple to the electron-rich aniline derivative **109** following one-electron oxidation at the aromatic ring by CAN (Scheme 4.3-6).[89]



Scheme 4.3-6. Oxidation of tyrosine **25z** and tryptophan **50z** by CAN, followed by coupling to an electron-rich aromatic compound **109**.[86]

In order to 'tame' their high reactivity, the side chains in tyrosine and tryptophan were acetylated to decrease the electron density of the side chains and reduce their susceptibility towards oxidation. The absolute second-order rate coefficients for the reaction of $NO_3^\bullet$ with *O*-acetyl tyrosine derivatives and *N*-acetyl tryptophan derivatives are compiled in Table 4.3-4.

Table 4.3-4. Absolute second-order rate coefficients, k, for the reaction of $NO_3^\bullet$ with tyrosine and tryptophan derivatives in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^8 M^{-1} s^{-1})
1	Ac-Tyr(*O*-Ac)-OMe (**25c'**)	1.7
2	Ac-Phe-Tyr(*O*-Ac)-OMe (**35c'**)	1.6
3	Ac-Tyr(*O*-Ac)-Leu-OMe (**36c'**)	2.5
4	Ac-Tyr(*O*-Ac)-Phe-OMe (**37c'**)	2.5
5	Ac-Tyr(*O*-Ac)-Tyr(*O*-Ac)-OMe (**38c'**)	3.3
6	Ac-Trp(*N*-Ac)-OMe (**50c'**)	–[b]

[a] Experimental error ±10%. [Substrate]:[CAN] = 3–30. [b] cannot be determined (see text below).

The rate coefficients obtained for the reaction of $NO_3^\bullet$ with O-acetylated tyrosine **25c'** is about one order of magnitude higher than that for phenylalanine **24c** (compare entry 1 with Table 4.3-1 entry 1, page 88). The higher reactivity of tyrosine residue can be rationalised by the strong electron-donating substituent on the aromatic ring of tyrosine, making the ring more susceptible to oxidation by $NO_3^\bullet$. The rate coefficient obtained for the reaction of $NO_3^\bullet$ with Ac-Phe-Tyr(*O*-Ac)-OMe (**35c'**) shows that this reaction is dominated by the tyrosine residue (entry 2). However, a slight increase of the rate coefficients by about 50% was observed in the reaction of $NO_3^\bullet$ with Ac-Tyr(*O*-Ac)-Leu-OMe (**36c'**) and Ac-Tyr(*O*-Ac)-Phe-OMe (**37c'**), compared to the C-terminal O-acetylated tyrosine **35c'** (entries 3 and 4 vs 2). In both dipeptides **36c'** and **37c'**, tyrosine is located at the N-terminus where it is flanked by two amide functional groups. There seems to be a small accelerating effect from the amide neighbouring group stabilisation as previously described in Section 4.3.2, page 91. A similar observation has been reported in the reaction of $SO_4^{\bullet-}$ with dipeptides containing glycine and tyrosine or tryptophan residues, which showed an increase of the rate coefficient by a factor of 1.5–2.3 when the aromatic amino acid was at the N-terminus.[47b]

No rate enhancement was observed in dipeptide **38c'**, containing two O-acetylated tyrosine residues (entry 5). The fast oxidation on both aromatic rings in dipeptide **38c'** might possibly mask any contribution from the neighbouring amide groups.

In the case of tryptophan, no rate data could be obtained for the reaction of $NO_3^\bullet$ with Ac-Trp(*N*-Ac)-OMe (**50c'**) due to the rapid dark reactions between CAN and the amino acid **50c'**, which could be visualised by a fast colour change from yellow to colourless in the mixture of CAN and substrate solution prior to the laser flash experiments. This colour change suggests that a large portion of CAN has

already been consumed through reduction to colourless Ce(III), which demonstrates the high reactivity of tryptophan residue towards oxidation even after reducing the electron density of the side chain.

4.4 Summary of findings

This chapter has explored the reactivity of aromatic amino acids and peptides towards $NO_3^\bullet$. Kinetic data have revealed that neighbouring amide groups are able to accelerate the rate of oxidation of the aromatic ring. These results suggest that aromatic amino acid residues, especially phenylalanine, are more prone to oxidative damage in a peptide environment.

The key findings of this chapter are:

- ✓ The rate coefficient for the reaction of $NO_3^\bullet$ with Phe **24c** is about one order of magnitude higher than that with aliphatic amino acids, suggesting that the reaction takes place at the aromatic ring through ET. HAT at the benzylic position an be excluded due to the absence of a primary H/D isotope effect.
- ✓ The presence of neighbouring amide groups speeds up the ET process due to more efficient stabilisation of the developing radical cation intermediate by the C-terminal amide as opposed to a C-terminal ester, as revealed through computational studies.
- ✓ The effect was also observed in a longer peptide chain. Dipeptides with N-terminal phenylalanine are oxidised about 2–3 times faster than those with C-terminal phenylalanine. This was rationalised by the fact that the stabilisation from two neighbouring amide groups in the N-terminal phenylalanine dipeptides is greater than that from one neighbouring amide and one neighbouring ester group in the C-terminal phenylalanine dipeptides.

- ✓ Phenylalanine-containing tripeptides also exhibit the same trend. The higher oxidation rate of phenylalanine in peptide environments suggests that phenylalanine can act as a relay amino acid in a long-distance ET.
- ✓ The amide neighbouring group effect was present in tyrosine, but due to the higher reactivity of tyrosine compared with that of phenylalanine, the effect is harder to observe.
- ✓ The high reactivity of tryptophan residue, even after acetylation of the side chain to reduce its electron density, prevents it from being studied under these conditions as CAN oxidises the side chain prior to the laser flash experiments.

Some results in this chapter were reported in the following publication:

"Amide Neighbouring-Group Effects in Peptides: Phenylalanine as Relay Amino Acid in Long-Distance Electron Transfer"

Nathanael, J. G.; Gamon, L. F.; Cordes, M.; Rablen, P. R.; Bally, T.; Fromm, K. M.; Giese, B.; Wille, U. *Chem. Bio. Chem.* **2018**, *19*, 922–926.

Chapter Five: Reaction of nitrate radicals with proline and proline-containing short peptides

5.1 Introduction

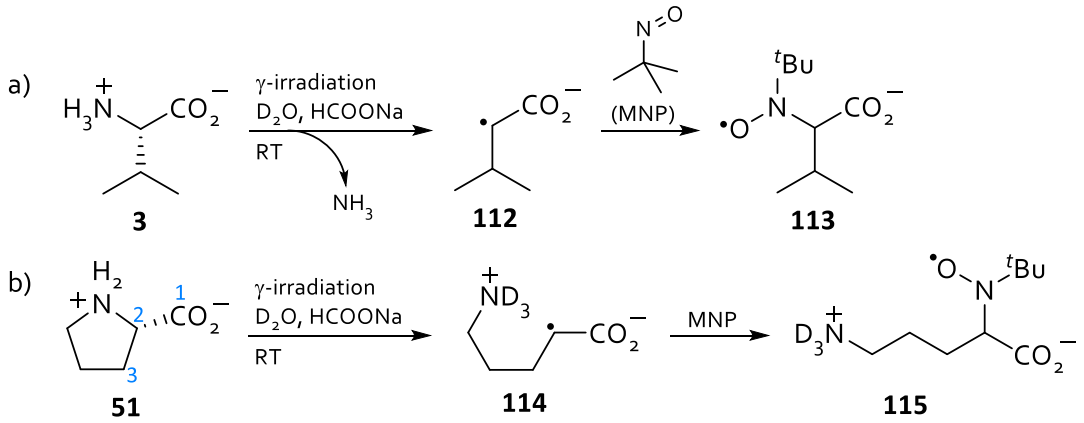
Among 22 proteinogenic amino acids, proline possesses some unique features that distinguish it from the rest. Proline is an aliphatic amino acid with its side chain covalently bound to the α-amino group, thus, forming a rigid cyclic moiety. Since its nitrogen atom is locked within a ring, the peptide bond is forced in a fixed conformation. This unique chemical structure of proline facilitates folding of many proteins, especially those rich in proline residues, by bending the regional amino acid alignment.[90]

Apart from these structural properties, proline also exhibits an exceptionally high reactivity and unique chemical pathways.[91–93] For example, in Maillard reactions where amino acids react with carbohydrates to form glycosylamines, proline was found to show the highest reactivity towards glucose among other amino acids studied, including glycine, asparagine and tryptophan.[91] This enhanced reactivity could arise from the unique cyclic secondary amine structure of proline and its ease of forming iminium salts with the carbonyl group found in an open-chain glucose (not shown).

Other than such ionic reactions, it has been shown that radical reactions involving proline also follow different reaction patterns. One example is the reaction of proline with hydrated electrons (e^-_{aq}) and highly oxidising hydroxyl radicals ($HO^\bullet$), which can be generated through γ-hydrolysis of water (eq 5.1). [92]

$$H_2O + \gamma\text{-irradiation} \rightarrow H^+ + HO^\bullet + e^-_{aq} \qquad \text{(eq 5.1)}$$

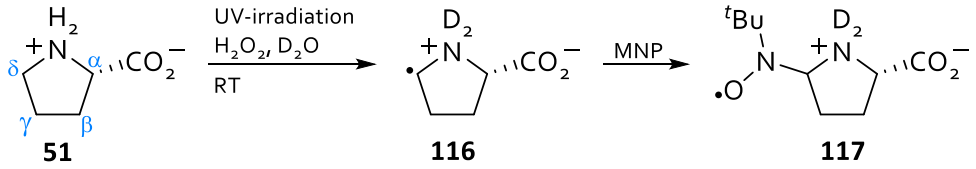
The reactions of e^-_{aq} were studied in the presence of sodium formate, HCOONa, which acts as a scavenger of H^+ and $HO^\bullet$. The reactions with amino acids in aqueous solutions lead to the formation of short-lived amino acid radicals, which are then trapped by 2-methyl-2-nitrosopropane (MNP) to yield relatively long-lived nitroxide radicals that were analysed by EPR (Scheme 5.1-1). Most of the common amino acids, for example valine (**3**), undergo reductive deamination to give a short-lived α-carbon radical **112**, which is immediately trapped by MNP to yield nitroxide **113** (Scheme 5.1-1a).[92a] In proline (**51**), on the other hand, ring opening occurs through the breaking of the C*2*–N bond to generate radical **114** (Scheme 5.1-1b). Radical-trapping of **114** by MNP then produces adduct **115**. Deuteration of the ammonium group in **114** and **115** originates from H/D exchange with the solvent, D_2O, which is required for the in situ EPR studies.



Scheme 5.1-1. Reactions of hydrated electrons (e^-_{aq}) with: a) valine (**3**) and b) proline (**51**) in the presence of 2-methyl-2-nitrosopropane (MNP) as the radical-trapping agent. Most amino acids react in the same fashion as valine (**3**).[92a]

In a similar EPR study, $HO^\bullet$ generated from UV irradiation of hydrogen peroxide (H_2O_2), was reported to react with amino acids through HAT at various sites.[92b] In the case of free amino acids, which are present as zwitterions in aqueous solution, abstraction only occurs from either the side chain or the α-carbon, but does not occur from the nitrogen of the positively charged ammonium group. In the case of proline, however, the reaction selectively occurred at the δ-position, as

it was confirmed by trapping of the radical **116** with MNP to form nitroxide **117**, which was analysed by EPR (Scheme 5.1-2).



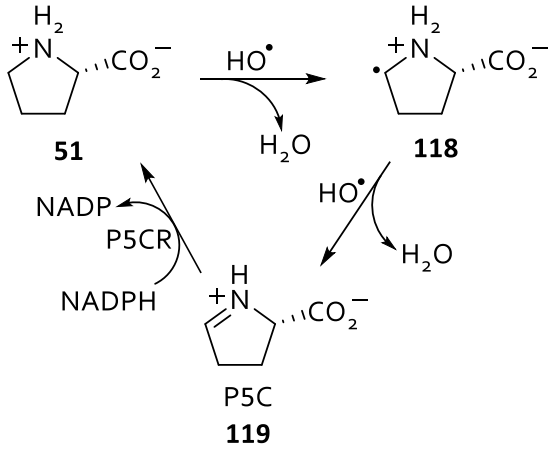
Scheme 5.1-2. Reaction of HO• with proline (**51**), followed by trapping with MNP.[92b]

Even though all amino acids were found to react with HO•, their reactivities towards HO• in proteins are not the same. Thus, in the reactions of proteins with HO• produced through the Fenton reaction of copper(I) with H_2O_2 or from γ-hydrolysis, proline and histidine were found to be the major sites of damage.[93] These reactions could lead to protein cleavage, but detailed identification of reaction products was not performed.

Because of the high reactivity of proline towards HO•, proline has been suggested to act as an antioxidant by suppressing free radical generation that is caused by oxidising radicals and closed-shell ROS.[94] Free proline is known to accumulate in plant tissues during abiotic stress, assisting in scavenging free radicals generated from the stress.[94a] Moreover, exogenously supplied proline at low concentrations (ca. 30 mM)[94b] was shown to promote the growth of plants exposed to salinity stress, decrease peroxidative damage to the lipid membranes in ground nuts and maintain nutrient status by increasing the uptake of minerals, such as potassium, calcium, phosphate, and ammonium ions in corns exposed to drought stress.[94c] On the other hand, at high concentrations, no benefit of proline was found to plants; in fact, it was shown that over-accumulated proline can be toxic for plants, for example by inhibiting growth and causing inorganic ion imbalances.[94c] The well-documented correlation between the accumulation of proline and overproduction of ROS by plants under stress conditions triggered a

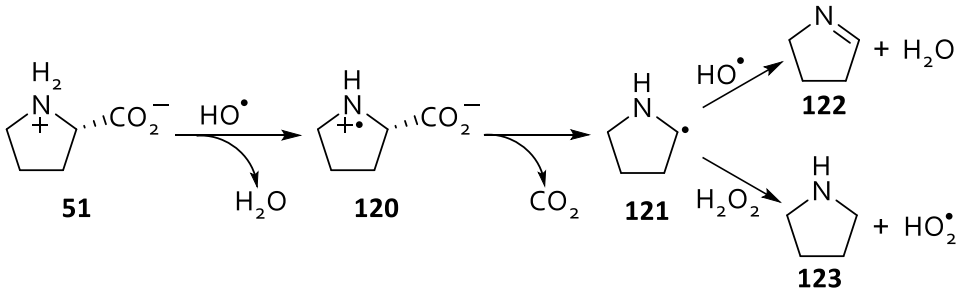
series of experimental and computational studies to elucidate the mechanisms of the reaction of HO• with proline.[95]

Computational studies revealed that the main pathway by which proline scavenges HO• is HAT.[95] This finding is in agreement with EPR studies outlined in Scheme 5.1-2 that showed that hydrogen atom at the δ-position is the most prone to abstraction.[92b] In one of the pathways proposed by Monza and co-workers, one proline molecule would scavenge at least two HO•. Thus, after initial formation of the δ-radical **118**, oxidation by a second molecule of HO• leads to formation of the iminium ion, Δ^1-pyrroline-5-carboxylate (P5C, **119**). In plants, P5C **119** can be recycled back to proline by the enzyme P5CR and consumption of one NADPH molecule per one P5C molecule (Scheme 5.1-3).[95a] This pathway is possible because the concentration of both proline and HO• significantly increase when plants are under any kind of stress.[96] In the cytosol and chloroplasts, proline concentration can reach up to 160 mM,[96b,97] higher than other common antioxidants (i.e., 1–5 mM for glutathione and 5–20 mM for ascorbate).[96c,98] Moreover, at these locations proline biosynthesis enzymes, such as P5CR, are localised, which enable the conversion of P5C (**119**) back to proline (**51**).[96c,99]



Scheme 5.1-3. Proline cycle in which proline (**51**) acts as a HO• scavenger and is recycled back by the P5CR/NADPH enzymatic system.[95a]

Another work by Monza and co-workers explored a possible pathway that involves initial HO• attack on the amino group of proline through, which leads to the decarboxylation of **51** through intermediate **120** and formation of the pyrrolidine-1-yl radical (**121**), as shown in Scheme 5.1-4.[95b] The latter can react with another HO• to yield Δ^1-pyrroline (**122**) and water. Alternatively, in the presence of hydrogen peroxide (H_2O_2), radical **121** can be oxidised to pyrrolidine (**123**), giving the conjugate acid of superoxide ($O_2^{\bullet-}$), $HO_2^{\bullet}$, as the by-product (Scheme 5.1-4).



Scheme 5.1-4. The fate of proline (**51**) upon HO• attack on the amino group.[95b]

Interestingly, the enzyme Δ^1-pyrroline dehydrogenase enables conversion of pyrroline **122** to γ-aminobutyric acid (GABA, not shown), which plays important protective roles in plants under stress conditions and has been known to accumulate in response to both abiotic and biotic stress.[95b,100] This non-enzymatic (except for the final step) alternative pathway to access GABA may highlight the role of proline as an antioxidant and a free radical scavenger in plants. Despite these comprehensive studies on the mechanism for the reaction of proline with HO• and the roles of proline in plants, far less efforts have been done to study the reactions of proline with other radicals.

Proline is abundant in the human body. Out of more than 18,000 human proteins studied, 99.8% of them contain proline.[90a] The proline metabolism plays an important and complex role in cell signalling, stress protection and energy production.[101] Moreover, proline-rich proteins have been found in the respiratory

tract and shown to be immunoreactive in serous cells of submucosal glands.[102] When air pollutants are inhaled, it is likely that proline-rich proteins will be exposed to the oxidising radicals found in pollutants, such as $NO_3^{\bullet}$, $NO_2^{\bullet}$ and O_3. Therefore, it is important to understand the mechanism for the reaction of proline with these oxidants, particularly the radical species. HAT has been thought to be the only reaction pathway proline undergoes. Kinetic studies by Venkatachalapathy and Ramamurthy suggested that $NO_3^{\bullet}$ reacts with proline through HAT, similar to any other aliphatic amino acids (see Table 3.1-2, page 47), but product analyses were not performed.[56] The work presented in this chapter will explore the reaction of $NO_3^{\bullet}$ with proline in both kinetic and product studies in order to gain fundamental understanding of the reaction mechanisms of this important amino acid with environmental radicals.

5.2 Aim

As outlined above, proline is the only proteinogenic amino acid with a side chain attached to the amino group, forming a rigid cyclic system. This unique structure enables a reaction mechanism different from the acyclic aliphatic amino acids. Kinetic studies were performed in conjunction with product studies to paint a complete picture of the reaction mechanism for the reaction of $NO_3^\bullet$ with proline. The reaction was carried out in acetonitrile and proline was protected at both N- and C-termini to increase its solubility in this solvent.

Specifically, the aims of this chapter were to:

- ✓ determine the absolute rate coefficients for the reactions of $NO_3^\bullet$ with N- and C-protected proline in acetonitrile and compare it with the other aliphatic amino acids,
- ✓ incorporate proline in a short peptide chain and use the knowledge gained from previous chapters to complement the study of proline behaviour towards $NO_3^\bullet$,
- ✓ carry out product studies by exposing proline-containing short peptides to $NO_2^\bullet/O_3$ mixtures as the source of $NO_3^\bullet$ and identify the products formed, and
- ✓ study the reactions with proline-like model systems to further consolidate the findings.

5.3 Results and Discussion

5.3.1 Kinetic studies of the reaction of $NO_3^\bullet$ with proline and proline-containing short peptides

The kinetic studies of the reaction of $NO_3^\bullet$ with N- and C-protected proline and proline-containing short peptides were performed under similar experimental conditions using laser flash photolysis described in Section 3.3.2, page 50. The absolute second-order rate coefficients, *k*, are compiled in Table 5.3-1.

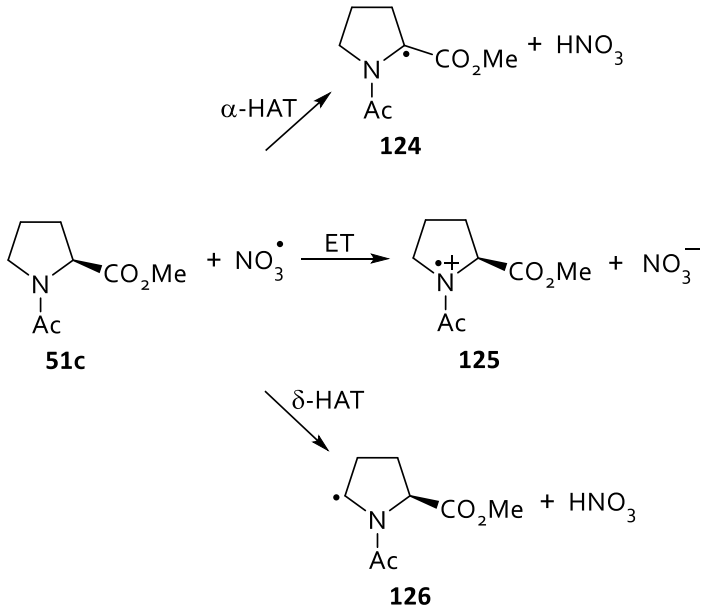
Table 5.3-1. Absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with proline and proline-containing di- and tripeptides in acetonitrile (298 ± 1 K).[a]

Entry	Substrate	k (10^8 M^{-1} s^{-1})
1	Ac-Pro-OMe (**51c**)	6.8
2	Ac-Gly-Pro-OMe (**52c**)	2.2
3	Ac-Val-Pro-OMe (**53c**)	8.3
4	Ac-Phe-Pro-OMe (**54c**)	1.7
5	Ac-Pro-Phe-OMe (**55c**)	7.2
6	Ac-Pro-Pro-OMe (**56c**)	16.7
7	Ac-Gly-Pro-Phe-OMe (**57c**)	3.3

[a] Experimental error ±5%. [Substrate]:[CAN] = 2–30.

Overall, the rate coefficient for the reaction of $NO_3^\bullet$ with the N- and C-protected proline **51c** was considerably higher than that for aliphatic amino acids [k = (1.4–3.3) x 10^6 M^{-1} s^{-1}, Table 3.3-1, page 54] and even higher than that for aromatic amino acids [k = (1.1–17.3) x 10^7 M^{-1} s^{-1}, Table 4.3-1, page 88 and Table 4.3-3, page 96]. This finding clearly shows that the cyclic system in proline significantly affects its reactivity.

It is possible that the reaction occurs at the amide nitrogen similar to the reaction of $NO_3^•$ with acyclic aliphatic amino acids through PCET (see Chapter 3). The main structural difference in proline compared to the other amino acids studied in this work is that the aliphatic side chain of proline is bound to the amino group to form a tertiary amide, which increases the electron density at the nitrogen atom of proline due to the inductive electron-donating effect from the side chain. Considering the oxidising nature of $NO_3^•$ [E^o ($NO_3^•$/ NO_3^-) = 2.3–2.5 V vs NHE],[14] it is likely that the higher rate coefficient is due to a direct ET process at the amide nitrogen to generate radical cation **125** (Scheme 5.3-1). Other possible pathways are: (i) α-HAT to yield radical **124** or (ii) δ-HAT to give radical **126**. The latter reaction has been found to be the preferred HAT pathway for the reaction of proline with oxygen-centred radicals, such as $HO^•$ or cumyloxyl ($PhC(CH_3)_2O^•$, $CumO^•$) radicals (see also Scheme 5.1-2, page 106).[47c,92b]



Scheme 5.3-1. Possible initial reaction pathways for the reaction of $NO_3^•$ with Ac-Pro-OMe (**5Ic**).

In order to identify the primary reaction pathway for proline, additional experiments involving proline-containing short peptides were performed. The kinetic data are included in Table 5.3-1. In di- and tripeptides consisting of proline

and aliphatic amino acid residues, the rate coefficient is essentially determined by the reaction at the proline moiety. Since the reactivity of Ac-Pro-OMe (**51c**) is more than two orders of magnitude higher than that of the other aliphatic amino acids, the rate coefficient determined for Ac-Val-Pro-OMe (**53c**) is within error the same as for Ac-Pro-OMe (**51c**; entry 3 vs 1). When two prolines were present in the peptide chain, such as in Ac-Pro-Pro-OMe (**56c**), the rate was found to be about twice as fast as with one proline (entry 6 vs 1).

Interestingly, a significant position dependence of the rate coefficient was found in dipeptides and tripeptides with proline and phenylalanine or glycine residues, respectively. The rate coefficient determined for the dipeptide with a C-terminal phenylalanine, Ac-Pro-Phe-OMe (**55c**), reflects the reactivity of proline alone (entry 5 vs 1). However, the rate coefficient for the dipeptide with the inverted sequence, Ac-Phe-Pro-OMe (**54c**), indicates a drop in reactivity by a factor of four (entry 4 vs 5), resembling more the rate coefficient determined for phenylalanine oxidation assisted by the neighbouring amide group (k = 1.6 x 10^8 M^{-1} s^{-1}, for the reaction of $NO_3^{\bullet}$ with Ac-Phe-NH_2; Table 4.3-2, page 91). From the kinetic studies alone, no plausible explanation for these findings can be given. Therefore, product studies were performed to provide insight into the possible reaction mechanisms, which will be outlined in the next section.

Furthermore, when glycine was located at the N-terminus of proline, such as that in Ac-Gly-Pro-OMe (**52c**) and Ac-Gly-Pro-Phe-OMe (**57c**), the rate coefficients decreased by a factor of 2–3 compared to Ac-Pro-OMe (**51c**; entries 2 and 7 vs 1). The presence of glycine at the N-terminus of proline residue seemingly reduces its reactivity. This retarding effect of glycine is not yet understood and could not be explored in the scope of this work. Thus, this finding will be further investigated in future studies.

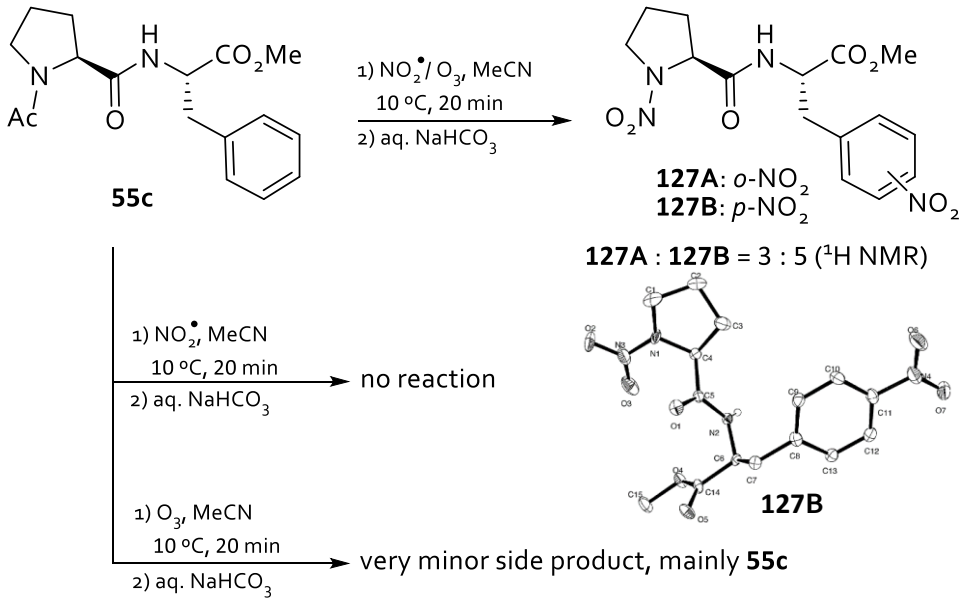
5.3.2 Product studies of the reaction of $NO_3^\bullet$ with proline-containing dipeptides

While kinetic studies reveal the reactivities of substrates towards the species under investigation and provide an indication of the initial reaction steps, it is through the combination with product studies that these methods become a powerful tool to obtain a detailed understanding of the reaction mechanism. Unfortunately, it was not possible to isolate products of the reaction of proline (either as a single amino acid or embedded in a di- or tripeptide) with $NO_3^\bullet$ obtained from CAN photolysis performed under preparative conditions (according to eq 3.1, page 47). The reaction resulted in formation of a large number of products, which could not be isolated and identified (not shown). Therefore, the product studies were performed with $NO_3^\bullet$ that was generated through the reaction of $NO_2^\bullet$ and O_3 (see eq 1.1, page 4)[11] to mimic the actual environmental exposure. It should be noted that generation of $NO_3^\bullet$ from $NO_2^\bullet/O_3$ mixtures can be advantageous, since under these conditions $NO_2^\bullet$ is usually present in excess, which could trap radical intermediates that would otherwise undergo uncontrolled degradation, therefore enabling isolation and identification of reaction products. This unique experimental setup may help to uncover the sites in proline-containing dipeptides most susceptible to the oxidative damage by $NO_3^\bullet$.

The product studies were performed at 10 °C through the addition of a measured excess amount of liquid $NO_2^\bullet$ (collected at –10 °C) to the solution of proline-containing dipeptides in acetonitrile, with a stream of ozonised oxygen passing through the solution.[78a] After 20 mins, the reaction mixtures were quenched by the addition of saturated aqueous sodium bicarbonate. When possible, the products were extracted and purified by column chromatography,

followed by HRMS and NMR analysis. Experimental details and spectroscopic data are given in Chapter 7, page 150.

Exposure of Ac-Pro-Phe-OMe (**55c**) to the $NO_2^{\bullet}/O_3$ mixtures resulted in complete consumption of the starting material and formation of two major products, which were identified as isomeric nitrated dipeptides **127A** and **127B** (Scheme 5.3-2), obtained in approximately a 3:5 ratio (according to the integration of the amide peaks in the ^1H NMR spectrum of the crude reaction mixture and HPLC analysis at λ = 220 nm; NMR spectra are shown in Figure 5.3-1). Control experiments with $NO_2^{\bullet}$ in the absence of O_3 showed no reaction with **55c**. In the reaction with O_3 and in the absence of $NO_2^{\bullet}$, HPLC analysis of the reaction mixture indicated unconsumed dipeptide **55c** as the major constituent with formation of a small amount of a side product, which was not identified.



Scheme 5.3-2. The reaction of $NO_2^{\bullet}$ and O_3 with Ac-Pro-Phe-OMe (**55c**).

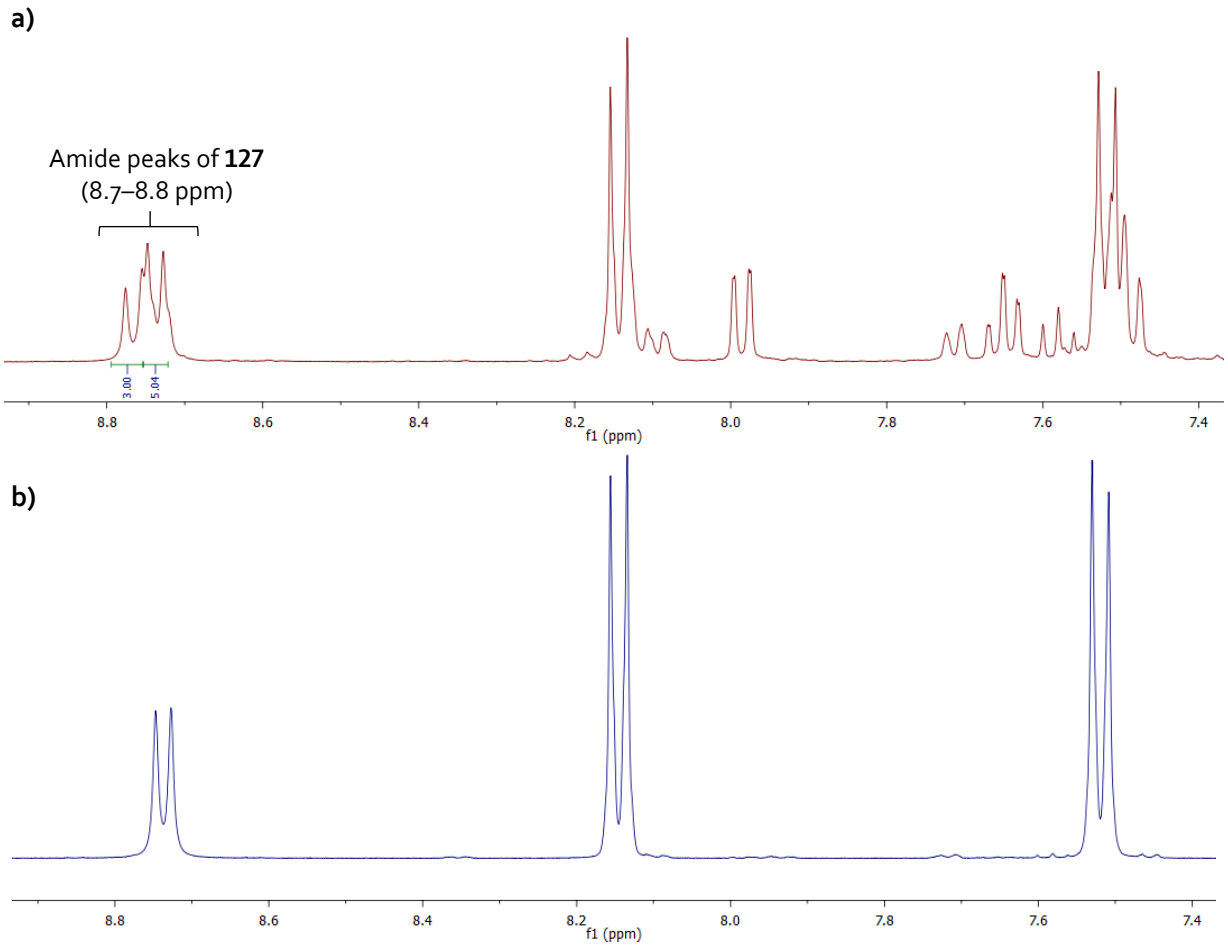
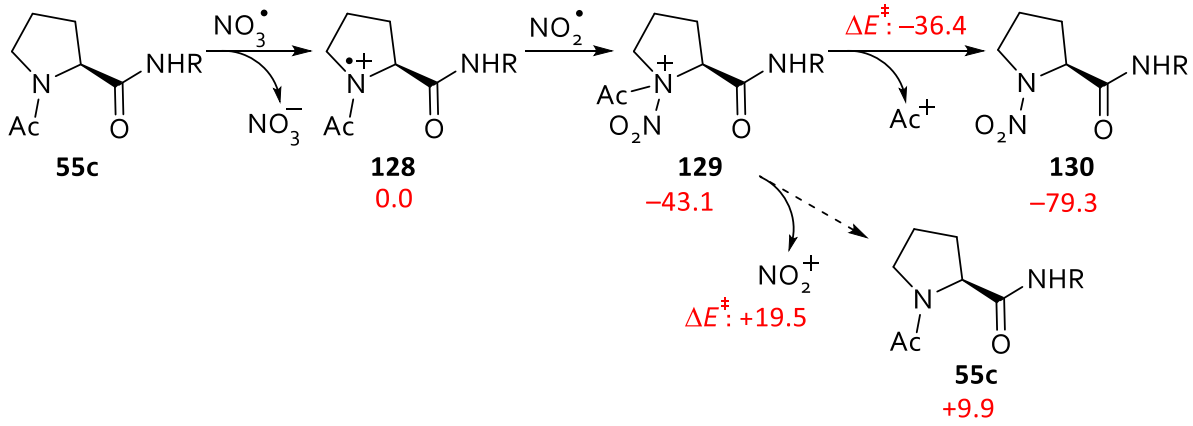


Figure 5.3-1. a) NMR spectrum of the crude reaction mixture, containing **127A** and **127B**; b) NMR spectrum of **127B** as a reference for the amide peaks in the crude reaction mixture. Peaks at 7.4–8.2 ppm belong to the aromatic protons.

The structure of product **127B** was identified by X-ray analysis (performed by Prof. Jonathan White), which showed that modifications happened at both proline and phenylalanine residues (Scheme 5.3-2). At the aromatic ring of phenylalanine, the reaction with $NO_3^{\bullet}$ takes place according to the previously discovered mechanism, which is through initial ET at the aromatic ring followed by radical trapping by $NO_2^{\bullet}$ and deprotonation (see Scheme 4.1-5, page 85). Meanwhile, the proline residue in **127B** shows replacement of the *N*-acetyl group at the proline moiety by a nitro (NO_2) substituent. Although the other isomer **127A** did not crystallise, HRMS analysis of this compound, which showed the same mass (*m*/*z* calcd. for $C_{15}H_{19}N_4O_7$ $[M + H]^+$: 367.1248, found: 367.1249; *m*/*z* calcd. for

$C_{15}H_{18}N_4O_7K$ $[M + K]^+$: 405.0807, found: 405.0809), and the very similar retention time from HPLC analysis (t_R = 10.741 mins for **127A** and 10.795 mins for **127B**) strongly suggest that **127A** should be the *ortho*-nitrosubstituted isomer of **127B**. The ratio of 3:5 is also in agreement with the ratio of *ortho*- and *para*-substituted isomers found when Ac-Phe-OMe (**24c**) was reacted with $NO_2^\bullet/O_3$ mixtures.[78a]

A mechanism for the reaction at the proline residue is proposed in Scheme 5.3-3, which consists of an initial electron transfer at the nitrogen of the proline to produce the radical cation **128**. This intermediate **128** is then trapped by $NO_2^\bullet$ to yield the cation **129**, which then loses the acetyl cation (Ac^+) through heterolytic cleavage, leading to the formation of the *N*-nitroproline product **130**.

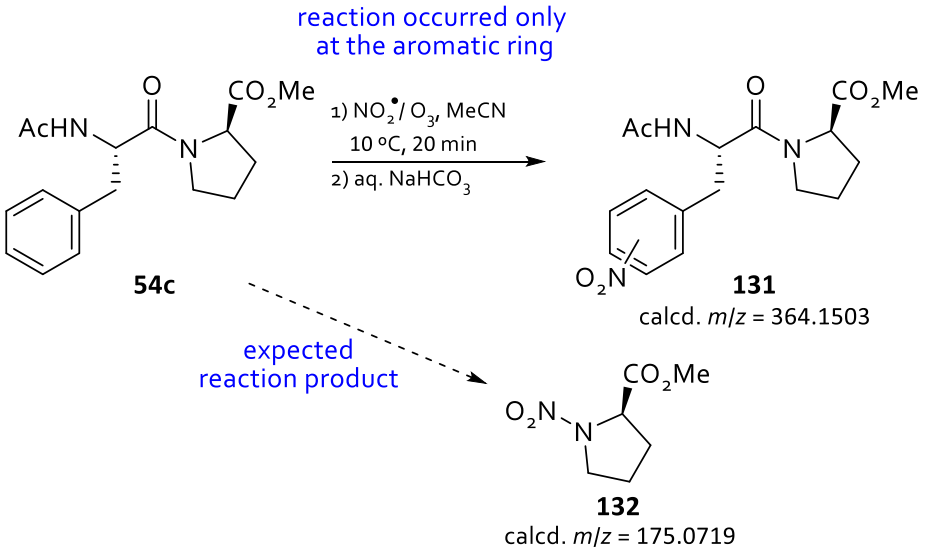


Scheme 5.3-3. Proposed mechanism with calculated energy profile (in red) for the reaction of $NO_3^\bullet$ with the proline residue at the N-terminus of dipeptide **55c** (R = Phe-OMe). M062X/6-31+G* free energies in kJ mol^{-1} in acetonitrile relative to **128**.

The transition state for radical-radical recombination of **128** and $NO_2^\bullet$ to form cation **129** could not be located, which suggests that such process is diffusion-controlled to generate a more stable adduct **129**. Computational studies (performed by Prof. Uta Wille) show that the loss of Ac^+ in adduct **129** happens with a low activation barrier (ca. 7 kJ mol^{-1}) and yields the *N*-nitroproline **130** exothermically. On the other hand, the loss of NO_2^+, which would restore **55c**, is unfavourable, having a much higher activation barrier of about 60 kJ mol^{-1}. The *N*-nitroproline product **130** highlights the amide nitrogen as the most reactive site of

proline residue towards $NO_3^\bullet$. Reaction through either α-HAT or δ-HAT would not give this product, clearly showing that ET is favoured over HAT in the reaction of $NO_3^\bullet$ with proline. It is likely that secondary reactions of $NO_3^\bullet$ with the *N*-nitroproline moiety through HAT at the α- or δ-carbon cannot occur due to the deactivating effect of the electron-withdrawing NO_2 group.

Interestingly, when the sequence in the dipeptide was inverted, no reaction at the proline residue was observed (Scheme 5.3-4). However, the reaction involving dipeptide **54c** was unfortunately not as clean as that of dipeptide **55c**. After 20 minutes of exposure to $NO_2^\bullet/O_3$ mixtures, incomplete consumption of starting material was observed. However, HRMS analysis showed formation of a new major product with m/z = 364.1502, which is about 45 mass units higher than **54c**. This finding suggests formation of the nitro-substituted phenylalanine **131**, which is supported by the isolation of a yellow oil upon the attempt of product purification. On the other hand, a reaction of $NO_3^\bullet$ at the proline residue should lead to the cleavage of peptide bond and yield the *N*-nitroproline **132** and water-soluble by-products (according to the mechanism in Scheme 5.3-3). This expected product **132** was not detected after HRMS analysis.



Scheme 5.3-4. Reaction of $NO_2^\bullet$ and O_3 with Ac-Phe-Pro-OMe (**54c**).

Since formation of the expected product **132** was not observed, the proline residue in **54c** appears to be relatively more inert towards ET than the aromatic ring. There was no indication of any other product (e.g., from δ-HAT), which probably because the product is water-soluble or undergoes decomposition during the workup. The finding from this product studies qualitatively confirms the low reactivity observed in the kinetic studies where the rate coefficient for the reaction of $NO_3^{\bullet}$ with Ac-Phe-Pro-OMe (**54c**) dropped by a factor of about four from that with the inverted sequence, Ac-Pro-Phe-OMe (**55c**; Table 5.3-1 entry 4 vs 5, page 111). The reactivity of Ac-Phe-Pro-OMe (**54c**; k = 1.7 x 10^8 M^{-1} s^{-1}) became similar to the reactivity of aromatic ET assisted by the neighbouring amide groups (k = 1.6 x 10^8 M^{-1} s^{-1} for Ac-Phe-NH_2 (**24q**)).

One possible explanation for this fundamental difference of the reaction outcome depending on the sequence of amino acid residues can be attributed to the amide neighbouring group effect discussed in Section 4.3.2, page 91–95, where the amide group of C-terminal proline leads to stabilisation of the developing positive charge on the aromatic ring thereby facilitating its ET (Figure 5.3-2).

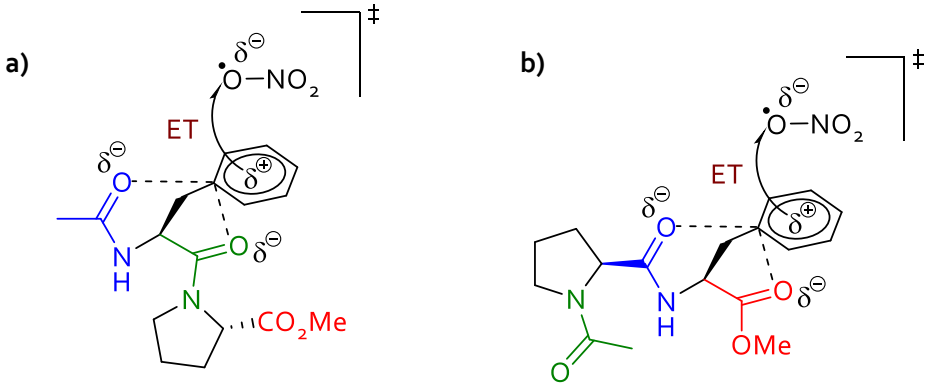


Figure 5.3-2. Amide neighbouring group effect in dipeptides **54c** and **55c**.

In both dipeptides Ac-Phe-Pro-OMe (**54c**) and Ac-Pro-Phe-OMe (**55c**), the aromatic ring in the phenylalanine residue receives stabilisation from the amide

neighbouring group effect, however, the effect is less pronounced in dipeptide **55c** because there is only one amide group in the vicinity of the aromatic ring. Moreover, it is a secondary amide (less nucleophilic) and would form a six-membered ring-like structure (amide in blue, Figure 5.3-2b), instead of a five-membered ring.

Meanwhile, in dipeptide **54c**, ET at the aromatic ring is facilitated by two amide groups (amides in blue and green, Figure 5.3-2a) flanking phenylalanine moiety, with one of them being a tertiary amide (more nucleophilic) and able to form a five-membered ring-like geometry (amide in green, Figure 5.3-2a). As a consequence, ET at the aromatic ring of **54c** is sped up due to these two neighbouring amide groups. However, the accelerating effect for the oxidation of the aromatic ring comes at a price. The assistance to the aromatic ring provided by the amide of the proline residue in turn decreases the electron density of the nitrogen of proline in dipeptide **54c** (Figure 5.3-3).

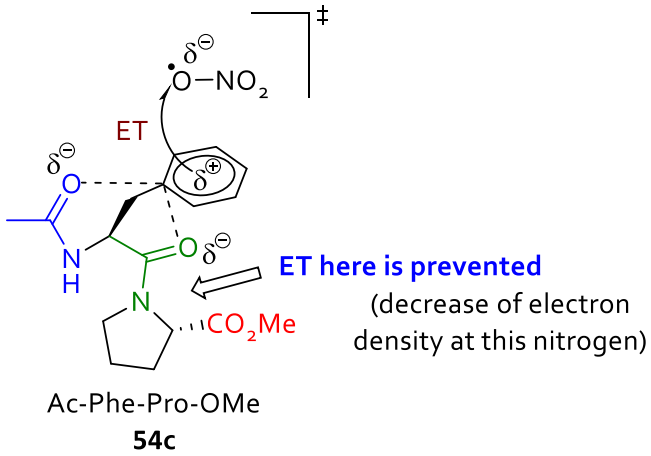


Figure 5.3-3. The amide neighbouring group effect lowers the electron density at the proline amide.

This decrease in electron density seems to prevent ET at this nitrogen, thus, the rate coefficient for the reaction of $NO_3^{\bullet}$ with dipeptide **54c** was found to be lower than that with dipeptide **55c** (Table 5.3-1 entry 4 vs 5, page 111). In the next section, non-proteinogenic amino acids are employed as the model compounds to

investigate the dependence of structural variations on ET at the tertiary amides. This would also include another example of the observed drop in reactivity similar to the one discussed in this section (Section 5.3.3.4, page 125).

5.3.3 Kinetic studies of the reaction of $NO_3^\bullet$ with proline-like model compounds

To gain more understanding of the particularly high reactivity of Ac-Pro-OMe (**51c**) compared with other aliphatic and aromatic amino acids, pyroglutamic acid (Glp) and *N*-methylglycine (sarcosine) were subjected to the kinetic studies based on some structural similarities. Figure 5.3-4 shows the four model systems used in this section and Ac-Pro-OMe (**51c**) for comparison. The structural similarities in the model compounds are colour-coded to aid visualisation.

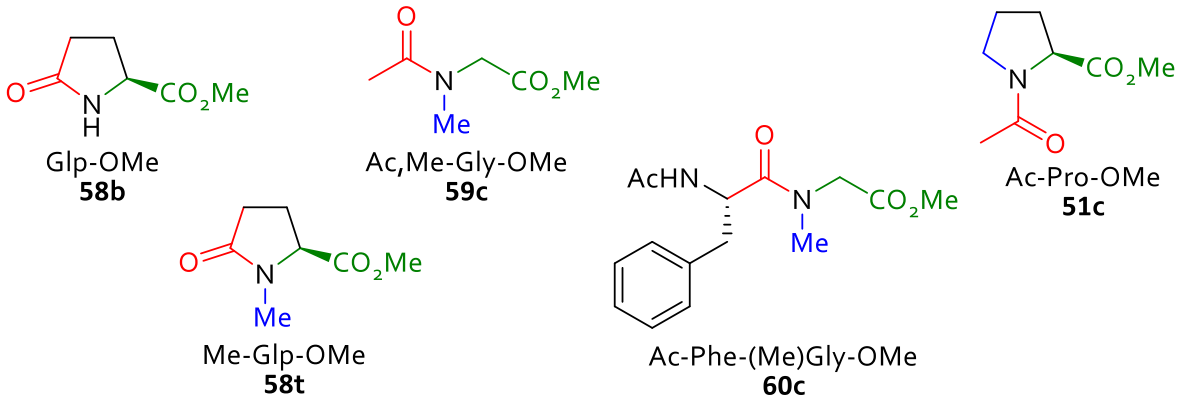


Figure 5.3-4. Model compounds used in this work and Ac-Pro-OMe (**51c**) for structural comparison. Red: carbonyl groups (EWG), blue: alkyl groups (EDG), green: α-carbons.

All derivatives were protected at the C-terminus as methyl esters. All model systems, except Glp-OMe (**58b**), have a tertiary amide moiety to mimic the proline amide. Their nitrogen atom each possesses one electron-withdrawing group (in red: carbonyl group), one alkyl group (in blue), and one α-carbon (in green). The absolute second-order rate coefficients, *k*, for the reaction of $NO_3^\bullet$ with the four model systems are compiled in Table 5.3-2. The influences of some structural variations are explored in this section.

Table 5.3-2. Absolute second-order rate coefficients, k, for the reaction of $NO_3^\bullet$ with pyroglutamic acid and sarcosine derivatives in acetonitrile (298 ± 1 K).[a]

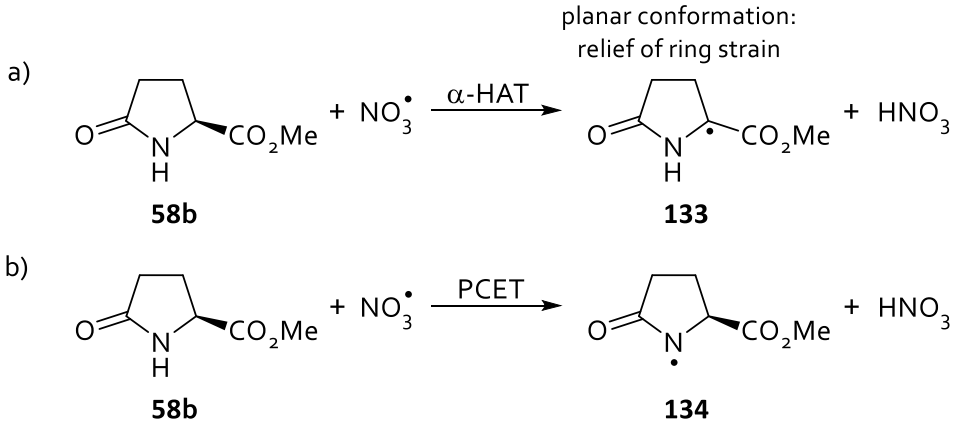
Entry	Substrate	k (10^8 M^{-1} s^{-1})
1	Glp-OMe (**58b**)	0.06
2	Me-Glp-OMe (**58t**)	2.0
3	Ac,Me-Gly-OMe (**59c**)	4.1
4	Ac-Phe-(Me)Gly-OMe (**60c**)	0.8

[a] Experimental error ±10%. [Substrate]:[CAN] = 3–50.

5.3.3.1 Influence between a secondary amide and a tertiary amide in a five-membered ring (entry 1 vs 2)

The lower reactivity of Glp-OMe (**58b**), which lacks the alkyl group at the nitrogen by more than one order of magnitude can be rationalised by the lower electron density on the nitrogen compared to the other model compounds, which are all tertiary amides. The rate coefficient obtained for Glp-OMe (**58b**) is of the same order of magnitude as that for the other acyclic aliphatic amino acids with secondary amide moieties [k = (1.4–3.3) x 10^6 M^{-1} s^{-1}, Table 3.3-1, page 54], although slightly faster by about a factor of 2–4. Easton and co-workers have previously reported based on the relative data for the reaction of $Br^\bullet$ with amino acid derivatives that Glp-OMe (**58b**) showed a higher reactivity by a factor of about three compared to the non-cyclic counterparts.[54b] It was proposed that α-HAT in **58b** is entropically favoured as it holds the amido group of **133** in a planar conformation, and thereby, relieving the ring strain and increasing the rate of the reaction (Scheme 5.3-5a). Since the rate enhancement found for the reaction of $NO_3^\bullet$ with Glp-OMe (**58b**) is about the same as that found by Easton and co-workers, it is highly likely that HAT at the α-carbon is the primary contributor to the higher rate coefficient obtained for Glp-OMe (**58b**) compared with the other

aliphatic amino acids. In addition to HAT from the α-carbon, $NO_3^{\bullet}$ can also react with Glp-OMe (**58b**) at the secondary amide through PCET as previously discussed in Chapter 3 (Scheme 5.3-5b). However, there seems to be no palpable difference between PCET in the cyclic and linear system.



Scheme 5.3-5. a) HAT at the α-carbon of Glp **58b**; b) PCET at the amide nitrogen of Glp **58b**.

The introduction of an *N*-methyl group in **58b** increases the electron density at the amide nitrogen, which in turn leads to an increase of the reactivity of Me-Glp-OMe (**58t**) towards $NO_3^{\bullet}$ by a factor of 30 compared to that of Glp-OMe (**58b**; entry 1 vs 2). This behaviour can be rationalised by a considerably faster ET at this relatively more electron-rich nitrogen in Me-Glp-OMe (**58t**), which signifies the mechanistic difference between secondary amides and tertiary amides. Secondary amides react with $NO_3^{\bullet}$ through PCET at the amide nitrogen and HAT from other sites in the amino acid (in this case from the α-carbon), while tertiary amides react mainly through ET at the electron-rich amide nitrogen (Figure 5.3-5).

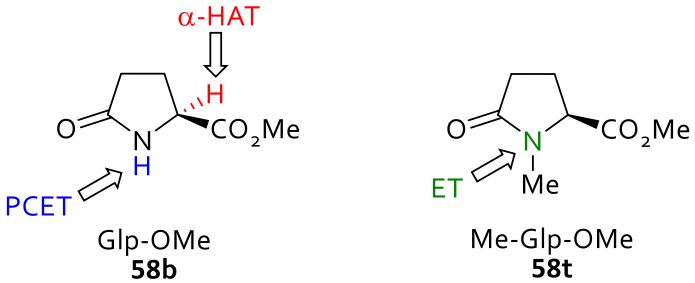


Figure 5.3-5. Reactive sites for the reactions of $NO_3^{\bullet}$ with amino acids **58b** and **58t**.

5.3.3.2 Influence between a cyclic and a linear system on ET at the tertiary amide (entry 2 vs 3)

Both Me-Glp-OMe (**58t**) and Ac,Me-Gly-OMe (**59c**) possess a tertiary amide moiety that contains a methyl group attached to the nitrogen atom (Figure 5.3-4). Ac,Me-Gly-OMe (**59c**) serves as a linear model for the tertiary amide bond in Me-Glp-OMe (**58t**). The rate coefficients determined for Me-Glp-OMe (**58t**) and Ac,Me-Gly-OMe (**59c**) were in the same order of magnitude (entry 2 vs 3), which is likely due to ET at the electron-rich nitrogen. The small discrepancy between the rate coefficients may arise from the number of abstractable α-hydrogens (one available hydrogen in Me-Glp-OMe and two available hydrogens in Ac,Me-Gly-OMe; see Figure 5.3-6). Another possibility is that ET rate for a cyclic amide differs from that for a linear amide. The structure of a cyclic amide is more rigid and gives less room for any possible conformation changes compared to that of a linear one.

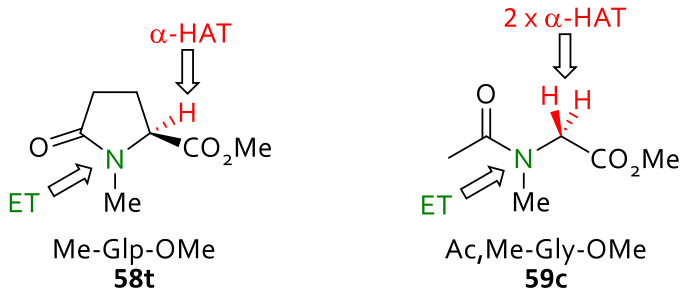
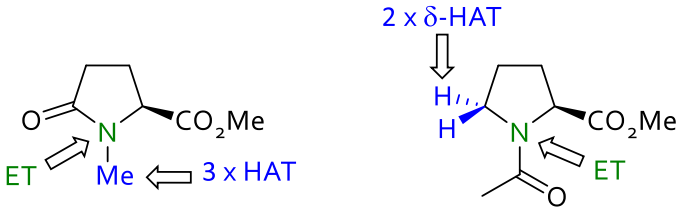


Figure 5.3-6. Reactive sites for the reactions of $NO_3^{\bullet}$ with amino acids **58t** and **59c**.

5.3.3.3 Influence of the nature of the alkyl group attached to the tertiary amide on the reactivity

Both Me-Glp-OMe (**58t**) and Ac-Pro-OMe (**51c**) have the same five-membered ring structure with different nature of the alkyl group attached to the tertiary amide. The alkyl group in Me-Glp-OMe (**58t**) is a primary group, whereas Ac-Pro-OMe (**51c**) possesses a secondary group embedded in the ring structure.

The same order of magnitude of the rate coefficients (compare Table 5.3-2 entry 2, page 122 with Table 5.3-1 entry 1, page 111) indicates that ET at the tertiary amide is the major contributor to the reactivity. The slight variation in rate coefficients might suggest that there is a higher contribution from a secondary alkyl group than a primary one (Figure 5.3-7). The hydrogens at δ-position are more vulnerable than hydrogens at the *N*-methyl position. This high reactivity of hydrogens of a secondary carbon adjacent to nitrogen has been shown in Ac-Lys(NHAc)-OMe (**8c**') in Section 3.3.3.3, page 67–69.



Me-Glp-OMe **58t** Ac-Pro-OMe **51c**

Figure 5.3-7. Reactive sites for the reactions of $NO_3^{\bullet}$ with amino acids **58t** and **51c**.

5.3.3.4 Influence of amide neighbouring group effects on the reactivity of the tertiary amide (entry 3 vs 4)

The incorporation of sarcosine residue in a peptide chain provides a non-cyclic mimic of Ac-Phe-Pro-OMe (**54c**) to probe if the decrease in reactivity would also be observed in this model system. Indeed, the rate coefficient of Ac-Phe-(Me)Gly-OMe (**60c**) decreased by a factor of 5 from Ac,Me-Gly-OMe (**59c**; entry 3 vs 4). This rate decrement might suggest the participation of the amide group of sarcosine in the neighbouring group stabilisation of the developing aromatic radical cation, which lowers the electron density at the nitrogen of sarcosine and prevents ET at that nitrogen (amide in green, Figure 5.3-8). The hypothesis is further supported by the fact that the rate coefficient determined for Ac-Phe-(Me)Gly-OMe (**60c**) was the same as Ac-Phe-NHMe (**24s**; k = 0.8 x 10^8 M^{-1} s^{-1}),

where the reaction proceeds through ET at the aromatic ring with the assistance of the neighbouring amide group. This finding also serves as an additional example for the similar drop of the rate coefficient (by a factor of 4) in Ac-Phe-Pro-OMe (**54c**) compared to that in Ac-Pro-OMe (**51c**; Table 5.3-1 entry 1 vs 4, page 111).

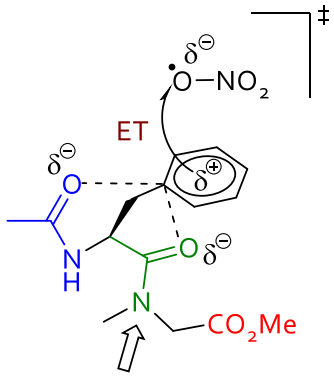


Figure 5.3-8. Participation of the amide groups of **60c** to assist ET at the aromatic ring.

5.4 Summary of findings

This chapter focuses on the reactivity of proline and its derivatives towards $NO_3^{\bullet}$. Electron transfer has been proposed to be the main pathway through which the reaction with $NO_3^{\bullet}$ takes place. Product studies were carried out to supplement the findings from the kinetic studies. Derivatives of other amino acids were employed as model compounds based on structural similarities with proline to further examine the proposed reaction pathway.

The key findings of this chapter are:

- ✓ Ac-Pro-OMe (**51c**) shows much higher reactivity, by about two orders of magnitude higher, than aliphatic acyclic amino acids. This immense rate enhancement is taken as an indication of a different reaction mechanism.

Electron transfer at the electron-rich nitrogen of the proline derivative is the main pathway as opposed to HAT or PCET for the reaction of $NO_3^{\bullet}$ with other aliphatic amino acids (Chapter 3).

- ✓ The reactivity of proline was observed to be dominating when proline is incorporated into a peptide chain. However, when glycine is present at the N-terminus of proline residue or if the amide of proline contributes to the stabilisation of the developing positive charge on the neighbouring aromatic ring following ET, rate retardation of proline residue was observed.
- ✓ This proposed 'deactivating' neighbouring group effect in proline was supported in the product studies where no reaction at the proline residue was detected after the exposure of Ac-Phe-Pro-OMe (**54c**) to $NO_3^{\bullet}$. The electron density at the proline nitrogen is decreased due to its role in assisting ET at the aromatic ring. Consequently, only the reaction at the aromatic ring was observed. On the other hand, in the dipeptide with an inverse sequence, Ac-Pro-Phe-OMe (**55c**), the reactions took place at both proline and phenylalanine residues through an oxidation/ET mechanism.
- ✓ N-and C-protected pyroglutamic acid (Glp) and sarcosine were employed as model compounds that possess structural resemblance to Ac-Pro-OMe (**51c**) and subjected to the laser flash photolysis experiments. The same order of magnitude for their rate coefficients strongly suggest that ET indeed occurs at the nitrogen atom with some possible minor reactions at different sites, hence, slight discrepancies in their values were obtained.

Chapter Six: Reaction of nitrogen dioxide with biological molecules

6.1 Introduction

Apart from $NO_3^\bullet$, nitrogen dioxide ($NO_2^\bullet$) is one of the most important components of air pollution. It is produced in large quantities (up to 0.08 ppm in large cities with heavy traffics) through combustion processes in automotive engines and motor vehicles.[103] Besides that, $NO_2^\bullet$ is present in unvented heating sources, tobacco smoke, and smoke from other burning organic materials.[104]

A large number of exposure studies have been performed that showed what $NO_2^\bullet$ can do to living organisms, ranging from plants over animals to humans.[105–108] For example, continuous exposure of pinto bean seedlings and tomato seedlings to low concentrations of $NO_2^\bullet$ significantly suppresses the growth of the plants and distorts the leaves to curve downward, which hinders access to sunlight and consequently reduces photosynthetic abilities.[105a] Carotenoids have been identified as an important group of compounds produced by plants to protect against oxidative damage.[105b,109] They are known to reduce $NO_2^\bullet$ to the less reactive $NO^\bullet$ in the presence of light.[105b] Therefore, plants are generally more susceptible to $NO_2^\bullet$ during nighttime when the concentration of $NO_2^\bullet$ is higher than daytime due to the absence of $NO_2^\bullet$ photochemical decomposition.[105a]

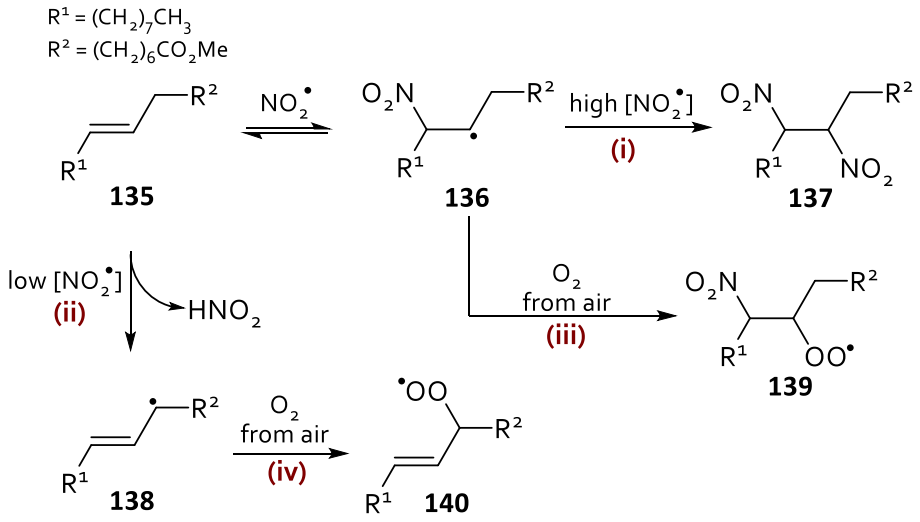
A large number of studies have shown that $NO_2^\bullet$ is responsible for a range of lung respiratory diseases.[104,106b,110–112] $NO_2^\bullet$ can induce bronchoconstriction and branchial hyperreactivity; cause pulmonary edema and pulmonary fibrosis; and trigger extracellular damage in the airways.[110,111] Experiments with cell cultures revealed an increase in epithelial cell membrane damage and permeability, a reduction of ciliary activity and depletion of protective antioxidants.[111a,d]

Furthermore, $NO_2^{\bullet}$ and SO_2 synergistically enhance the airway response to inhaled allergens in asthmatic patients.[111b]

Despite the considerable number of adverse health effects caused by $NO_2^{\bullet}$, there is still a lack of understanding of the detailed mechanism on how exposure to $NO_2^{\bullet}$ leads to these health complications. Several studies have used different model systems to further comprehend the chemical pathway that leads to oxidative damage of biological systems upon exposure to $NO_2^{\bullet}$.[111c,113] In animals and humans, lungs are the primary target for inhaled $NO_2^{\bullet}$ due to their large surface area and their role in gas exchange.[108,110] Pulmonary damage has been found to be the major adverse effect upon the exposure of animal lung cells to $NO_2^{\bullet}$, involving some damage in the bronchioles and alveolar epithelial cells.[106b,c] DNA damage was also demonstrated in mice lungs through an increase in single-strand breaks after exposure.[108] This finding highlights $NO_2^{\bullet}$ solubility in lipids as it can be transported through cell membranes to access DNA. Lipid peroxidation is considered to be an important pathway, which then subsequently leads to cell injury and death.[107] The presence of protective enzymes in the lungs, such as glutathione peroxidase, prevents cell damage by lipid peroxides.[107a] However, the activities of this system are shown to gradually decrease after a 16-week exposure to 4.0 ppm $NO_2^{\bullet}$.[107b]

A work by Pryor et al. suggested that $NO_2^{\bullet}$ reacts with double bonds in unsaturated lipids through a free radical mechanism, which was found to depend on the concentration of $NO_2^{\bullet}$ the lipids are exposed to (Scheme 6.1-1).[113a] Product studies showed that $NO_2^{\bullet}$ reversibly adds to the double bonds of the lipid **135** to form a carbon-centred radical **136**. At high $NO_2^{\bullet}$ concentrations, intermediate **136** is immediately trapped by $NO_2^{\bullet}$ to give the addition product **137** (pathway i). On

the other hand, at low $NO_2^\bullet$ concentrations, the rate of radical trapping slows down and the intermediate **136** can revert to reform the lipid **135**.



Scheme 6.1-1. Different reaction pathways for the reaction of $NO_2^\bullet$ with double bonds in the unsaturated lipid **135** depending on the concentration of $NO_2^\bullet$.[113a]

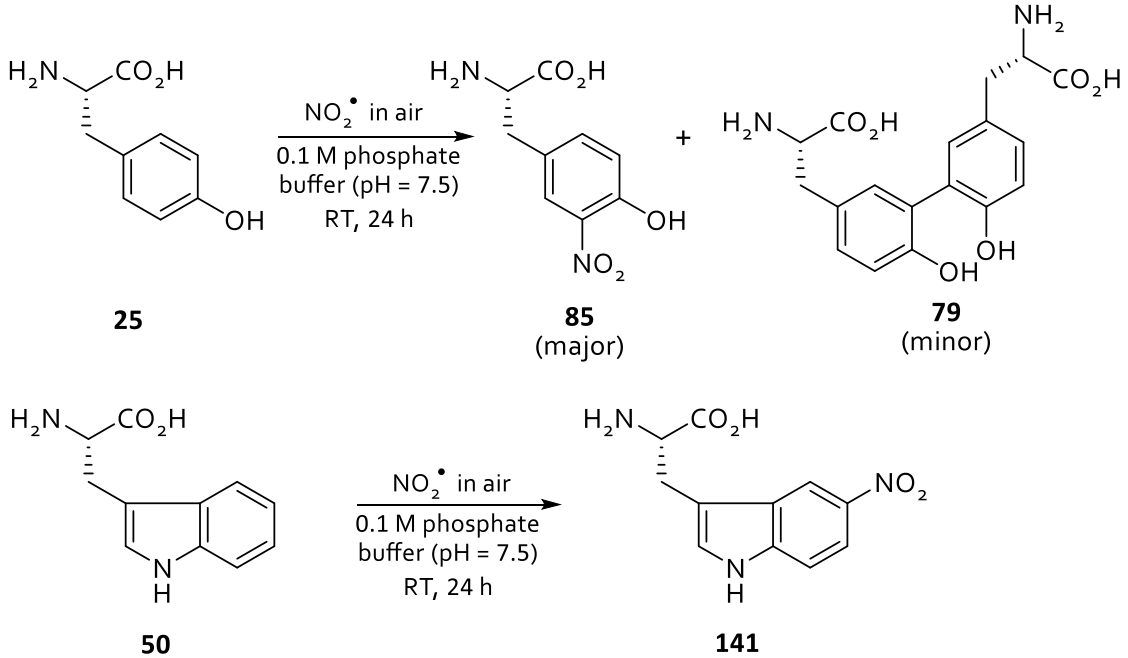
Alternatively, lipid **135** can undergo HAT at its allylic position (pathway ii), which leads to the allyl radical **138** and nitrous acid (HNO_2). This process is irreversible because HNO_2 readily decomposes into gaseous nitrogen oxides ($NO_2^\bullet$ and $NO^\bullet$) and water according to eq. 6.1.

$$2HNO_2 \rightarrow NO_2^\bullet + NO^\bullet + H_2O \qquad \text{(eq. 6.1)}$$

When the radical intermediates **136** and **138** are trapped by O_2 instead, reactive peroxyl radicals **139** and **140** are generated (pathways iii and iv), which initiate the radical chain mechanism of lipid peroxidation.[113]

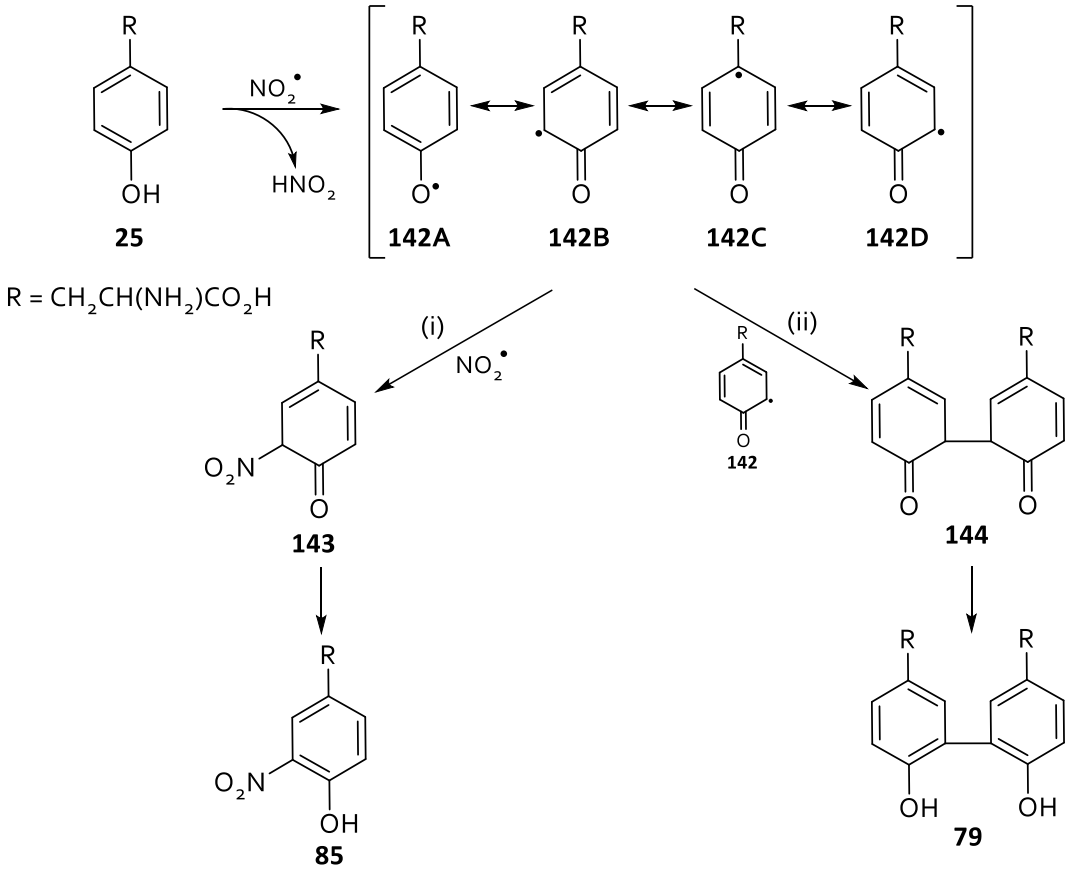
Apart from lipids, proteins have been shown to be damaged by $NO_2^\bullet$. Exposure of various proteins, such as bovine serum albumin (BSA), γ-globulin and eye lens protein α-crystallin, to $NO_2^\bullet$ in air demonstrated that tyrosine and tryptophan residues are the main target of damage.[111c] Kikugawa and co-workers performed exposure studies by passing gaseous $NO_2^\bullet$ (40–90 ppm in air) through a buffered (pH = 7.5) aqueous solution of tyrosine (**25**) or tryptophan (**50**) at room

temperature for 24 hours (Scheme 6.1-2).[111c] Upon exposure, 3-nitrotyrosine **85** and dityrosine **79** were obtained from the reaction involving tyrosine (**25**), while 5-nitroindole derivative **141** was found to be the major product in the reaction of $NO_2^\bullet$ with tryptophan (**50**). In a separate exposure experiment under the same reaction conditions, dityrosine **79** was also found to be degraded by $NO_2^\bullet$ into unidentified products. Therefore, the formation of dityrosine **79** as a minor product can be rationalised by its further degradation. The mechanism of these reactions was not explored, but HAT was proposed as the main pathway due to the free-radical nature of $NO_2^\bullet$.



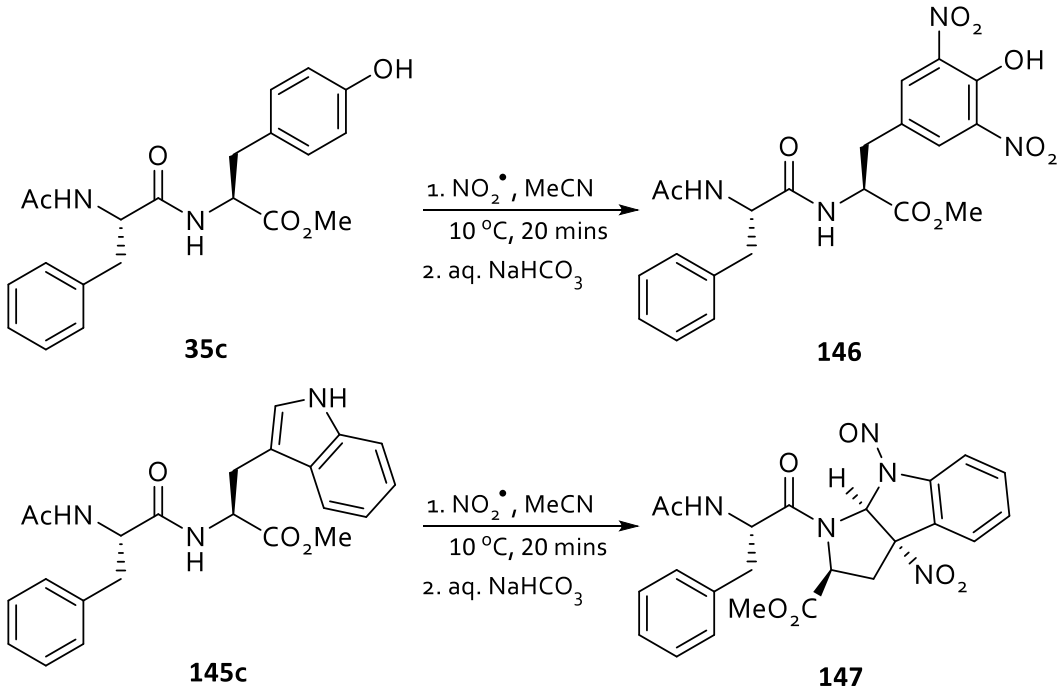
Scheme 6.1-2. Reaction of $NO_2^\bullet$ with aromatic amino acids **25** and **50**.[111c]

Some other studies also proposed HAT as the initial step of the reaction of $NO_2^\bullet$ with aromatic amino acids **25** and **50**.[113c,e] For example, Pfeiffer et al. proposed that HAT from the phenolic hydrogen of **25** would generate the phenolic radical **142A** (Scheme 6.1-3). This resonance-stabilised tyrosyl radical **142A–D** can then either be trapped by $NO_2^\bullet$ to produce adduct **143** (pathway i) or dimerise to afford dimer **144** (pathway ii). Tautomerisation of **143** and **144** to rearomatise the ring structure in both cases yields nitrotyrosine **85** and dityrosine **79**, respectively.



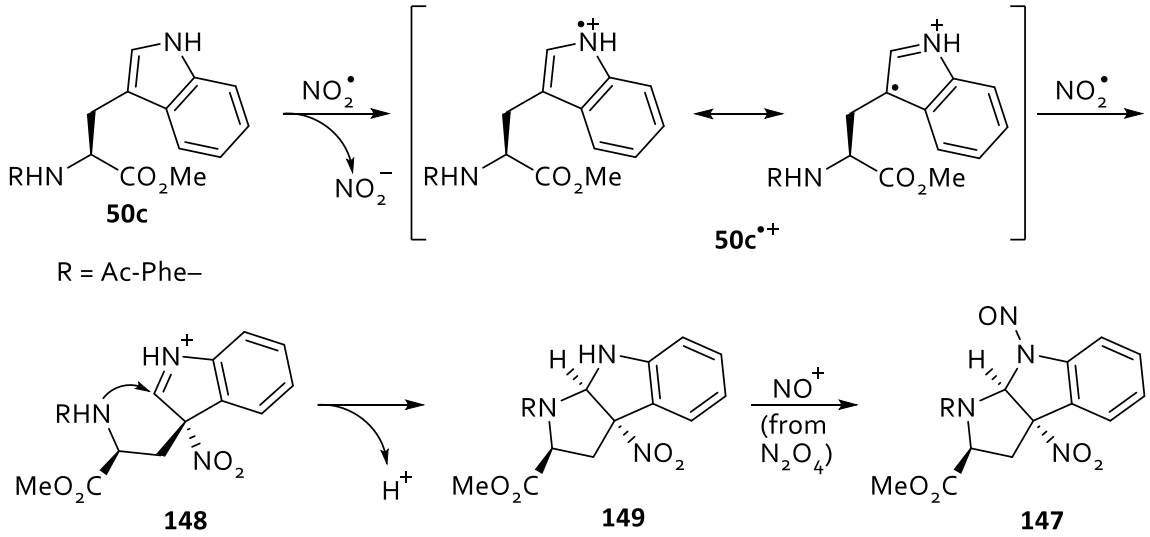
Scheme 6.1-3. Proposed mechanism for the reaction of $NO_2^•$ with tyrosine (**25**).[113c]

The moderate oxidising nature of $NO_2^•$ [E^o ($NO_2^•$/ NO_2^-) = 0.9 V][5] enables oxidative electron transfer as an alternative reaction pathway with the most reactive amino acids. The reactions of $NO_2^•$ were further explored by Gamon et al. through the product studies of dipeptides Ac-Phe-Tyr-OMe (**35c**) and Ac-Phe-Trp-OMe (**145c**) in acetonitrile.[25] Even though Chapter 4 has shown that a neighbouring amide group increases the oxidation rate of N-terminal phenylalanine by $NO_3^•$, phenylalanine residue in dipeptides **35c** and **145c** was found to be completely inert towards $NO_2^•$. Modification of dipeptides **35c** and **145c** only takes place at the tyrosine and tryptophan residues, respectively (Scheme 6.1-4). The reaction at the tyrosine residue shows a similar nitration product as previously found in the literature (Scheme 6.1-2).[111c] The higher degree of nitration in **146** was rationalised by the considerably higher amount of $NO_2^•$ used in this work (0.5 mL of liquid $NO_2^•$ injected into the reaction mixture).[25]



Scheme 6.1-4. Reaction of $NO_2^{\bullet}$ with aromatic-containing dipeptides **35c** and **145c** in acetonitrile.[25]

A new product **147** was isolated from the reaction of $NO_2^{\bullet}$ with dipeptide **145c** (multiple products were originally obtained but only one was isolated, the others are likely diastereomers of **147**). The initial step for the formation of the pyrroloindoline **147** was proposed to be an oxidative ET at the indole nitrogen to produce the radical cation **50c**$^{\bullet+}$ (Scheme 6.1-5).

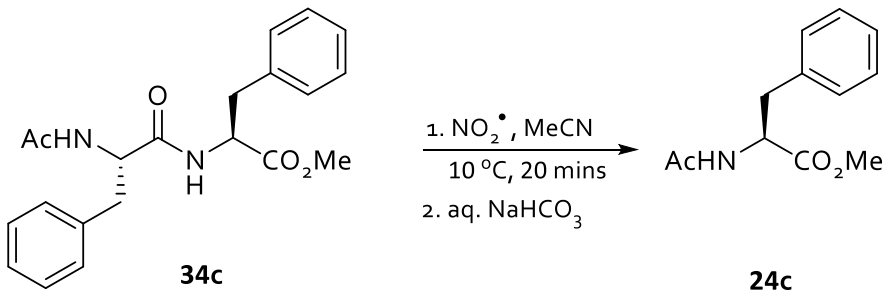


Scheme 6.1-5. Proposed mechanism for the reaction of $NO_2^{\bullet}$ with dipeptide **50c** in acetonitrile.[25]

The intermediate **50c**$^{\bullet+}$ is immediately trapped by $NO_2^{\bullet}$ from below the molecular plane of **50c**$^{\bullet+}$ to generate cation **148**, followed by a subsequent nucleophilic cyclisation from the neighbouring amide nitrogen and deprotonation to yield bicyclic **149**. The final step involves *N*-nitrosation of the free amine in **149** by NO^+, which originates from the ionic form of N_2O_4 (dimer of $NO_2^{\bullet}$; eq. 6.2), to yield the *N*-nitroso pyrroloindoline **147**.[114] The equilibrium constant of $NO_2^{\bullet}$ dimerisation in acetonitrile is 3.3×10^4 L mol^{-1} at room temperature.[114b] It is an exothermic reaction, where at higher temperature, more $NO_2^{\bullet}$ would be produced.

$$2NO_2^{\bullet} \rightleftarrows N_2O_4 \rightleftarrows NO^+ + NO_3^- \qquad \text{(eq. 6.2)}$$

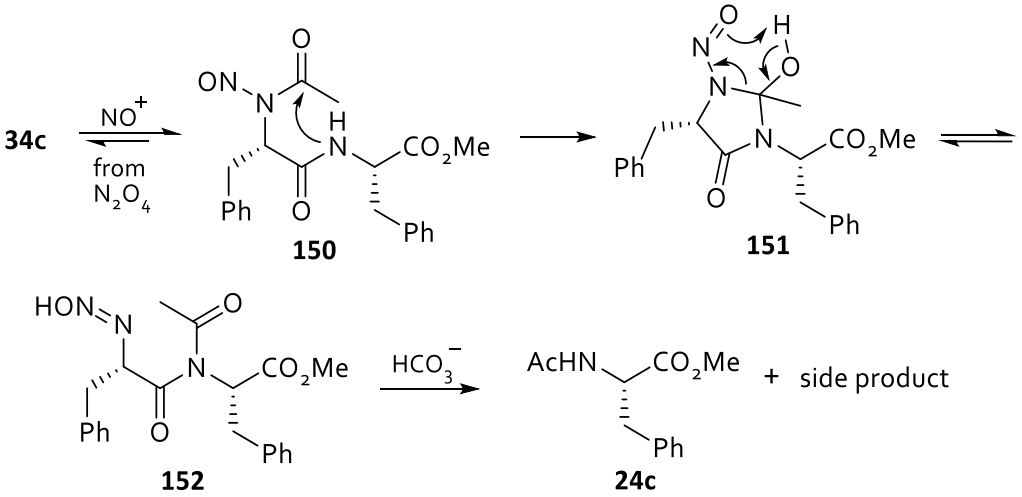
Apart from protein modification at side chains, such as in tyrosine and tryptophan residues, little is known about the reaction of $NO_2^{\bullet}$ with the peptide backbone itself. It has been shown that poly-L-lysine and poly-L-arginine undergo degradation to smaller peptides upon $NO_2^{\bullet}$ exposure.[110] The peptide bond scission could be initiated by reactions of $NO_2^{\bullet}$ with amide nitrogens exposed along the peptide backbone. Following this finding, it was recently discovered that in the absence of reactive side chains, such as that in tyrosine or tryptophan, exposure of aromatic dipeptide **34c** to $NO_2^{\bullet}$ led to backbone fragmentation through cleavage of the peptide bond, yielding amino acid **24c** as the only product (Scheme 6.1-6).[25]



34c **24c**

Scheme 6.1-6. Reaction of $NO_2^{\bullet}$ with Ac-Phe-Phe-OMe (**34c**).[25]

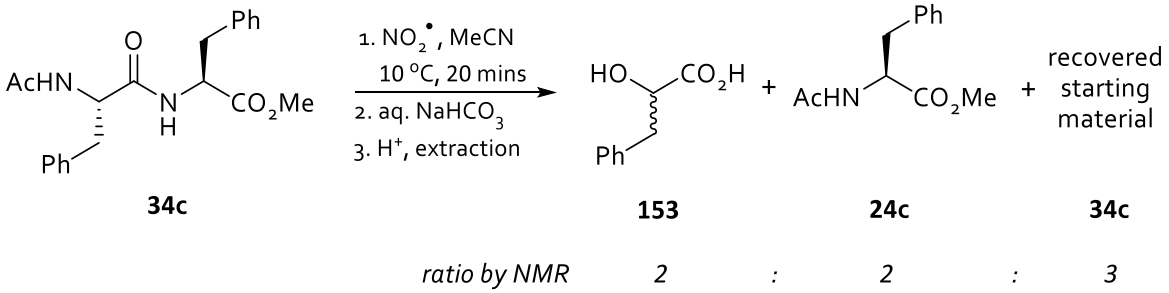
The mechanism for this fragmentation, which is proposed in Scheme 6.1-7, involves initial *N*-nitrosation of the *N*-acetyl terminus of dipeptide **34c** by NO^+, originating from the ionic form ($NO^+NO_3^-$) of N_2O_4, which exists in equilibrium with $NO_2^\bullet$ according to eq. 6.2.[114] This nitrosation yields *N*-nitrosamide **150**, which now has an activated adjacent acyl moiety for nucleophilic attack due to the electron-withdrawing effect of the *N*-nitroso group.[78a] This activated acyl group enables a nucleophilic, intramolecular cyclisation of the peptide backbone to give the five-membered cyclic intermediate **151**.



Scheme 6.1-7. Proposed mechanism for the $NO_2^\bullet$-driven fragmentation-rearrangement reaction of dipeptide **34c** to afford amino acid **24c**.[25]

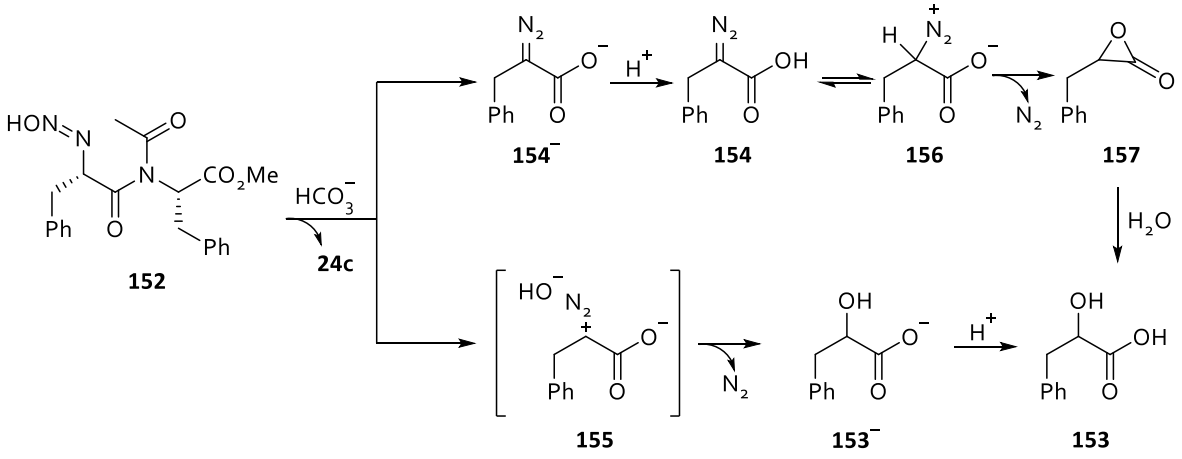
Rearrangement of the intermediate **151** results in the diazotic acid **152** through a concerted hydrogen transfer/fragmentation sequence.[115] Neutralisation of the reaction mixtures under aqueous environment can drive regioselective hydrolysis of the imide moiety present in the diazotic acid **152** to give the N- and C-protected amino acid **24c** and a water-soluble side product. Upon the modification of the work-up procedure, which is now through neutralisation and acidification of the reaction mixture, followed by extraction with ethyl acetate, 2-hydroxy-3-phenylpropionic acid (**153**) was also obtained along with the amino acid **24c** (ratio of **153**:**24c** is 1:1 by NMR) and starting material **34c** (Scheme 6.1-8).[116] This

α-hydroxy carboxylic acid **153** was identified as the side product, which is water-soluble prior to acidification of the reaction mixture (**153** exists in its ionic form under basic conditions).



Scheme 6.1-8. Modified workup conditions to identify the side product resulting from the fragmentation-rearrangement reaction of $NO_2^•$ with phenylalanine dipeptide **34c**. Stereochemistry is not determined.[116]

Scheme 6.1-9 shows how hydrolysis of the diazotic acid **152** can lead to the formation of the side product **153**. The reaction could give either the diazo **154**$^-$ or the nitrogen-separated ion **155** and liberate the N- and C-protected amino acid **24c**.[116]



Scheme 6.1-9. Mechanism for the formation of the side product **153**.[116]

The α-diazo acid **154** would remain in its anionic form in the aqueous phase until the mixture is acidified. Once protonated, the diazo acid **154** tautomerises to form the zwitterionic diazonium carboxylate **156**. Cyclisation of this intermediate **156** releases N_2 and affords a strained 3-membered lactone **157**. The lactone **157** is prone

to hydrolysis and, hence, undergoes a rapid ring opening under aqueous conditions to yield the α-hydroxy acid **153**.

Alternatively, if the reaction proceeds via the nitrogen-separated ion **155**, a concerted process would take place through an intramolecular collapse of the intermediate **155**.[117] The ion pair releases N_2 to yield the α-hydroxy acid **153** upon the protonation of the reaction mixture.

This finding shows an alternative pathway for the reaction of $NO_2^{\bullet}$ with some organic molecules besides the commonly known free radical mechanism, such as addition to the double bond or HAT. In this pathway, $NO_2^{\bullet}$ reacts as an ionic species through NO^+, which is formed upon the dissociation of the dimer N_2O_4.

6.2 Aim

The air pollutant $NO_2^\bullet$ has been shown to interact with different biological molecules in living organisms through both radical and ionic pathways. In most cases, radical pathways are widely accepted, especially for the reaction of $NO_2^\bullet$ with reactive species, such as tyrosine and tryptophan side chains. However, ionic pathways have not received much attention, as it was only recently proposed for the reactions of $NO_2^\bullet$ with the peptide backbone. This work envisages to find more examples of such reactions to give further support for the ionic pathway.

Specifically, the aims of this chapter were to:

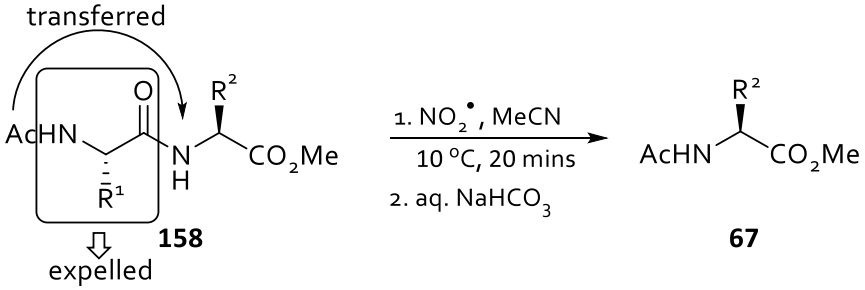
- ✓ perform product studies for the reactions of $NO_2^\bullet$ with dipeptides possessing 'non-reactive' side chains in acetonitrile, where there is no solvent interference with the reaction intermediates,
- ✓ vary the side chains of the N- and C-protected dipeptides to investigate if the backbone fragmentation found previously (see Scheme 6.1-6) would be affected by different steric demands,
- ✓ evaluate the reaction outcomes involving tetrapeptides with different steric hindrance, and
- ✓ explore $NO_2^\bullet$-induced oxidative damage in biomolecules other than amino acids and peptides, such as cholesterol and its derivatives.

The work in this chapter is a part of collaborative works with Dr. Luke Gamon (for peptides) and Anton Zalewski (for cholesterol).

6.3 Results and Discussion

6.3.1 Reaction of $NO_2^{\bullet}$ with dipeptides

Overall reactions of $NO_2^{\bullet}$ with dipeptides **158** resulted in the shortening of the peptide chain by one amino acid moiety. This process was accompanied with simultaneous acyl migration in the N→C direction, yielding N- and C-protected amino acid **67** as the major product (Scheme 6.3-1).



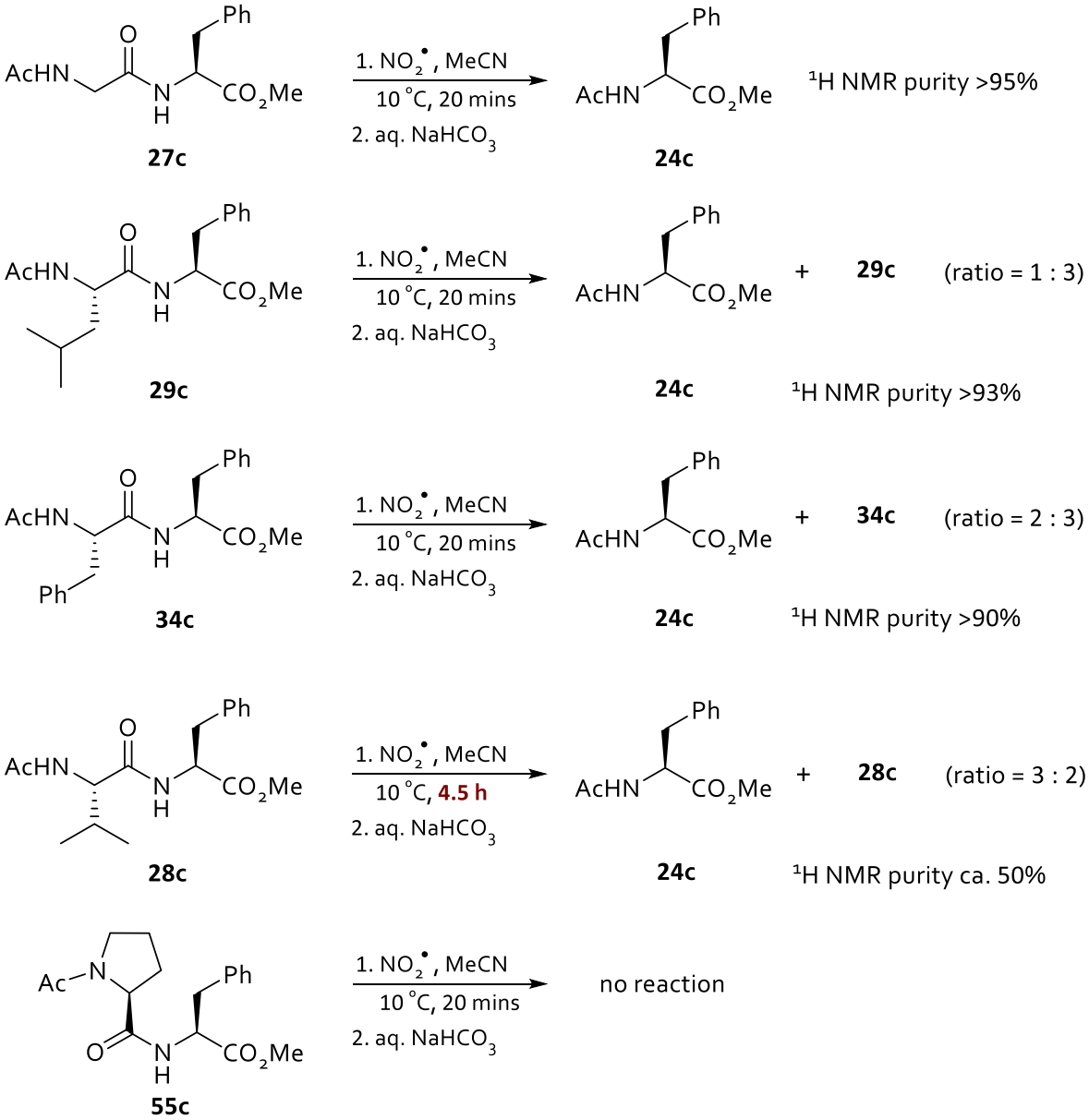
Scheme 6.3-1. Degradation pathway of dipeptide **158** upon exposure to $NO_2^{\bullet}$.

The effect of steric hindrance was investigated by varying the side chain of the N-terminal amino acid residue (i.e., various R^1 in dipeptide **158** was used). To ease the product detection by diode array HPLC, phenylalanine was installed at the C-terminal residue in dipeptide **158** (R^2 = Bn).

The reaction time was kept at 10 °C for 20 minutes to consistently compare the reactivity of dipeptides **158** upon exposure to $NO_2^{\bullet}$. The reaction was quenched with saturated aqueous sodium bicarbonate, extracted with ethyl acetate, dried and evaporated. The reaction mixtures were analysed by comparing the ^{1}H NMR spectra and HPLC retention times with authentic samples. In all cases, Ac-Phe-OMe (**24c**) would be generated as the product of this fragmentation-rearrangement reaction. However, some reactions did not go to completion after 20 minutes. In these cases, the ratios of the product and the starting material were determined from the ^{1}H NMR spectra of the reaction mixtures by integrating the

signals of the methyl protons of either the N-terminal acetyl group or the C-terminal methyl group.

Scheme 6.3-2 shows the results of the exposure of dipeptides Ac-Gly-Phe-OMe (**27c**), Ac-Leu-Phe-OMe (**29c**), Ac-Phe-Phe-OMe (**34c**), Ac-Val-Phe-OMe (**28c**) and Ac-Pro-Phe-OMe (**55c**) to $NO_2^{\bullet}$.



Scheme 6.3-2. Reaction of $NO_2^{\bullet}$ with dipeptides **27c**–**29c**, **34c**, **55c**.

Exposure of the dipeptides Ac-Gly-Phe-OMe (**27c**) to $NO_2^{\bullet}$ resulted in a complete consumption of the starting material after 20 minutes, generating Ac-Phe-OMe (**24c**) as the only detected product. This fragmentation-rearrangement reaction of dipeptide **27c** proceeded with high efficiency to give the expected product **24c** which constituted >95% of the crude product mixture, based on ^{1}H NMR analysis.

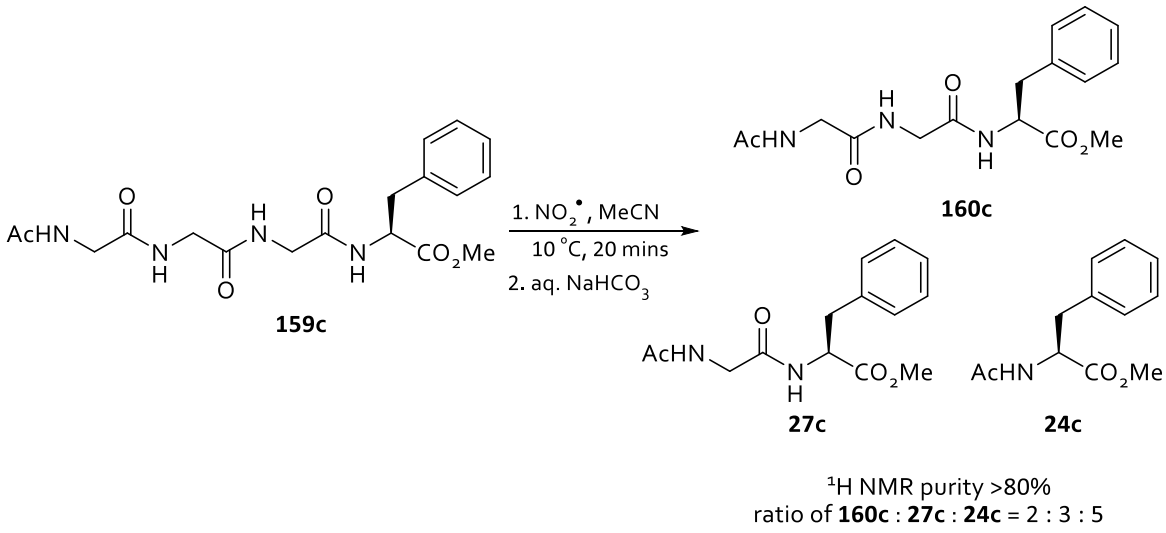
When dipeptides Ac-Leu-Phe-OMe (**29c**) and Ac-Phe-Phe-OMe (**34c**) were treated with $NO_2^{\bullet}$, incomplete conversion was observed, which can be rationalised as a result of an increasing steric demand at the N-terminal α-carbon. The steric hindrance provided from the N-terminal side chain slows down the initial *N*-nitrosation by hindering the access to the amide nitrogen. The ratio of the fragmentation product **24c** and the unreacted starting materials **29c** and **34c** was determined to be 1:3 and 2:3, respectively. This gives evidence that the rate of the fragmentation reaction is affected by these isobutyl and benzyl substituents, with the isobutyl group hinders N-nitrosation to a greater extent.

Finally, the bulky isopropyl substituent in Ac-Val-Phe-OMe (**28c**) hampers the cleavage process considerably as no fragmentation product **24c** was detected after the typical reaction time of 20 minutes. Only after an extensive reaction time (up to 4.5 hours) was employed, the presence of protected phenylalanine **24c** was observed to a certain extent as the fragmentation product (about 60% conversion from the starting material **28c**). However, this extended exposure time also resulted in additional unidentified side reactions, which gave rise to formation of many amide-containing by-products (according to ^{1}H NMR analysis). The unreacted starting material **28c** and the fragmentation product **24c** only constituted approximately 50% of the crude reaction mixture. It is possible that acidification of the reaction system led to decomposition of the peptide, especially when the $NO_2^{\bullet}$-mediated peptide fragmentation process was slow.

There is no reaction of $NO_2^\bullet$ with Ac-Pro-Phe-OMe (**55c**), which can be rationalised by the absence of the N-terminal amide. This means that the initial *N*-nitrosation would not occur, which consequently prevents the whole fragmentation-rearrangement process.

6.3.2 Reaction of $NO_2^\bullet$ with tetrapeptides

Elongating the length of the peptide chain provides an opportunity to assess the effectiveness of the fragmentation-rearrangement reaction mediated by $NO_2^\bullet$. Exposure of the tetrapeptide Ac-Gly-Gly-Gly-Phe-OMe (**159c**) to $NO_2^\bullet$ under the same reaction conditions for Ac-Gly-Phe-OMe (**27c**) led to the formation of three cleavage products: tripeptide Ac-Gly-Gly-Phe-OMe (**160c**), dipeptide Ac-Gly-Phe-OMe (**27c**) and Ac-Phe-OMe (**24c**), in a 2:3:5 ratio respectively (Scheme 6.3-3). These products were confirmed by comparison with authentic samples. The preparation of tetrapeptide **159c**, its reaction with $NO_2^\bullet$ and HRMS analysis was performed by Dr. Luke Gamon.[41]



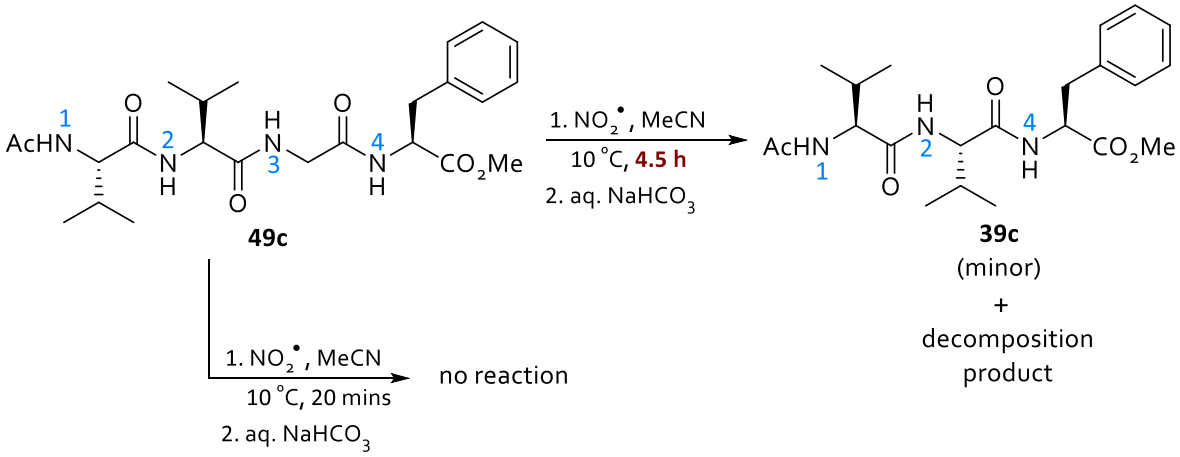
Scheme 6.3-3. Reaction of $NO_2^\bullet$ with Ac-Gly-Gly-Gly-Phe-OMe (**159c**).[41]

^{1}H NMR analysis showed that these three sequential fragmentation products **160c**, **27c**, and **24c** contributed to about 80% of the crude reaction mixture. Even though the remaining 20% constituted various decomposition products that could not be identified, the fragmentation-rearrangement reaction is still the most dominant reaction pathway for tetrapeptide **159c**.

Tripeptide **160c** resulted from excision of one glycine residue after a single *N*-nitrosation process, whereas the shorter chain dipeptide **27c** and phenylalanine **24c** required several successive *N*-nitrosation and fragmentation-rearrangement reactions. The amount of $NO_2^\bullet$ used in this reaction was in excess compared to the substrate (0.5 mL of liquid $NO_2^\bullet$ is roughly about 17 equivalents relative to 1 mmol of substrate). Assuming that the fragmentation is driven by an initial nitrosation with NO^+, every residue requires two equivalents of $NO_2^\bullet$ to facilitate the reaction. This is because NO^+ originates from dissociation of N_2O_4, which is formed from dimerisation of two molecules of $NO_2^\bullet$ (eq. 6.2, page 134). Theoretically, the amount of $NO_2^\bullet$ used would be more than enough to facilitate total fragmentation of 8 amino acid residues. However, this incomplete conversion clearly shows that the ideal nitrosation case does not take into account the evaporation of $NO_2^\bullet$ into the headspace of the reaction flask. The availability of NO^+ would indeed be reduced in the solution because of some evaporated $NO_2^\bullet$ in the reaction flask, and thus, incomplete conversion was observed.

As previously outlined in page 140, an isopropyl side chain hinders *N*-nitrosation at the amide nitrogen. Therefore, when Ac-Val-Val-Gly-Phe-OMe (**49c**) was treated with $NO_2^\bullet$, no cleavage was observed within 20 minutes of exposure due to considerable steric hindrance at the amide nitrogens. However, ESI-HRMS analysis of the crude mixture, which was performed after a reaction time of 4.5 hours, indicated the formation of Ac-Val-Val-Phe-OMe (**39c**) as a minor

component of a complex mixture of decomposition product, as shown in Scheme 6.3-4 (*m*/*z* calcd. for $C_{22}H_{34}N_3O_5$ $[M + H]^+$ = 420.2493, found: 420.2498; *m*/*z* calcd. for $C_{22}H_{33}N_3O_5Na$ $[M + Na]^+$ = 442.2312, found: 442.2312).

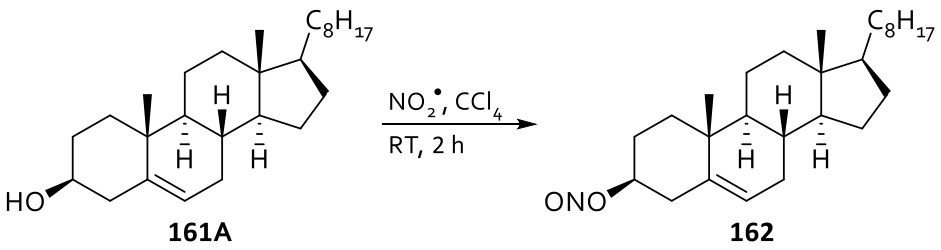


Scheme 6.3-4. Reaction of $NO_2^\bullet$ with Ac-Val-Val-Gly-Phe-OMe (**49c**).

These findings suggest that formation of tripeptide **39c** was initiated by *N*-nitrosation of the sterically least hindered amide nitrogen N-3, followed by excision of the glycine residue from tetrapeptide **49c**. This observation is consistent with the reactions concerning dipeptides with $NO_2^\bullet$ (Section 6.3.1, page 139), that steric bulk prevents the fragmentation-rearrangement reaction from occurring (such as that in Ac-Val-Phe-OMe, **28c**). Hence, the significant steric hindrance provided by two valine residues increases the resistance of the tetrapeptide Ac-Val-Val-Gly-Phe-OMe (**49c**) towards peptide cleavage mediated by $NO_2^\bullet$ and drives the excision of an internal glycine residue chemoselectively at a slow rate.

6.3.3 Reaction of $NO_2^{\bullet}$ with cholesterol and its O-protected derivatives

Cholesterol is the most abundant neutral lipid in the epithelial lining fluid, which is directly exposed to inhaled pollutants, such as $NO_2^{\bullet}$.[24] Therefore, it is important to understand the degradation pathway of cholesterol upon exposure to $NO_2^{\bullet}$. A significant diminution of the cholesterol content was found in the brains of $NO_2^{\bullet}$-intoxicated guinea pigs.[118] Moreover, loss of cholesterol from a cholesterol monomolecular film has been reported and its rate was found to be dependent on the concentration of gaseous $NO_2^{\bullet}$.[119] However, only few studies explored the reaction products. Upon the exposure of cholesterol **161A** to $NO_2^{\bullet}$ in carbon tetrachloride, cholesteryl nitrite **162** was determined to be the major product (Scheme 6.3-5).[120] The product formed was clearly due to the reaction at the hydroxyl group.

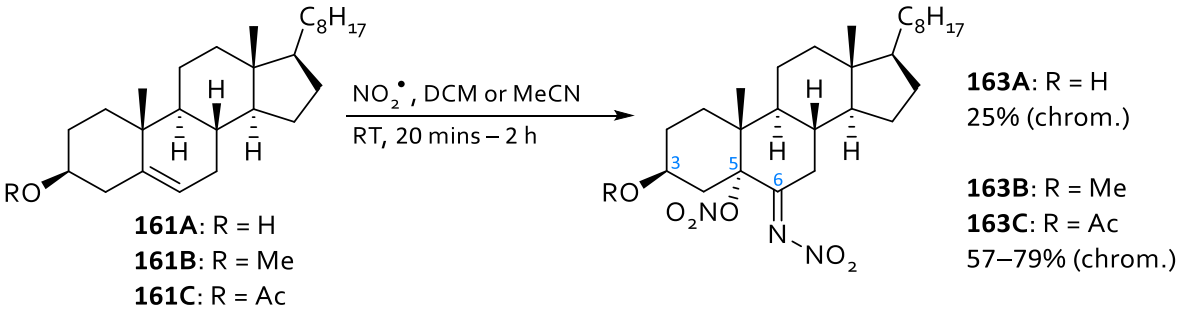


161A **162**

Scheme 6.3-5. Reaction of $NO_2^{\bullet}$ with cholesterol (**161A**).[120]

The reaction also takes place at the other reactive site in **161A**, the electron-rich π system. However, the products from this reaction were not identified. Therefore, in order to gain more mechanistic understanding and assess the potential biological implications, a product study was performed employing the O-protected cholesterols **161B** and **161C**. This system 'isolates' the alkene moiety to be the only reactive site towards $NO_2^{\bullet}$. The synthesis and reaction of **161B** and **161C** were performed by Anton Zalewski as a part of his Masters project.

The reaction was carried out in dichloromethane or acetonitrile to yield nitroimine nitrates **163B** and **163C** as the only detected product (Scheme 6.3-6) in a moderate yield (57–79%). Even though the formation of these products **163B** and **163C** was the major pathway (^{1}H NMR purity of the crude mixture >94%), some decomposition on the silica column during product purification lowered the isolated yield of **163B** and **163C**.

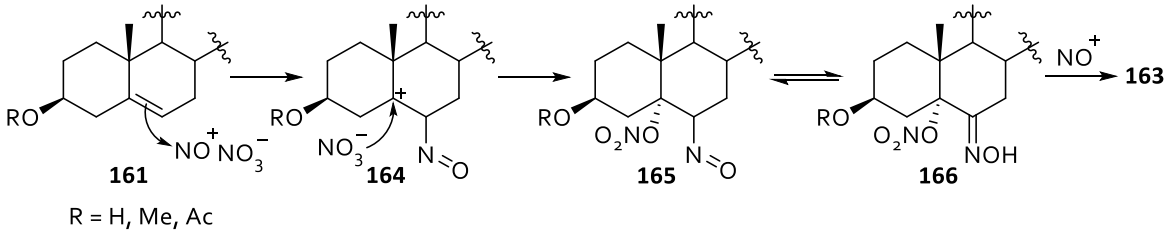


Scheme 6.3-6. Reaction of $NO_2^\bullet$ with cholesterol derivatives **161A–C**.

HRMS analysis for **163B** revealed that the protonated molecular ion $[M + H]^+$ at *m*/*z* 522.35370 is consistent with the molecular formula $[C_{28}H_{48}N_3O_6]^+$ (calcd. 522.35373) and confirms the addition of three nitrogen and five oxygen atoms to the molecular framework. The ^{1}H NMR spectrum showed only loss of the vinyl proton at C-6, while the ^{13}C NMR data revealed that the NMR signals of the former alkene carbon atoms have been shifted quite significantly (C-5: from δ = 140.85 to 92.99 ppm and C-6: from δ = 121.56 to 176.01 ppm). These results suggest that the modifications have taken place exclusively at the double bond of **161B**. Upon standing, crystals of **163B** were obtained. X-ray analysis indicated a nitroimine group at C-6 and a nitrate ester at C-5, which sits *trans* to the angular methyl group and the methoxy substituent at C-3. Similarly, HRMS analysis for **163C** showed the protonated molecular ion $[M + H]^+$ at *m*/*z* 550.34955, which is consistent with the molecular formula $[C_{29}H_{48}N_3O_7]^+$ (*m*/*z* calcd. 550.34868). No crystal structure could be obtained for **163C**, however, the comparable NMR and IR data of **163B** and **163C** support the presence of the same nitroimine nitrate moiety in **163C**.

The work in this chapter employs cholesterol **161A** to investigate if the presence of a hydroxyl group would affect the reactivity of the double bond in **161A** towards $NO_2^{\bullet}$. The reaction of $NO_2^{\bullet}$ with cholesterol (**161A**) in dichloromethane gave a complex mixture of products, unquestionably because of the reactions at the hydroxyl group. Nevertheless, the major product was isolated in 25% yield, and unambiguously identified as the nitroimine nitrate **161A** based on the NMR, IR and HRMS data (Scheme 6.3-6). The finding clearly shows that even in the presence of the reactive hydroxyl group, the selective reaction of $NO_2^{\bullet}$ with the π system in **161A** still occurs to a significant extent. Unfortunately, products arising from the reactions involving the hydroxyl group could not be isolated/identified, probably due to some decomposition on the silica column during attempted purification.

The proposed mechanism for this transformation by $NO_2^{\bullet}$ is outlined in Scheme 6.3-7, picturing only the structural motif relevant to this reaction.



Scheme 6.3-7. Proposed mechanism for the formation of the nitroimine nitrate **163** in the reaction of $NO_2^{\bullet}$ with cholesterol **161** in dichloromethane or acetonitrile.

The reaction is initiated by electrophilic addition of NO^+ to the alkene moiety in **161** to produce the cationic intermediate **164**, which is immediately trapped by the existing NO_3^- from the sterically less hindered side to yield the nitroso nitrate ester **165**. Tautomerisation of this nitroso group in **165** generates oxime **166**, which is then subsequently converted to the nitroimine nitrate **163** through the reaction with NO^+.[121] This mechanism clearly differs from the radical addition/allylic hydrogen abstraction pathway proposed by Pryor et al. (see page 129–130). Instead of the reaction by the free radical $NO_2^{\bullet}$, the ionic species NO^+ resulting from

dissociation of dimer N_2O_4 is involved. This finding is supported by a recent suggestion that $NO_2^{\bullet}$ absorption in the airway fluids is catalysed by antioxidants through hydrolytic dissociation to NO^+ and NO_3^-.[122]

6.4 Summary of findings and future directions

In this chapter, the reactions of $NO_2^{\bullet}$ with peptides and cholesterol have been explored. Short-chain peptides were shown to undergo a fragmentation-rearrangement reaction upon exposure to $NO_2^{\bullet}$. Meanwhile, alkene moiety in cholesterol shows a high reactivity towards $NO_2^{\bullet}$ in the absence or presence of the reactive hydroxyl group.

The key findings of this chapter are:

- ✓ The air pollutant $NO_2^{\bullet}$ facilitates the cleavage of peptide backbones through a novel fragmentation-rearrangement mechanism. During this process, one amino acid residue is expelled from the peptide chain and converted to an α-hydroxy carboxylic acid derivative. The remaining residue, which is one amino acid short, is simultaneously fused through formation of a new peptide bond.
- ✓ This fragmentation-rearrangement process mediated by $NO_2^{\bullet}$ is greatly influenced by the steric hindrance at the peptide bond. Using various di- and tetrapeptide model systems, it was found that the relative rate of this reaction decreases with increasing steric bulk at the α-carbon in the order Gly>Phe>Leu>Val due to the slowing down of the initial amide nitrosation.
- ✓ The process can be repeated in a longer peptide chain until all available $NO_2^{\bullet}$ has been exhausted. This reaction is the major pathway when there is no easily oxidisable side-chain, such as that in tyrosine and tryptophan.

- ✓ The reactive sites of cholesterol are located at both the hydroxyl group and the double bond moiety. The double bond in cholesterol reacts with $NO_2^{\bullet}$ to form a stable nitroimine nitrate moiety. This reaction is not affected by the presence of a hydroxyl group, although the yield was found to be lower.
- ✓ Oxidative damage by $NO_2^{\bullet}$ in some biological molecules, such as peptides possessing non-reactive side chains and cholesterols, was found to take place through an ionic pathway rather than a well-known radical pathway. This pathway has to be considered when assessing $NO_2^{\bullet}$ toxicity.

Further work is necessary to determine if this $NO_2^{\bullet}$-mediated fragmentation-rearrangement is responsible for $NO_2^{\bullet}$ pronounced toxicity in vivo. Besides that, there is a possibility of utilising this reaction to access shorter peptides, for example a smaller cyclic peptide, that is not readily accessible due to the ring constraint or other factors. The sacrificial glycine residue can be installed in the cyclic peptide, which then will be expelled following the fragmentation-rearrangement reaction.

Some results in this chapter were reported in the following publications:

"Fragmentation-Rearrangement of Peptide Backbones Mediated by the Air Pollutant $NO_2^{\bullet}$"

Gamon, L. F.; Nathanael, J. G.; Taggert, B. I.; Henry, F. A.; Bogena, J.; Wille, U. *Chem. Eur. J.* **2015,** *21,* 14924–14930.

"Oxidation of cholesterol and *O*-protected derivatives by the environmental pollutant $NO_2^{\bullet}$"

Zalewski, A. N.; Nathanael, J. G.; White, J. M.; Wille, U. *Chem. Commun.* **2016,** *52,* 4060–4063.

Chapter Seven: Experimental section

Free amino acids and some amino acid derivatives were purchased from Sigma Aldrich Australia, Chem-Supply Australia, AK Scientific USA, or Chem-Impex USA. All reagents for synthesis were obtained from various commercial suppliers (Sigma-Aldrich Australia, Merck Australia, Thermo Fisher Scientific Australia, or AK Scientific USA), and used without further purification. All solvents for synthesis were of analytical reagent (AR) grade and purchased from a variety of manufacturers, including Chem-Supply Australia, RCI Labscan Australia, or Fisher Chemical Australia. For laser flash photolysis experiments and product studies, only HPLC grade acetonitrile (Honeywell B&J Australia) was used as the solvent.

Thin layer chromatography (TLC) was performed to monitor the reactions using aluminium plates coated with silica gel 60 F_{254} (Merck) and visualised with UV light at 254 nm or stained with phosphomolybdic acid (PMA), ninhydrin, $KMnO_4$ or Hanessian's stain followed by heating. The crude products were purified by recrystallisation from hot solvent or silica column chromatography with approximately 30 g of dry silica (Davisil Chromatography Silica Gel LC60A, 40–60 micron, 230–400 mesh) per 1 g of crude product mixture. The eluting solvent consisted of a mixture of either petroleum ether and ethyl acetate or chloroform/dichloromethane and methanol. Solvents were removed under reduced pressure using a rotary evaporator (Buchi). The purity of model compounds was assessed by analytical reversed-phase HPLC on an Alltech Hypersil BDS-C18 5 µm 150 x 4.6 mm (gradient: 100% water buffered with 0.1% TFA to 100% acetonitrile buffered with 0.1% TFA over 25 minutes, 4% min^{-1}, flow rate: 1 mL min^{-1}).

^{1}H and ^{13}C NMR spectra were recorded using an Agilent MR 400 MHz NMR spectrometer, an Agilent DD2 500 MHz NMR spectrometer or a Varian Inova 600 MHz NMR spectrometer in deuterated dimethylsulfoxide (DMSO-d_6), deuterated acetonitrile (CD_3CN), deuterated chloroform ($CDCl_3$), deuterated methanol (CD_3OD), or deuterated water (D_2O) at 25.0 °C at the Bio21 Institute, the University of Melbourne, Australia. Chemical shifts are reported in ppm (δ) using respective residual solvents as the reference (^{1}H NMR δ 2.50 ppm for DMSO, δ 1.94 ppm for CH_3CN, δ 7.26 ppm for $CHCl_3$, δ 3.31 ppm for CH_3OH, and δ 4.79 ppm for H_2O; ^{13}C NMR δ 39.5 ppm for DMSO, δ 118.3 ppm for CH_3CN, δ 77.2 ppm for $CHCl_3$, and δ 49.0 ppm for CH_3OH). ^{1}H NMR data are reported as follows: chemical shift (δ), multiplicity, coupling constants (in Hz, if any), and relative integral. High-Resolution Mass Spectrometry (HRMS) was conducted at the Bio21 Institute, the University of Melbourne, Australia on an Agilent 6220 ESI-TOF mass spectrometer or a Thermo Scientific Exactive Plus Orbitrap mass spectrometer. Parent ions are denoted by $[M + H]^+$, $[M + Na]^+$, or $[M + K]^+$, and in the case of salts, they are $[M - Cl]^+$ or $[M - CF_3CO_2]^+$.

Spectroscopic details are provided for all isolated intermediates and substrates used in the laser flash photolysis and product studies. Compounds for which no spectroscopic data are provided, were obtained from suppliers mentioned on the previous page. Detailed experimental procedures and spectroscopic data for each substrate are given below according to its appearance in Chapter 2, page 13.

GP1. General procedure for the N-acetylation of the amino acids[25]

The amino acid (53.4 mmol) was suspended in 5% aqueous sodium bicarbonate (150 mL) and cooled to 0 °C. Acetic anhydride (6.1 mL, 64.5 mmol) was added dropwise over a period of 1 hour. The mixture was stirred for 4 hours at room

temperature until starting material was fully consumed as monitored using TLC (9:1 EtOH/1 M AcOH, ninhydrin stain). The reaction mixture was then acidified to pH 2–3 with 6 M hydrochloric acid and cooled overnight. The resulting precipitate was collected by filtration, washed with ice-cold water (2 x 10 mL) and dried to give the N-acetylated amino acid. In the absence of solid, precipitation was induced by partial solvent removal.

GP2. General procedure for the esterification of the amino acids[25]

The amino acid (78.8 mmol) was suspended in methanol (250 mL) and cooled to 0 °C. Thionyl chloride (10 mL, 138 mmol) was added dropwise. The mixture was stirred overnight at room temperature until starting material was fully consumed as monitored using TLC (9:1 EtOH/1 M AcOH, ninhydrin stain). The solvent was removed under reduced pressure to give the hydrochloride salt of the amino acid methyl ester.

GP3. General procedure for the esterification of N-protected amino acids[26]

The N-protected amino acid (12.0 mmol) was suspended in methanol (11 mL) and cooled to 0 °C. Thionyl chloride (1.0 mL, 13.8 mmol) was added dropwise. The mixture was stirred for 4 hours at room temperature until starting material was fully consumed as monitored using TLC (EtOAc, PMA or $KMnO_4$ stain). The solvent was removed under reduced pressure and water (20 mL) was added. The product was extracted with dichloromethane (3 x 15 mL). The combined organic extracts were dried over anhydrous sodium sulfate and concentrated under reduced pressure to give the N- and C-protected amino acids.

_N_-Acetylglycine methyl ester (1c)

Prepared according to GP3. White crystals, yield: 0.59 g (37%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.28 (t, J = 6.0 Hz, 1H), 3.80 (d, J = 6.0 Hz, 2H), 3.62 (s, 3H), 1.85 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.5, 169.7, 51.6, 40.6, 22.2 ppm. HRMS (ESI) m/z calcd. for $[C_5H_9NO_3Na]^+$: 154.0480 [M + Na]$^+$, found 154.0477, HRMS (ESI) m/z calcd. for $[C_{10}H_{19}N_2O_6]^+$: 263.1243 [2M + H]$^+$, found 263.1241.

_N_-Acetyl-L-alanine methyl ester (2c)

*(a) N-Acetyl-L-alanine (**2a**):* Prepared according to GP1. A colourless liquid, yield: 1.38 g (20%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.44 (s, 1H), 8.13 (d, J = 7.2 Hz, 1H), 4.16 (p, J = 7.3 Hz, 1H), 1.82 (s, 3H), 1.24 ppm (d, J = 7.3 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.3, 167.0, 47.4, 22.3, 17.2 ppm.

*(b) N-Acetyl-L-alanine methyl ester (**2c**):* Prepared according to GP3 using 1.38 g (10.5 mmol) of *N*-acetyl-L-alanine. A colourless liquid, yield: 1.12 g (74%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.26 (d, J = 6.9 Hz, 1H), 4.23 (p, J = 7.2 Hz, 1H), 3.61 (s, 3H), 1.83 (s, 3H), 1.25 ppm (d, J = 7.3 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.3, 169.1, 51.8, 47.5, 22.2, 17.0 ppm. HRMS (ESI) m/z calcd. for $[C_6H_{12}NO_3]^+$: 146.0817 [M + H]$^+$, found 146.0820, HRMS (ESI) m/z calcd. for $[C_{12}H_{23}N_2O_6]^+$: 291.1556 [2M + H]$^+$, found 291.1564.

_N_-Acetyl-L-valine methyl ester (3c)

Prepared according to GP3. White crystals, yield: 1.87 g (90%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.12 (d, J = 8.1 Hz, 1H), 4.14 (dd, J = 8.1, 6.3 Hz, 1H), 3.62 (s, 3H), 2.05–1.94 (m, 1H), 1.87 (s, 3H), 0.88 (d, J = 6.9 Hz, 3H), 0.86 ppm (d, J = 6.8 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 169.6, 57.4, 51.6, 29.8, 22.2, 19.0, 18.3 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{16}NO_3]^+$: 174.1130 [M + H]$^+$, found 174.1128, HRMS (ESI) m/z calcd. for $[C_8H_{15}NO_3Na]^+$: 196.0950 [M + Na]$^+$, found 196.0958.

_N_-Acetyl-L-leucine methyl ester (4c)

*(a) N-Acetyl-L-leucine (**4a**):* Prepared according to GP1 using 5.65 g (43.1 mmol) of L-leucine. A white solid, yield: 6.61 g (89%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.46 (s, 1H), 8.07 (d, J = 7.9 Hz, 1H), 4.19 (ddd, J = 9.1, 8.1, 6.0 Hz, 1H), 1.83 (s, 3H), 1.69–1.54 (m, 1H), 1.48 (ddd, J = 8.5, 5.5, 3.0 Hz, 2H), 0.89 (d, J = 6.5 Hz, 3H), 0.84 ppm (d, J = 6.5 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.3, 169.2, 50.2, 40.0, 24.3, 22.8, 22.3, 21.3 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{16}NO_3]^+$: 174.1130 [M + H]$^+$, found 174.1133, HRMS (ESI) m/z calcd. for $[C_{16}H_{31}N_2O_6]^+$: 347.2182 [2M + H]$^+$, found 347.2182.

*(b) N-Acetyl-L-leucine methyl ester (**4c**):* Prepared according to GP3. White crystals, yield: 1.91 g (88%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.21 (d, J = 7.7 Hz, 1H), 4.26 (ddd, J = 9.7, 7.6, 5.3 Hz, 1H), 3.61 (s, 3H), 1.84 (s, 3H), 1.67–1.55 (m, 1H), 1.56–1.41 (m, 2H), 0.89 (d, J = 6.7 Hz, 3H), 0.84 ppm (d, J = 6.5 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.2, 169.4, 51.7, 50.2, 39.7, 24.2, 22.7, 22.2, 21.3 ppm. HRMS (ESI) m/z calcd. for $[C_9H_{18}NO_3]^+$: 188.1287 [M + H]$^+$, found 188.1287, HRMS (ESI) m/z calcd. for $[C_{18}H_{35}N_2O_6]^+$: 375.2495 [2M + H]$^+$, found 375.2517.

_N_-Acetyl-L-isoleucine methyl ester (5c)

*(a) N-Acetyl-L-isoleucine (**5a**):* Prepared according to GP1 using 4.76 g (36.3 mmol) of L-isoleucine. White shards, yield: 3.27 g (52%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.50 (s, 1H), 7.99 (d, J = 8.4 Hz, 1H), 4.16 (dd, J = 8.4, 6.1 Hz, 1H), 1.86 (s, 3H), 1.82–1.67 (m, 1H), 1.39 (dqd, J = 14.8, 7.4, 4.4 Hz, 1H), 1.26–1.10 (m, 1H), 0.88–0.79 ppm (m, 6H). ^{113}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.2, 169.4, 56.2, 36.3, 24.7, 22.3, 15.6, 11.3 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{16}NO_3]^+$: 174.1130 [M + H]$^+$, found 174.1126, HRMS (ESI) m/z calcd. for $[C_{16}H_{31}N_2O_6]^+$: 347.2182 [2M + H]$^+$, found 347.2178.

*(b) N-Acetyl-L-isoleucine methyl ester (**5c**):* Prepared according to GP3 using 1.02 g (5.88 mmol) of *N*-acetyl-L-isoleucine. A white solid, yield: 0.91 g (82%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.14 (d, *J* = 8.0 Hz, 1H), 4.19 (dd, *J* = 8.1, 6.5 Hz, 1H), 3.62 (s, 3H), 1.86 (s, 3H), 1.82–1.66 (m, 1H), 1.39 (dqd, *J* = 17.8, 7.5, 4.4 Hz, 1H), 1.26–1.10 (m, 1H), 0.88–0.78 ppm (m, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 169.5, 56.3, 51.5, 36.3, 24.8, 22.2, 15.4, 11.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_9H_{18}NO_3]^+$: 188.1287 $[M + H]^+$, found 188.1284, HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{35}N_2O_6]^+$: 375.2495 $[2M + H]^+$, found 375.2578.

N-Acetyl-L-glutamic acid 1,5-dimethyl ester (6c')

*(a) N-Acetyl-L-glutamic acid (**6a**):* Prepared according to GP1. A white solid, yield: 6.10 g (61%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.53–12.09 (m, 2H), 8.11 (d, *J* = 7.7 Hz, 1H), 4.18 (ddd, *J* = 9.0, 7.7, 5.1 Hz, 1H), 2.26 (td, *J* = 8.0, 2.8 Hz, 2H), 2.02–1.87 (m, 1H), 1.84 (s, 3H), 1.81–1.67 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.7, 173.5, 169.4, 51.2, 30.1, 26.4, 22.4 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_7H_{12}NO_5]^+$: 190.0716 $[M + H]^+$, found 190.0712, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{23}N_2O_{10}]^+$: 379.1353 $[2M + H]^+$, found 379.1349.

*(b) N-Acetyl-L-glutamic acid 1,5-dimethyl ester (**6c'**):* Prepared according to GP3 using 2.0 mL (27.5 mmol) of thionyl chloride. A white solid, yield: 2.21 g (84%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.25 (d, *J* = 7.5 Hz, 1H), 4.24 (ddd, *J* = 8.9, 7.5, 5.4 Hz, 1H), 3.61 (s, 3H), 3.59 (s, 3H), 2.40–2.32 (m, 2H), 2.03–1.90 (m, 1H), 1.84 (s, 3H), 1.84–1.73 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.5, 172.3, 169.5, 51.9, 51.4, 51.1, 29.6, 26.1, 22.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_9H_{16}NO_5]^+$: 218.1029 $[M + H]^+$, found 218.1018, HRMS (ESI) *m*/*z* calcd. for $[C_9H_{15}NO_5Na]^+$: 240.0848 $[M + Na]^+$, found 240.0836.

_N_-Acetyl-_O_-methyl-L-serine methyl ester (7c')[27]

*(a) L-Serine methyl ester hydrochloride salt (**7b**):* Prepared according to GP2 using 4.17 g (39.7 mmol) of L-serine. White crystals, yield: 6.11 g (99%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.61 (s, 3H), 5.63 (t, J = 5.3 Hz, 1H), 4.08 (t, J = 3.5 Hz, 1H), 3.82 (t, J = 3.6 Hz, 2H), 3.73 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 168.5, 59.4, 54.4, 52.7 ppm. HRMS (ESI) m/z calcd. for $[C_4H_{10}NO_3]^+$: 120.0661 $[M - Cl]^+$, found 120.0657.

*(b) N-Acetyl-L-serine methyl ester (**7c**):* L-Serine methyl ester hydrochloride salt (2.01 g, 12.9 mmol) was suspended in dichloromethane and cooled to 0 °C. Triethylamine (3.8 mL, 27.4 mmol), followed by acetyl chloride (1.0 mL, 14.1 mmol), was added dropwise and the resulting mixture was stirred at room temperature for 1 hour. The solvent was then removed under reduced pressure and ethyl acetate (15 mL) was added. The mixture was filtered, and the filtrate was concentrated under reduced pressure. The crude product was purified by column chromatography (24:1 $CHCl_3$/MeOH, $KMnO_4$ stain) to give the desired product as a colourless liquid (1.67 g, 80%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.17 (d, J = 7.7 Hz, 1H), 5.02 (t, J = 5.7 Hz, 1H), 4.32 (dt, J = 7.7, 5.1 Hz, 1H), 3.67 (dt, J = 11.2, 5.6 Hz, 1 H), 3.62 (s, 3H), 3.58 (dd, J = 10.9, 5.3 Hz, 1H), 1.86 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.2, 169.4, 61.3, 54.6, 51.8, 22.3 ppm. HRMS (ESI) m/z calcd. for $[C_6H_{12}NO_4]^+$: 162.0766 $[M + H]^+$, found 162.0762, HRMS (ESI) m/z calcd. for $[C_6H_{11}NO_4Na]^+$: 184.0586 $[M + Na]^+$, found 184.0581.

*(c) N-Acetyl-O-methyl-L-serine methyl ester (**7c'**): N*-Acetyl-L-serine methyl ester (1.66 g, 10.3 mmol) was dissolved in acetonitrile (100 mL) at room temperature under inert atmosphere. Methyl iodide (4.5 mL, 72.3 mmol) and silver(I) oxide (8.39 g, 36.2 mmol) were added. The reaction mixture was stirred for 5 days under the exclusion of light at room temperature and filtered through a pad of celite. The

filtrate was concentrated under reduced pressure and the crude product was purified by column chromatography (3:7 Pet. ether/EtOAc, $KMnO_4$ stain) to give the desired product as a white solid (1.17 g, 65%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.32 (d, *J* = 7.8 Hz, 1H), 4.49 (ddd, *J* = 7.7, 5.6, 4.3 Hz, 1H), 3.66–3.59 (m, 4H), 3.50 (dd, *J* = 9.9, 4.3 Hz, 1H), 3.25 (s, 3H), 1.86 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.7, 169.5, 71.5, 58.4, 52.1, 52.0, 22.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_7H_{14}NO_4]^+$: 176.0923 [M + H]$^+$, found 176.0918, HRMS (ESI) *m*/*z* calcd. for $[C_7H_{13}NO_4Na]^+$: 198.0742 [M + Na]$^+$, found 198.0737.

N_α,N_ε-Diacetyl-L-lysine methyl ester (8c')

*(a) N_α,N_ε-Diacetyl-L-lysine (**8a'**):* Prepared according to GP1 using 3.28 g (17.4 mmol) of L-lysine monohydrochloride salt and 4.0 mL (42.4 mmol) of acetic anhydride. A colourless liquid, yield: 1.00 g (24%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.93 (d, *J* = 7.8 Hz, 1H), 7.81 (t, *J* =5.5 Hz, 1H), 4.07 (td, *J* = 8.2, 5.0 Hz, 1H), 2.98 (q, *J* = 6.6 Hz, 3H), 1.83 (s, 3H), 1.77 (s, 3H), 1.64 (ddd, *J* = 12.6, 8.9, 4.9 Hz, 1H), 1.52 (td, *J* = 14.7, 14.1, 8.0 Hz, 1H), 1.39–1.32 (m, 2H), 1.29–1.23 ppm (m, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.12, 169.04, 168.97, 52.39, 38.39, 31.23, 28.87, 22.94, 22.63, 22.50 ppm.

*(b) N_α,N_ε-Diacetyl-L-lysine methyl ester (**8c'**):* Prepared according to GP3 using 1.00 g (4.34 mmol) of N_α,N_ε-diacetyl-L-lysine. A white solid, yield: 0.18 g (17%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.20 (d, *J* = 7.4 Hz, 1H), 7.79 (t, *J* = 5.3 Hz, 1H), 4.17 (ddd, *J* = 8.8, 7.3, 5.3 Hz, 1H), 3.61 (s, 3H), 2.99 (q, *J* = 6.5 Hz, 2H), 1.84 (s, 3H), 1.77 (s, 3H), 1.71–1.58 (m, 1H), 1.55 (ddt, *J* = 13.8, 8.8, 3.9 Hz, 1H), 1.42–1.30 (m, 2H), 1.33–1.20 ppm (m, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 169.4, 168.9 51.9, 51.7, 38.2, 30.6, 28.7, 22.8, 22.6, 22.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{11}H_{20}N_2O_4]^+$: 245.1501 [M + H]$^+$, found 245.1496, HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{41}N_4O_8]^+$: 489.2924 [2M + H]$^+$, found 489.2920.

GP4. General procedure for the peptide coupling[25]

The N-protected amino acid/peptide (10.0 mmol), the salt of amino acid/peptide methyl ester (10.1 mmol) and HBTU (3.75 g, 9.9 mmol) were suspended in dimethylformamide (15 mL) and cooled to 0 °C. Triethylamine (4.2 mL, 30.0 mmol) was added dropwise and the resulting mixture was stirred overnight at room temperature. The reaction was monitored using TLC (EtOAc, PMA or ninhydrin stain) until a large amount of starting material was consumed (some remained due to the excess amount of starting material). The mixture was then partitioned between 1 M hydrochloric acid (100 mL) and ethyl acetate (100 mL). The aqueous layer was extracted with ethyl acetate (2 x 100 mL), and the combined organic layers were washed sequentially with 5% aqueous sodium bicarbonate (100 mL) and brine (100 mL). The organic layer was dried over anhydrous magnesium sulfate and concentrated under reduced pressure to give a sticky oil or a white solid. The crude product was purified by column chromatography or by recrystallisation to give the N- and C-protected short peptide.

GP5. General procedure for the N-Boc deprotection[28]

The N-Boc protected peptide methyl ester (10.0 mmol, synthesised according to GP4) was dissolved in dichloromethane (8 mL) and cooled to 0 °C. Trifluoroacetic acid (8 mL, 104.5 mmol) was added dropwise. The resulting mixture was stirred for 4 hours at room temperature until starting material was fully consumed as monitored using TLC (EtOAc, ninhydrin stain). The solvent was removed under reduced pressure to give a yellow oil. Toluene (10 mL) was repeatedly added and removed under reduced pressure to get rid of residual trifluoroacetic acid and give the trifluoroacetate salt of the peptide methyl ester.

GP6. General procedure for the hydrolysis of C-terminal methyl esters[29]

The N-acetylated peptide methyl ester (1.82 mmol, synthesised according to GP4) was dissolved in tetrahydrofuran (28 mL) and cooled to 0 °C. To this solution was added an ice-cold solution of lithium hydroxide (0.60 g, 25.2 mmol) in water (28 mL) and the resulting mixture was stirred overnight at room temperature. A solution of 3 M hydrochloric acid (10 mL) and brine (20 mL) were then added. The mixture was extracted with ethyl acetate (4 x 20 mL) and the combined organic extracts were washed with brine (50 mL). The organic layer was then dried over anhydrous magnesium sulfate and concentrated under reduced pressure to give the N-acetylated peptide.

***N*-Acetyl-L-valyl-L-valine methyl ester (9a)**

*(a) L-Valine methyl ester hydrochloride salt (**3b**):* Prepared according to GP2. A white solid, yield: 12.8 g (99%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.66 (s, 3H), 3.83 (d, *J* = 4.7, 1H), 3.74 (s, 3H), 2.19 (heptd, *J* = 6.9, 4.5 Hz, 1H), 0.98 (d, *J* = 6.9 Hz, 3H), 0.93 ppm (d, *J* = 6.9 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.2, 57.2, 52.2, 29.3, 18.4, 17.6 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_6H_{14}NO_2]^+$: 132.1019 [M – Cl]$^+$, found 132.1021.

*(b) N-Acetyl-L-valyl-L-valine methyl ester (**9a**):* Prepared according to GP4, purified by column chromatography (1:9 Pet. ether/EtOAc, PMA stain) and recrystallised from hot ethyl acetate. A white powder, yield: 2.11 g (77%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.13 (d, *J* = 7.7 Hz, 1H), 7.84 (d, *J* = 8.9 Hz, 1H), 4.27 (dd, *J* = 8.9, 7.0 Hz, 1H), 4.16–4.03 (m, 1H), 3.61 (s, 3H), 2.03 (dp, *J* = 13.5, 6.9 Hz, 1 H), 1.92 (dq, *J* = 14.3, 7.2 Hz, 1H), 1.85 (s, 3H), 0.89 (d, *J* = 6.8 Hz, 3H), 0.87 (d, *J* = 5.3 Hz, 3H), 0.85 (d, *J* = 5.3 Hz, 3H), 0.83 ppm (d, *J* = 6.9 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.8, 171.7, 169.1, 57.5, 57.2, 51.5, 30.6, 29.6, 22.4, 19.1, 18.9, 18.3, 18.2 ppm.

HRMS (ESI) *m/z* calcd. for $[C_{13}H_{25}N_2O_4]^+$: 273.1814 $[M + H]^+$, found 273.1810, HRMS (ESI) *m/z* calcd. for $[C_{13}H_{24}N_2O_4Na]^+$: 295.1634 $[M + Na]^+$, found 295.1628.

_N_-Acetyl-L-valyl-L-leucine methyl ester (10c)

*(a) L-Leucine methyl ester hydrochloride salt (**4b**):* Prepared according to GP2 using 4.98 g (38.0 mmol) of L-leucine. A white solid, yield: 6.72 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.73 (s, 3H), 3.90 (t, *J* = 7.0 Hz, 1H), 3.73 (s, 3H), 1.81–1.72 (m, 1H), 1.72–1.58 (m, 2H), 0.88 ppm (d, *J* = 6.4 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.3, 52.7, 50.5, 39.1, 23.7, 22.2, 22.0 ppm. HRMS (ESI) *m/z* calcd. for $[C_7H_{16}NO_2]^+$: 146.1181 $[M - Cl]^+$, found 146.1182.

*(b) N-Acetyl-L-valyl-L-leucine methyl ester (**10c**):* Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white powder, yield: 0.71 g (25%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.27 (d, *J* = 7.3 Hz, 1H), 7.84 (d, *J* = 8.9 Hz, 1H), 4.25 (ddd, *J* = 9.7, 7.3, 5.1 Hz, 1H), 4.19 (dd, *J* = 8.9, 7.0 Hz, 1H), 3.60 (s, 3H), 1.93 (h, *J* = 6.8 Hz, 1 H), 1.85 (s, 3H), 1.68–1.57 (m, 1H), 1.59–1.52 (m, 1H), 1.52–1.42 (m, 1H), 0.89 (d, *J* = 6.5 Hz, 3H), 0.86 (d, *J* = 6.8 Hz, 3H), 0.83 ppm (d, *J* = 6.5 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, $CDCl_3$): δ 173.2, 171.5, 170.4, 58.5, 52.4, 51.0, 41.2, 31.6, 24.9, 23.3, 22.9, 22.0, 19.1, 18.4 ppm. HRMS (ESI) *m/z* calcd. for $[C_{14}H_{27}N_2O_4]^+$: 287.1971 $[M + H]^+$, found 287.1965, HRMS (ESI) *m/z* calcd. for $[C_{14}H_{26}N_2O_4Na]^+$: 309.1790 $[M + Na]^+$, found 309.1783.

_N_-Acetyl-L-leucyl-L-alanine methyl ester (11c)

*(a) L-Alanine methyl ester hydrochloride (**2b**):* Prepared according to GP2. A white solid, yield: 11.2 g (99%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.71 (s, 3H), 4.04 (q, *J* = 7.2, 1H), 3.72 (s, 3H), 1.42 ppm (d, *J* = 7.2 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.4, 52.8, 47.8, 15.7 ppm. HRMS (ESI) *m/z* calcd. for $[C_4H_{10}NO_2]^+$: 104.0712 $[M - Cl]^+$, found 104.0709.

*(b) N-Acetyl-L-leucyl-L-alanine methyl ester (**11c**):* Prepared according to GP4, purified by column chromatography (EtOAc, PMA stain) and recrystallised from hot water. White crystals, yield: 0.35 g (14%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.36 (d, *J* = 6.9 Hz, 1H), 7.94 (d, *J* = 8.4 Hz, 1H), 4.33 (td, *J* = 9.0, 5.8 Hz, 1H), 4.24 (p, *J* = 7.2 Hz, 1H), 3.60 (s, 3H), 1.81 (s, 3H), 1.68–1.52 (m, 1H), 1.49–1.32 (m, 2H), 1.27 (d, *J* = 7.3 Hz, 3H), 0.89 (d, *J* = 6.5 Hz, 3H), 0.84 ppm (d, *J* = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.9, 172.2, 169.0, 51.8, 50.4, 47.5, 41.0, 24.1, 23.0, 22.5, 21.7, 16.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{23}N_2O_4]^+$: 259.1652 [M + H]$^+$, found 259.1653, HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{22}N_2O_4Na]^+$: 281.1472 [M + Na]$^+$, found 281.1470.

N-Acetyl-L-leucyl-L-valine methyl ester (12c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white powder, yield: 0.62 g (21%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.10 (d, *J* = 8.0 Hz, 1H), 7.96 (d, *J* = 8.2 Hz, 1H), 4.39 (q, *J* = 7.8 Hz, 1H), 4.13 (dd, *J* = 8.0, 6.4 Hz, 1H), 3.61 (s, 3H), 2.03 (h, *J* = 6.7 Hz, 1H), 1.82 (s, 3H), 1.59 (dp, *J* = 13.4, 6.7 Hz, 1H), 1.40 (t, *J* = 7.3 Hz, 2H), 0.88 (d, *J* = 6.7 Hz, 6H), 0.86 (d, *J* = 5.2 Hz, 3H), 0.84 ppm (d, *J* = 4.9 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.6, 171.9, 169.0, 57.4, 51.6, 50.6, 40.8, 29.8, 24.1, 23.0, 22.4, 21.7, 18.9, 18.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{27}N_2O_4]^+$: 287.1971 [M + H]$^+$, found 287.1979, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{26}N_2O_4Na]^+$: 309.1790 [M + Na]$^+$, found 309.1797.

N-Acetyl-L-leucyl-L-leucine methyl ester (13c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white hygroscopic solid, yield: 2.14 g (72%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.23 (d, *J* = 7.6 Hz, 1H), 7.93 (d, *J* = 8.3 Hz, 1H), 4.33 (ddd, *J* = 9.7, 8.2, 5.6 Hz, 1H), 4.26 (ddd, *J* = 10.0, 7.5, 4.9 Hz, 1H), 3.60 (s, 3H), 1.82 (s,

3H), 1.65–1.53 (m, 3H), 1.51–1.34 (m, 3H), 0.89 (d, *J* = 6.4 Hz, 3H), 0.88 (d, *J* = 6.5 Hz, 3H), 0.85 (d, *J* = 6.6 Hz, 3H), 0.83 ppm (d, *J* = 6.8 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.9, 172.5, 169.0, 51.8, 50.6, 50.2, 41.0, 39.6, 24.2, 24.1, 23.0, 22.8, 22.5, 21.8, 21.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{29}N_2O_4]^+$: 301.2127 [M + H]$^+$, found 301.2127, HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{28}N_2O_4Na]^+$: 323.1947 [M + Na]$^+$, found 323.1953.

_N_-Acetyl-L-valyl-L-leucyl-L-valine methyl ester (14c)

Synthesised from the C-terminus (Scheme 2.1-2, page 18)

*(a) N-Boc-L-leucyl-L-valine methyl ester (**12e**):* Prepared according to GP4, purified by column chromatography (4:1 Pet. ether/EtOAc, ninhydrin stain). A white powder, yield: 2.60 g (76%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.88 (d, *J* = 8.3 Hz, 1H), 6.91 (d, *J* = 8.5 Hz, 1H), 4.19 (dd, *J* = 8.3, 6.2 Hz, 1H), 4.03 (td, *J* = 9.1, 5.8 Hz, 1H), 3.62 (s, 3H), 2.09–1.97 (m, 1H), 1.66–1.54 (m, 1H), 1.46–1.37 (m, 2H), 1.37 (s, 9H), 0.88 (d, *J* =1.7 Hz, 3H), 0.87 (d, *J* = 1.6 Hz, 3H), 0.86 (d, *J* = 3.2 Hz, 3H), 0.85 ppm (d, *J* = 3.2 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 171.9, 155.2, 78.0, 57.1, 52.7, 51.7, 40.6, 30.1, 28.2, 24.2, 22.9, 21.6, 18.9, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{33}N_2O_5]^+$: 345.2390 [M + H]$^+$, found 345.2385, HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{32}N_2O_5Na]^+$: 367.2209 [M + Na]$^+$, found 367.2204.

*(b) L-Leucyl-L-valine methyl ester trifluoroacetate salt (**12b**):* Prepared according to GP5 using 2.58 g (7.48 mmol) of *N*-Boc-L-leucyl-L-valine methyl ester. A white powder, yield: 2.61 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.72 (d, *J* = 7.7 Hz, 1H), 8.20 (s, 3H), 4.21 (dd, *J* = 7.7, 6.0 Hz, 1H), 3.89 (t, *J* = 7.6 Hz, 1H), 3.64 (s, 3H), 2.15–2.01 (m, 1H), 1.67 (dp, *J* = 13.3, 6.7 Hz, 1H), 1.63–1.45 (m, 2H), 0.93 (d, *J* = 5.9 Hz, 6H), 0.89 ppm (d, *J* = 7.3 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.3, 169.4, 57.7, 51.8, 50.6, 40.3, 29.8, 23.4, 22.6, 22.0, 18.9, 18.2 ppm. HRMS (ESI) *m*/*z*

calcd. for $[C_{12}H_{25}N_2O_3]^+$: 245.1865 $[M - CF_3CO_2]^+$, found 245.1859, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{49}N_4O_6]^+$: 489.3652 $[2M - C_4HO_4F_6]^+$, found 489.3654.

*(c) N-Acetyl-L-valyl-L-leucyl-L-valine methyl ester (**14c**):* Prepared according to GP4 using 0.53 g (3.30 mmol) of *N*-acetyl-L-valine and 1.18 g (3.30 mmol) of L-leucyl-L-valine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 63.6 mg (5%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.97 (d, *J* = 3.2 Hz, 1H), 7.95 (d, *J* = 3.3 Hz, 1H), 7.86, (d, *J* = 8.8 Hz, 1H), 4.38 (q, *J* = 7.8 Hz, 1H), 4.14 (ddd, *J* = 9.3, 6.7, 2.6 Hz, 2H), 3.61 (s, 3H), 2.03 (dq, *J* = 13.2, 6.6 Hz, 1H), 1.93 (h, *J* = 6.8 Hz, 1H), 1.85 (s, 3H), 1.60 (dp, *J* = 13.3, 6.7 Hz, 1H), 1.44 (dd, *J* = 8.3, 6.2 Hz, 2H), 0.91–0.83 (m, 12H), 0.83–0.78 ppm (m, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 171.7, 170.9, 169.2, 57.7, 57.3, 51.6, 50.8, 40.7, 30.3, 29.8, 24.1, 23.0, 22.5, 21.7, 19.2, 18.9, 18.2, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{19}H_{36}N_3O_5]^+$: 386.2655 $[M + H]^+$, found 386.2651, HRMS (ESI) *m*/*z* calcd. for $[C_{19}H_{35}N_3O_5Na]^+$: 408.2474 $[M + Na]^+$, found 408.2469.

_N_-Acetyl-L-leucyl-L-leucyl-L-leucine methyl ester (15c)

Synthesised from the N-terminus (Scheme 2.1-3, page 18).

*(a) N-Acetyl-L-leucyl-L-leucine (**13a**):* Prepared according to GP6 using 0.27 g (0.90 mmol) of *N*-acetyl-L-leucyl-L-leucine methyl ester. A white powder, yield: 0.25 g (95%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.44 (s, 1H), 8.07 (d, *J* = 7.8 Hz, 1H), 7.92 (d, *J* = 8.3 Hz, 1H), 4.33 (td, *J* = 8.9, 5.8 Hz, 1H), 4.19 (td, *J* = 8.7, 5.2 Hz, 1H), 1.81 (s, 3H), 1.67–1.55 (m, 2H), 1.58–1.45 (m, 2H), 1.46–1.33 (m, 2H), 0.88 (dd, *J* = 6.5, 4.7 Hz, 6H), 0.83 ppm (t, *J* = 5.4 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.9, 172.3, 168.9, 50.6, 50.1, 41.0, 39.8, 24.2, 24.1, 23.1, 22.9, 22.5, 21.7, 21.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{27}N_2O_4]^+$: 287.1971 $[M + H]^+$, found 287.1964, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{26}N_2O_4Na]^+$: 309.1790 $[M + Na]^+$, found 309.1783.

*(b) N-Acetyl-L-leucyl-L-leucyl-L-leucine methyl ester (**15c**):* Prepared according to GP4 using 0.23 g (0.81 mmol) of *N*-acetyl-L-leucyl-L-leucine and 0.15 g (0.82 mmol) of L-leucine methyl ester hydrochloride salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.11 g (32%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.09 (d, J = 7.7 Hz, 1H), 7.96 (d, J = 8.0 Hz, 1H), 7.91 (d, J = 8.3 Hz, 1H), 4.36–4.21 (m, 3H), 3.59 (s, 3H), 1.82 (s, 3H), 1.66–1.50 (m, 4H), 1.52–1.34 (m, 5H), 0.92–0.79 ppm (m, 18H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.79, 173.59, 173.29, 172.36, 52.89, 52.44, 52.00, 51.35, 40.88, 40.60, 40.31, 25.00, 24.92 (2C), 22.76, 22.75, 22.74, 22.41, 21.50, 21.41, 21.18 ppm. HRMS (ESI) m/z calcd. for $[C_{21}H_{40}N_3O_5]^+$: 414.2968 $[M + H]^+$, found 414.2967, HRMS (ESI) m/z calcd. for $[C_{21}H_{39}N_3O_5Na]^+$: 436.2787 $[M + Na]^+$, found 436.2782.

_N_-Acetyl-2-aminoisobutyric acid methyl ester (16c)

*(a) N-Acetyl-2-aminoisobutyric acid (**16a**):* Prepared according to GP1. A white solid, yield: 2.97 g (38%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 12.08 (s, 1H), 8.00 (s, 1H), 1.77 (s, 3H), 1.31 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 175.6, 168.6, 54.6, 25.0, 22.6 ppm. HRMS (ESI) m/z calcd. for $[C_6H_{12}NO_3]^+$: 146.0817 $[M + H]^+$, found 146.0812, HRMS (ESI) m/z calcd. for $[C_6H_{11}NO_3Na]^+$: 168.0637 $[M + Na]^+$, found 168.0631.

*(b) N-Acetyl-2-aminoisobutyric acid methyl ester (**16c**):* Prepared according to GP3 using 1.02 g (7.01 mmol) of *N*-acetyl-2-aminoisobutyric acid. White crystals, yield: 0.53 g (47%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.17 (s, 1H), 3.54 (s, 3H), 1.77 (s, 3H), 1.31 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.5, 168.7, 54.7, 51.7, 24.9, 22.3 ppm. HRMS (ESI) m/z calcd. for $[C_7H_{14}NO_3]^+$: 160.0974 $[M + H]^+$, found 160.0969, HRMS (ESI) m/z calcd. for $[C_7H_{13}NO_3Na]^+$: 182.0793 $[M + Na]^+$, found 182.0789.

_N_-Acetyl-2-aminoisobutyric acid methyl(d_3) ester (16h)

Prepared according to GP3 using 0.79 g (5.44 mmol) of *N*-acetyl-2-aminoisobutyric acid and methanol-d_4 (6.0 mL). White crystals, yield: 0.53 g (60%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.17 (s, 1H), 1.78 (s, 3H), 1.31 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.5, 168.7, 54.7, 24.9, 22.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_7H_{11}D_3NO_3]^+$: 163.1157 $[M + H]^+$, found 163.1156, HRMS (ESI) *m*/*z* calcd. for $[C_7H_{10}D_3NO_3Na]^+$: 185.0976 $[M + Na]^+$, found 185.0976.

_N_-Trifluoroacetyl-2-aminoisobutyric acid methyl ester (16i)

*(a) N-Trifluoroacetyl-2-aminoisobutyric acid (**16f**):*[30] 2-Aminoisobutyric acid (2.52 g, 24.2 mmol) was suspended in trifluoroacetic acid (3.75 mL, 49.0 mmol) under inert atmosphere and cooled to 0 °C. Trifluoroacetic anhydride (5.2 mL, 36.8 mmol) was added dropwise and the resulting mixture was stirred for 40 hours at room temperature. Water (5 mL) was then added to the yellow solution to quench the reaction. The mixture was stirred for a further 1 hour and half of the solvent was removed under reduced pressure. The solution was cooled to 0 °C, and the resulting crystals were filtered and washed with ice-cold 1 M hydrochloric acid (2 x 5 mL) to give the desired product as white needles (3.51 g, 72%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.66 (s, 1H), 9.48 (s, 1H), 1.42 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.0, 155.6 (q, *J* = 36.5 Hz), 115.6 (q, *J* = 288.6 Hz), 56.2, 24.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_6H_7NO_3F_3]^-$: 198.0383 $[M – H]^-$, found 198.0378, HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{15}N_2O_6F_6]^-$: 397.0839 $[2M – H]^-$, found 397.0856.

*(b) N-Trifluoroacetyl-2-aminoisobutyric methyl ester (**16i**):* Prepared according to GP3 using 1.20 g (6.04 mmol) of *N*-trifluoroacetyl-2-aminoisobutyric acid. White crystals, yield: 0.49 g (38%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.70 (s, 1H), 3.61 (s,

3H), 1.43 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.9, 155.8 (q, J = 36.8 Hz), 115.6 (q, J = 288.2 Hz), 56.3, 52.3, 24.3 ppm. HRMS (ESI) m/z calcd. for $[C_7H_{11}NO_3F_3]^+$: 214.0686 $[M + H]^+$, found 214.0701, HRMS (ESI) m/z calcd. for $[C_7H_{10}NO_3F_3Na]^+$: 236.0505 $[M + Na]^+$, found 236.0509.

_N_-Acetyl(d_2)-2-aminoisobutyric acid methyl ester (16j)

*(a) 2-Aminoisobutyric acid methyl ester hydrochloride salt (**16b**):* Prepared according to GP2 using 4.84 g (46.9 mmol) of 2-aminoisobutyric acid. White crystals, yield: 6.94 g (96%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.82 (s, 3H), 3.74 (s, 3H), 1.49 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.0, 55.8, 53.1, 23.4 ppm. HRMS (ESI) m/z calcd. for $[C_5H_{12}NO_2]^+$: 118.0868 $[M - Cl]^+$, found 118.0865.

*(b) N-Acetyl(d_2)-2-aminoisobutyric acid methyl ester (**16j**):* Prepared according to the procedure for *N*-acetyl-L-serine methyl ester using 0.58 g (3.78 mmol) of 2-aminoisobutyric methyl ester hydrochloride salt and acetyl(d_3) chloride (0.25 mL, 3.39 mmol), purified by column chromatography (97:3 $CHCl_3$/MeOH, $KMnO_4$ stain). White crystals, yield: 0.33 g (60%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.17 (s, 1H), 3.54 (s, 3H), 1.74 (s, 1H), 1.31 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.5, 168.7, 54.7, 51.8, 24.9 ppm. HRMS (ESI) m/z calcd. for $[C_7H_{12}D_2NO_3]^+$: 162.1094 $[M + H]^+$, found 162.1095, HRMS (ESI) m/z calcd. for $[C_7H_{11}D_2NO_3Na]^+$: 184.0914 $[M + Na]^+$, found 184.0914.

_N_-Acetyl(d_3)-2-aminoisobutyric acid methyl(d_3) ester (16k)

*(a) 2-Aminoisobutyric acid methyl(d_3) ester hydrochloride salt (**16g**):* Prepared according to GP2 using 0.48 g (4.65 mmol) of 2-aminoisobutyric acid and methanol-d_4 (15 mL). White crystals, yield: 0.71 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.76 (s, 3H), 1.48 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ

172.1, 55.9, 23.4 ppm. HRMS (ESI) *m/z* calcd. for $[C_5H_9D_3NO_2]^+$: 121.1051 $[M - Cl]^+$, found 121.1053.

*(b) N-Acetyl(d_3)-2-aminoisobutyric acid methyl(d_3) ester (**16k**):* Prepared according to the procedure for *N*-acetyl-L-serine methyl ester using 0.71 g (4.51 mmol) of 2-aminoisobutyric methyl(d_3) ester hydrochloride salt and acetyl(d_3) chloride (0.36 mL, 4.88 mmol), purified by column chromatography (97:3 $CHCl_3$/MeOH, $KMnO_4$ stain). White crystals, yield: 0.60 g (81%). ^1H NMR (400 MHz, DMSO-d_6): δ 8.17 (s, 1H), 1.31 ppm (s, 6H). ^{13}C {^1H} NMR (101 MHz, DMSO-d_6): δ 174.5, 168.7, 54.7, 24.9 ppm. HRMS (ESI) *m/z* calcd. for $[C_7H_8D_6NO_3]^+$: 166.1345 $[M + H]^+$, found 166.1344, HRMS (ESI) *m/z* calcd. for $[C_7H_7D_6NO_3Na]^+$: 188.1165 $[M + Na]^+$, found 188.1163.

_N_-Pivaloyl-2-aminoisobutyric acid methyl ester (16l)

Prepared according to the procedure for *N*-acetyl-L-serine methyl ester using 0.81 g (5.25 mmol) of 2-aminoisobutyric methyl ester hydrochloride salt and pivaloyl chloride (0.8 mL, 6.50 mmol), purified by column chromatography (99:1 $CHCl_3$/MeOH, $KMnO_4$ stain). White crystals yield 0.58 g (55%). ^1H NMR (400 MHz, DMSO-d_6): δ 7.45 (s, 1H), 3.53 (s, 3H), 1.33 (s, 6H), 1.07 (s, 9H) ppm. ^{13}C {^1H} NMR (101 MHz, DMSO-d_6): δ 177.0, 174.6, 55.0, 51.6, 37.8, 27.2, 24.9 ppm. HRMS (ESI) *m/z* calcd. for $[C_{10}H_{20}NO_3]^+$: 202.1438 $[M + H]^+$, found 202.1440, HRMS (ESI) *m/z* calcd. for $[C_{10}H_{19}NO_3Na]^+$: 224.1257 $[M + Na]^+$, found 224.1257.

GP7. General procedure for the N-Boc protection of amino acid methyl esters[31]

The hydrochloride salt of amino acid methyl ester (14.1 mmol) and triethylamine (4.0 mL, 28.7 mmol) were added to a solution of Boc anhydride (2.77 g, 12.7 mmol) in dichloromethane (40 mL). The resulting mixture was stirred for 18 hours at room temperature. The solution was cooled to 0 °C and washed with ice-cold 0.5 M hydrochloric acid (2 x 50 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow oil. The crude product was purified by column chromatography (9:1 Pet. ether/EtOAc, ninhydrin stain) to give the N-Boc protected amino acid methyl ester.

***N*-Boc-glycine methyl ester (1e)**

*(a) Glycine methyl ester hydrochloride salt (**1b**):* Prepared according to GP2. A white solid, yield: 9.64 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.58 (s, 3H), 3.76 (s, 2H), 3.72 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 168.0, 52.5, 39.5 ppm.

*(b) N-Boc-glycine methyl ester (**1e**):* Prepared according to GP7 using 1.99 g (15.9 mmol) of glycine methyl ester hydrochloride salt. A colourless liquid, yield: 1.85 g (77%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.18 (t, J = 6.2 Hz, 1H), 3.65 (d, J = 6.2 Hz, 2H), 3.60 (s, 3H), 1.36 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.9, 155.8, 78.2, 51.6, 41.8, 28.1 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{15}NO_4Na]^+$: 212.0899 $[M + Na]^+$, found 212.0896.

***N*-Boc-L-alanine methyl ester (2e)**

Prepared according to GP7. A colourless liquid, yield: 0.80 g (30%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.26 (d, J = 7.5 Hz, 1H), 4.01 (p, J = 7.4 Hz, 1H), 3.61 (s, 3H), 1.37 (s, 9H), 1.22 ppm (d, J = 7.4 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.6, 155.2, 78.1, 51.7, 48.9, 28.2, 16.9 ppm. HRMS (ESI) m/z calcd. for $[C_9H_{18}NO_4]^+$:

204.1236 [M + H]$^+$, found 204.1231, HRMS (ESI) *m*/*z* calcd. for [$C_9H_{17}NO_4Na$]$^+$: 226.1055 [M + Na]$^+$, found 226.1050.

_N_-Boc-2-aminoisobutyric acid methyl ester (16e)

Prepared according to GP7. A white powder, yield: 2.13 g (77%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.22 (s, 1H), 3.57 (s, 3H), 1.35 (s, 9H), 1.30 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 175.0, 154.6, 78.0, 55.2, 51.7, 28.1, 25.2 ppm. HRMS (ESI) *m*/*z* calcd. for [$C_{10}H_{20}NO_4$]$^+$: 218.1392 [M + H]$^+$, found 218.1387, HRMS (ESI) *m*/*z* calcd. for [$C_{10}H_{19}NO_4Na$]$^+$: 240.1212 [M + Na]$^+$, found 240.1206.

_N_-Boc-L-leucine methyl ester (4e)

Prepared according to GP7. A colourless oil, yield: 2.87 g (92%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.22 (d, *J* = 8.0 Hz, 1H), 3.98 (ddd, *J* = 10.1, 8.0, 4.9 Hz, 1H), 3.61 (s, 3H), 1.68–1.55 (m, 1H), 1.56–1.46 (m, 1H), 1.45–1.39 (m, 1H), 1.37 (s, 9H), 0.87 (d, *J* = 6.7 Hz, 3H), 0.84 ppm (d, *J* = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.7, 155.7, 78.9, 52.1, 51.6, 40.4, 27.5, 24.5, 22.1, 20.7 ppm. HRMS (ESI) *m*/*z* calcd. for [$C_{12}H_{24}NO_4$]$^+$: 246.1705 [M + H]$^+$, found 246.1699, HRMS (ESI) *m*/*z* calcd. for [$C_{12}H_{23}NO_4Na$]$^+$: 268.1525 [M + Na]$^+$, found 268.1519.

_N_-Boc-L-alanyl-L-alanine methyl ester (17e)

Prepared according to GP4, purified by column chromatography (3:2 Pet. ether/EtOAc, ninhydrin stain). A white solid, yield: 0.73 g (28%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.16 (d, *J* = 7.1 Hz, 1H), 6.85 (d, *J* = 7.8 Hz, 1H), 4.26 (p, *J* = 7.3 Hz, 1H), 3.98 (p, *J* = 7.2 Hz, 1H), 3.61 (s, 3H), 1.37 (s, 9H), 1.27 (d, *J* = 7.3 Hz, 3H), 1.16 ppm (d, *J* = 7.2 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.0, 172.7, 155.0, 77.9, 51.8, 49.2, 47.4, 28.2, 18.1, 17.0 ppm. HRMS (ESI) *m*/*z* calcd. for [$C_{12}H_{23}N_2O_5$]$^+$: 275.1607 [M + H]$^+$, found 275.1599, HRMS (ESI) *m*/*z* calcd. for [$C_{12}H_{22}N_2O_5Na$]$^+$: 297.1426 [M + Na]$^+$, found 297.1418.

_N_-Boc-2-aminoisobutyryl-2-aminoisobutyric acid methyl ester (18e)

*(a) N-Boc-2-aminoisobutyric acid (**16d**):*[32] 2-Aminoisobutyric acid (1.05 g, 10.2 mmol) was suspended in a 1:1 water/1,4-dioxane mixture (25 mL). Sodium hydroxide (0.79 g, 19.8 mmol) was added at 0 °C and the resulting mixture was stirred for 5 minutes. Boc anhydride (2.74 g, 12.6 mmol) was then added in portions and the resulting solution was stirred overnight at room temperature. The organic solvent (1,4-dioxane) was removed under reduced pressure and the remaining solution was washed with diethyl ether (3 x 10 mL). The aqueous layer was then slowly acidified with 1 M hydrochloric acid at 0 °C. The product was extracted with ethyl acetate (3 x 20 mL) and the combined organic extracts were dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure to give the desired product as a white solid (1.66 g, 80%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.15 (s, 1H), 7.02 (s, 1H), 1.36 (s, 9H), 1.29 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 176.1, 154.5, 77.7, 54.9, 28.2, 25.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_9H_{18}NO_4]^+$: 204.1236 $[M + H]^+$, found 204.1231, HRMS (ESI) *m*/*z* calcd. for $[C_9H_{17}NO_4Na]^+$: 226.1055 $[M + Na]^+$, found 226.1050.

*(b) N-Boc-2-aminoisobutyryl-2-aminoisobutyric acid methyl ester (**9c**):* Prepared according to GP4 using 1.17 g (5.76 mmol) of *N*-Boc-2-aminoisobutyric acid and 0.90 g (5.84 mmol) of 2-aminoisobutyric acid methyl ester hydrochloride salt, purified by column chromatography (7:3 Pet. ether/EtOAc, ninhydrin stain). A white solid, yield: 1.19 g (69%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.58 (s, 1H), 6.68 (s, 1H), 3.54 (s, 3H), 1.37 (s, 9H), 1.33 (s, 6H), 1.28 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.5, 174.0, 154.1, 78.0, 55.5, 55.1, 51.8, 28.1, 24.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{27}N_2O_5]^+$: 303.1915 $[M + H]^+$, found 303.1913, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{26}N_2O_5Na]^+$: 325.1734 $[M + Na]^+$, found 325.1732.

_N_-Boc-L-valyl-L-valine methyl ester (9e)

Prepared according to GP4, purified by column chromatography (3:1 Pet. ether/EtOAc, ninhydrin stain). A white solid, yield: 2.94 g (89%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.97 (d, J = 7.9 Hz, 1H), 6.69 (d, J = 9.1 Hz, 1H), 4.18 (dd, J = 8.0, 6.2 Hz, 1H), 3.86 (dd, J = 9.2, 7.0 Hz, 1H), 3.61 (s, 3H), 2.04 (hept, J = 6.8 Hz, 1H), 1.91 (hept, J = 6.8 Hz, 1H), 1.38 (s, 9H), 0.91–0.80 ppm (m, 12H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.83, 171.78, 155.36, 77.96, 59.49, 57.24, 51.58, 30.32, 29.86, 28.15, 19.12, 18.86, 18.21, 18.17 ppm. HRMS (ESI) m/z calcd. for $[C_{16}H_{31}N_2O_5]^+$: 331.2233 $[M + H]^+$, found 331.2227, HRMS (ESI) m/z calcd. for $[C_{16}H_{30}N_2O_5Na]^+$: 353.2052 $[M + Na]^+$, found 353.2046.

_N_-Boc-L-leucyl-L-leucine methyl ester (13e)

Prepared according to GP4, purified by column chromatography (7:3 Pet. ether/EtOAc, ninhydrin stain). A white solid, yield: 2.56 g (72%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.08 (d, J = 7.8 Hz, 1H), 6.82 (d, J = 8.5 Hz, 1H), 4.30 (ddd, J = 10.2, 7.8, 4.9 Hz, 1H), 3.97 (q, J = 7.9 Hz, 1H), 3.60 (s, 3H), 1.73–1.51 (m, 3H), 1.47 (ddd, J = 13.9, 9.1, 5.0 Hz, 1H), 1.37 (s, 11H), 0.88 (dd, J = 6.6, 4.0 Hz, 6H), 0.84 ppm (dd, J = 11.6, 6.3 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.94, 173.69, 156.55, 79.82, 53.89, 52.61, 51.51, 41.81, 41.20, 28.50, 25.40, 25.39, 23.23, 23.14, 21.97, 21.70 ppm. HRMS (ESI) m/z calcd. for $[C_{18}H_{35}N_2O_5]^+$: 359.2546 $[M + H]^+$, found 359.2541, HRMS (ESI) m/z calcd. for $[C_{18}H_{34}N_2O_5Na]^+$: 381.2365 $[M + Na]^+$, found 381.2360.

_N_-Boc-L-leucyl-L-leucyl-L-leucine methyl ester (15e)

*(a) L-Leucyl-L-leucine methyl ester trifluoroacetate salt (**13b**):* Prepared according to GP5 using 1.79 g (5.01 mmol) of *N*-Boc-L-leucyl-L-leucine methyl ester. A white powder, yield: 1.83 g (98%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.83 (d, J = 7.4 Hz, 1H), 8.20 (s, 3H), 4.34 (ddd, J = 9.9, 7.4, 5.0 Hz, 1H), 3.80 (t, J = 7.2 Hz, 1H), 3.63

(s, 3H), 1.77–1.64 (m, 2H), 1.65–1.46 (m, 4H) 0.94–0.84 ppm (m, 12H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.39, 170.20, 52.77, 52.76, 52.05, 41.21, 40.74, 25.31, 24.87, 23.05, 22.72, 22.27, 21.62 ppm. HRMS (ESI) *m/z* calcd. for $[C_{13}H_{27}N_2O_3]^+$: 259.2022 $[M - CF_3CO_2]^+$, found 259.2015.

*(b) N-Boc-L-leucyl-L-leucyl-L-leucine methyl ester (**15e**):* Prepared according to GP4 using 1.01 g (4.37 mmol) of *N*-acetyl-L-leucine and 1.64 g (4.40 mmol) of L-leucyl-L-leucine methyl ester trifluoroacetate salt, purified by column chromatography (3:2 Pet. ether/EtOAc, ninhydrin stain). A white solid, yield: 1.04 g (50%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.21 (d, *J* = 7.6 Hz, 1H), 7.72 (d, *J* = 8.5 Hz, 1H), 6.91, (d, *J* = 8.4 Hz, 1H), 4.36 (q, *J* = 7.8 Hz, 1H), 4.26 (ddd, *J* = 9.8, 7.6, 5.0 Hz, 1H), 3.93 (q, *J* = 7.9 Hz, 1H), 3.60 (s, 3H), 1.67–1.50 (m, 4H), 1.52–1.38 (m, 3H), 1.36 (s, 11H), 0.93–0.78 ppm (m, 18H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.7, 172.1, 172.0, 155.2, 78.0, 52.9, 51.8, 50.4, 50.1, 41.1, 40.6, 39.6, 28.1, 24.2, 24.1, 23.9, 23.0, 22.9, 22.8, 21.8, 21.7, 21.2 ppm. HRMS (ESI) *m/z* calcd. for $[C_{24}H_{46}N_3O_6]^+$: 472.3387 $[M + H]^+$, found 472.3382, HRMS (ESI) *m/z* calcd. for $[C_{24}H_{45}N_3O_6Na]^+$: 494.3206 $[M + Na]^+$, found 494.3197.

GP8. General procedure for the synthesis of N-pthaloylated amino acid methyl esters

*(a) Synthesis of the phthaloylating agent (**20**):*[33] Trimethyl phosphite (8.0 mL, 67.7 mmol) was added dropwise to 90% phthaloyl chloride (10 mL, 62.5 mmol) over 30 minutes, keeping the temperature below 50 °C. The mixture was stirred for 3 hours at room temperature and excess trimethyl phosphite was removed under reduced pressure to give a sticky oil. Petroleum ether (10 mL) was repeatedly added and removed under reduced pressure to induce precipitation. The resulting solid was

then washed with cold petroleum ether (2 x 10 mL) to give 3-chloro-3-(dimethoxyphosphoryl)isobenzofuran-1(3*H*)-one (**20**) as a white solid (16.1 g, 86%). This acylphosphonate derivative **20** was used without further purification to introduce the phthaloyl group in the amino acid.

(b) Synthesis of N-phthaloylated amino acids:[34] The acylphosphonate derivative **20** (1.38 g, 5.0 mmol) and the amino acid (5.3 mmol) were dissolved in a 1:1 acetonitrile/water mixture (25 mL). Diisopropylethylamine (3.5 mL, 20.1 mmol) was added dropwise and the mixture was stirred for 30 minutes at room temperature. Acetonitrile was removed under reduced pressure and the remaining solution was acidified with 1 M hydrochloric acid dropwise. The resulting precipitate was filtered and washed with ice-cold 1 M hydrochloric acid to give the N-Phth protected amino acid.

(c) Synthesis of N-Phthaloylated amino acid methyl esters:[35] (This method is an alternative to GP3) The N-phthaloylated amino acid (4.2 mmol) was suspended in dichloromethane (10 mL) and cooled to 0 °C. EDC·HCl (0.65 g, 4.2 mmol) and DMAP (0.50 g, 4.2 mmol) were then added, followed by the addition of methanol (0.17 mL, 4.2 mmol). The resulting mixture was stirred at room temperature for 4 hours. The solvent was removed under reduced pressure and ethyl acetate (40 mL) was added to the residue. The solution was washed sequentially with 10% aqueous potassium hydrogen sulfate (20 mL), 5% aqueous sodium bicarbonate (20 mL) and water (20 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow oil. The crude product was purified by column chromatography to give the N-phthaloylated amino acid methyl ester.

_N_-Phthaloylglycine methyl ester (1n)

*(a) N-Phthaloylglycine (**1m**):* Provided by Dr. Luke Gamon. ^{1}H NMR (400 MHz, DMSO-d_6): δ 13.26 (s, 1H), 7.97–7.89 (m, 2H), 7.91–7.84 (m, 2H), 4.32 ppm (s, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 168.9, 167.2, 134.8, 131.4, 123.4, 38.9 ppm.

*(b) N-Phthaloylglycine methyl ester (**1n**):* Prepared according to GP3 using 1.24 g (6.1 mmol) of *N*-phthaloylglycine. A white powder, yield: 1.21 g (91%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.98–7.90 (m, 2H), 7.94–7.86 (m, 2H), 4.45 (s, 2H), 3.70 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 168.0, 167.1, 134.9, 131.3, 123.5, 52.5, 38.7 ppm. HRMS (ESI) *m/z* calcd. for $[C_{11}H_{10}NO_4]^+$: 220.0610 $[M + H]^+$, found 220.0603, HRMS (ESI) *m/z* calcd. for $[C_{11}H_9NO_4Na]^+$: 242.0429 $[M + Na]^+$, found 242.0422.

_N_-Phthaloyl-2-aminoisobutyric acid methyl ester (16n)

*(a) N-Phthaloyl-2-aminoisobutyric acid (**16m**):* Prepared according to GP8(b). A white solid, yield: 0.54 g (47%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.92 (s, 1H), 7.85 (s, 4H), 1.72 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.9, 168.0, 134.7, 131.2, 123.0, 59.8, 24.3 ppm. HRMS (ESI) *m/z* calcd. for $[C_{12}H_{12}NO_4]^+$: 234.0761 $[M + H]^+$, found 234.0761, HRMS (ESI) *m/z* calcd. for $[C_{12}H_{11}NO_4Na]^+$: 256.0581 $[M + Na]^+$, found 256.0580.

*(b) N-Phthaloyl-2-aminoisobutyric acid methyl ester (**16n**):* Prepared according to GP8(c), purified by column chromatography (1:4 Pet. ether/$CHCl_3$, UV light). White crystals, yield: 0.47 g (46%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.86 (s, 4H), 3.65 (s, 3H), 1.72 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 167.8, 134.8, 131.0, 123.1, 59.6, 52.5, 24.1 ppm. HRMS (ESI) *m/z* calcd. for $[C_{13}H_{14}NO_4]^+$: 248.0918 $[M + H]^+$, found 248.0916, HRMS (ESI) *m/z* calcd. for $[C_{13}H_{13}NO_4Na]^+$: 270.0737 $[M + Na]^+$, found 270.0735.

_N_-Phthaloyl-L-valine methyl ester (3n)

*(a) N-Phthaloyl-L-valine (**3m**):* Prepared according to GP8(b). A white solid, yield: 0.52 g (42%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 13.03 (s, 1H), 8.00–7.84 (m, 4H), 4.45 (d, J = 7.9 Hz, 1H), 2.56 (dq, J = 14.5, 7.2 Hz, 1H), 1.07 (d, J = 6.6 Hz, 3 H), 0.82 ppm (d, J = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.8, 167.5, 135.0, 131.0, 123.5, 56.9, 28.0, 20.9, 19.3 ppm. HRMS (ESI) m/z calcd. for $[C_{13}H_{14}NO_4]^+$: 248.0918 $[M + H]^+$, found 248.0917, HRMS (ESI) m/z calcd. for $[C_{13}H_{13}NO_4Na]^+$: 270.0737 $[M + Na]^+$, found 270.0736.

*(b) N-Phthaloyl-L-valine methyl ester (**3n**):* Prepared according to GP3 using 0.49 g (1.97 mmol) of *N*-phthaloyl-L-valine, purified by column chromatography ($CHCl_3$, UV light). A colourless oil, yield: 0.41 g (81%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.97–7.86 (m, 4H), 4.59 (d, J = 7.7 Hz, 1H), 3.61 (s, 3H), 2.58 (h, J = 6.9 Hz, 1H), 1.05 (d, J = 6.6 Hz, 3H), 0.82 ppm (d, J = 6.8 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 168.8, 167.4, 135.1, 130.9, 123.6, 56.6, 52.3, 28.2, 20.5, 18.9 ppm. HRMS (ESI) m/z calcd. for $[C_{14}H_{16}NO_4]^+$: 262.1074 $[M + H]^+$, found 262.1074, HRMS (ESI) m/z calcd. for $[C_{14}H_{15}NO_4Na]^+$: 284.0894 $[M + Na]^+$, found 284.0894.

_N_-Phthaloyl-L-leucine methyl ester (4n)

*(a) N-Phthaloyl-L-leucine (**4m**):* Prepared according to GP8(b). A white solid, yield: 0.82 g (63%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 13.20 (s, 1H), 8.04–7.62 (m, 4H), 4.79 (dd, J = 11.5, 4.3 Hz, 1H), 2.17 (ddd, J = 13.8, 11.6, 4.1 Hz, 1H), 1.84 (ddd, J = 14.2, 10.2, 4.3 Hz, 1 H), 1.51–1.36 (m, 1H), 0.86 ppm (t, J = 6.7 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.9, 167.5, 135.0, 131.1 123.5, 50.1, 36.7, 24.7, 23.1, 20.8 ppm. HRMS (ESI) m/z calcd. for $[C_{14}H_{16}NO_4]^+$: 262.1074 $[M + H]^+$, found 262.1074, HRMS (ESI) m/z calcd. for $[C_{14}H_{15}NO_4Na]^+$: 284.0894 $[M + Na]^+$, found 284.0892.

*(b) N-Phthaloyl-L-leucine methyl ester (**4n**):* Prepared according to GP3 using 0.80 g (3.05 mmol) of *N*-phthaloyl-L-leucine, purified by column chromatography ($CHCl_3$, UV light). A colourless oil, yield: 0.59 g (70%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.97–7.86 (m, 4H), 4.92 (dd, *J* = 11.43, 4.3 Hz, 1H), 3.64 (s, 3H), 2.13 (ddd, *J* = 15.0, 11.4, 4.2 Hz, 1H), 1.86 (ddd, *J* = 14.2, 10.1, 4.4 Hz, 1H), 1.50–1.36 (m, 1H), 0.88 (d, *J* = 6.7 Hz, 3H), 0.85 ppm (d, *J* = 6.7 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.8, 167.3, 135.1, 131.0, 123.6, 52.7, 49.9, 36.8, 24.5, 23.0, 20.9 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{18}NO_4]^+$: 276.1231 [M + H]$^+$, found 276.1231, HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{17}NO_4Na]^+$: 298.1050 [M + Na]$^+$, found 298.1050.

N-Phthaloyl-2-aminoisobutyric acid methyl(d_3) ester (16p)

Prepared according to GP8(c) using 1.10 g (4.72 mmol) of *N*-phthaloyl-2-aminoisobutyric acid and methanol-d_4 (0.20 mL), purified by column chromatography (1:4 Pet. ether/$CHCl_3$, UV light). Colourless crystals, yield: 0.68 g (62%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.86 (s, 4H), 1.72 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 167.8, 134.8, 131.0, 123.1, 59.6, 24.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{13}H_{11}D_3NO_4]^+$: 251.1106 [M + H]$^+$, found 251.1105, HRMS (ESI) *m*/*z* calcd. for $[C_{13}H_{10}D_3NO_4Na]^+$: 273.0925 [M + Na]$^+$, found 273.0924.

N-(o-carbomethoxy)benzoyl-2-aminoisobutyric acid methyl ester (16o)

Prepared according to GP3 using 0.53 g (2.28 mmol) of *N*-phthaloyl-2-aminoisobutyric acid and 0.5 mL (6.88 mmol) of thionyl chloride, purified by column chromatography (99:1 $CHCl_3$/MeOH, UV light). A white solid, yield: 0.29 g (45%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.75 (s, 1H), 7.73 (d, *J* = 7.4 Hz, 1H), 7.62 (t, *J* = 7.5 Hz, 1H), 7.55 (t, *J* = 7.5 Hz, 1H), 7.46 (d, *J* = 7.3 Hz, 1H), 3.74 (s, 3H), 3.61 (s, 3H), 1.43 ppm (s, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.4, 167.5, 167.3,

137.5, 131.5, 129.9, 129.7, 129.1, 128.0, 55.4, 52.1, 51.9, 24.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{18}NO_5]^+$: 280.1180 $[M + H]^+$, found 280.1180, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{17}NO_5Na]^+$: 302.0999 $[M + Na]^+$, found 302.0999.

_N_-Phthaloylglycylglycine methyl ester (21n)

Prepared according to GP4, recrystallised from hot ethyl acetate. A white powder, yield: 1.91 g (69%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.71 (t, *J* = 5.9 Hz, 1H), 7.95–7.86 (m, 2H), 7.92–7.83 (m, 2H), 4.26 (s, 2H), 3.88 (d, *J* = 5.8 Hz, 2H), 3.63 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.0, 167.4, 166.7, 134.6, 131.7, 123.2, 51.8, 40.6, 39.9 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{13}H_{13}N_2O_5]^+$: 277.0825 $[M + H]^+$, found 277.0819, HRMS (ESI) *m*/*z* calcd. for $[C_{13}H_{12}N_2O_5Na]^+$: 299.0644 $[M + Na]^+$, found 299.0637.

_N_-Phthaloylglycyl-L-valine methyl ester (22n)

Prepared according to GP4 using 1.03 g (5.01 mmol) of *N*-phthaloylglycine and 0.85 g (5.06 mmol) of L-valine methyl ester hydrochloride salt, purified by column chromatography (3:2 Pet. ether/EtOAc, UV light). A white powder, yield: 1.40 g (89%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.62 (d, *J* = 8.3 Hz, 1H), 7.94–7.86 (m, 2H), 7.91–7.83 (m, 2H), 4.30 (s, 2H), 4.20 (dd, *J* = 8.3, 6.4 Hz, 1H), 3.65 (s, 3H), 2.09–1.99 (m, 1H), 0.88 ppm (t, *J* = 7.3 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.7, 167.5, 166.4, 134.6, 131.7, 123.2, 57.5, 51.8, 39.8, 30.1, 18.9, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{16}H_{19}N_2O_5]^+$: 319.1294 $[M + H]^+$, found 319.1286, HRMS (ESI) *m*/*z* calcd. for $[C_{16}H_{18}N_2O_5Na]^+$: 341.1113 $[M + Na]^+$, found 341.1106.

_N_-Phthaloylglycyl-L-leucine methyl ester (23n)

Prepared according to GP4 using 1.02 g (4.99 mmol) of *N*-phthaloylglycine and 0.91 g (5.03 mmol) of L-leucine methyl ester hydrochloride salt, purified by column chromatography (3:2 Pet. ether/EtOAc, UV light). A white powder, yield: 1.01 g (61%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.66 (d, J = 7.7 Hz, 1H), 7.93–7.89 (m, 2H), 7.88–7.85 (m, 2H), 4.37–4.26 (m, 1H), 4.25 (s, 2H), 3.63 (s, 3H), 1.66–1.61 (m, 1H), 1.57–1.44 (m, 2H),0.90 (d, J = 5.3 Hz, 3H), 0.84 ppm (d, J = 5.4 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.6, 167.4, 166.2, 134.6, 131.7, 123.2, 51.9, 50.4, 39.9, 39.8, 24.2, 22.7, 21.3 ppm. HRMS (ESI) m/z calcd. for $[C_{17}H_{21}N_2O_5]^+$: 333.1451 [M + H]$^+$, found 333.1443, HRMS (ESI) m/z calcd. for $[C_{17}H_{20}N_2O_5Na]^+$: 355.1270 [M + Na]$^+$, found 355.1261.

GP9. General procedure for the synthesis of N-acetylated amino acid amides[36]

The N-acetylated amino acid (10.0 mmol) was dissolved in anhydrous tetrahydrofuran (20 mL) at 0 °C. Triethylamine (4.2 mL, 30.1 mmol), followed by isobutyl chloroformate (2.6 mL, 20.0 mmol), was added dropwise. The resulting suspension was stirred for 30 minutes at 0 °C and 30% aqueous ammonia (2 mL, 30.0 mmol) was then added. The reaction mixture was stirred for 30 minutes at room temperature. The solvent was removed under reduced pressure and the crude product was purified by column chromatography to give the N-acetylated amino acid amide.

GP10. General procedure for the synthesis of N-acetylated amino acid *t*-butyl amides[37]

The N-acetylated amino acid (42.4 mmol) and *p*-nitrophenol (5.89 g, 42.4 mmol) were dissolved in dichloromethane (75 mL). DCC (8.79 g, 42.6 mmol) was added in portions. The resulting mixture was heated at 40 °C for 2 hours and filtered through a bed of celite. The filter cake was washed with dichloromethane (20 mL) and the filtrate was concentrated under reduced pressure. The crude product was recrystallised from hot ethyl acetate (60 mL) to give the *p*-nitroester intermediate as a yellow solid. The solid was then dissolved in dichloromethane (110 mL) and *t*-butylamine (7.26 mL, 69.1 mmol) was added dropwise. The resulting mixture was stirred at room temperature for 40 minutes and then filtered through a bed of celite. The filtrate was concentrated under reduced pressure and the crude product was recrystallised from hot ethanol or purified by column chromatography to give the N-acetylated amino acid *t*-butyl amide.

***N*-Acetyl-L-alanine amide (2q)**

Prepared according to GP9, purified by column chromatography (9:1 $CHCl_3$/MeOH, $KMnO_4$ stain). A white solid, yield: 0.24 g (18%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.93 (d, J = 7.6 Hz, 1H), 7.29 (s, 1H), 6.92 (s, 1H), 4.17 (p, J = 7.1 Hz, 1H), 1.82 (s, 3H), 1.17 ppm (d, J = 7.0 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.5, 168.9, 47.9, 22.6, 18.3 ppm. HRMS (ESI) m/z calcd. for $[C_5H_{11}N_2O_2]^+$: 131.0816 $[M + H]^+$, found 131.0814, HRMS (ESI) m/z calcd. for $[C_5H_{10}N_2O_2Na]^+$: 153.0635 $[M + Na]^+$, found 153.0634.

_N_-Acetyl-L-leucine amide (4q)

Prepared by Julius Kerstien according to GP9. A white solid, yield: 0.36 g (21%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 7.87 (d, J = 8.3 Hz, 1H), 7.31 (s, 1H), 6.90 (s, 1H), 4.20 (td, J = 8.7, 5.9 Hz, 1H), 1.83 (s, 3H), 1.62–1.52 (m, 1H), 1.47–1.34 (m, 2H), 0.87 (d, J = 6.6 Hz, 3H), 0.83 ppm (d, J = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.4, 169.0, 50.7, 41.0, 24.3, 23.0, 22.6, 21.6 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{17}N_2O_2]^+$: 173.1285 $[M + H]^+$, found 173.1286, HRMS (ESI) m/z calcd. for $[C_8H_{16}N_2O_2Na]^+$: 195.1104 $[M + Na]^+$, found 195.1105.

_N_-Acetylglycine _t_-butyl amide (1r)

Prepared according to GP10, purified by column chromatography (49:1 EtOAc/MeOH, $KMnO_4$ stain). Colourless crystals, yield: 0.12 g (2%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.92 (t, J = 5.6 Hz, 1H), 7.40 (s, 1H), 3.59 (d, J = 5.6 Hz, 2H), 1.83 (s, 3H), 1.24 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.3, 168.1, 50.0, 42.2, 28.5, 22.5 ppm. HRMS (ESI) m/z calcd. for $[C_8H_{17}N_2O_2]^+$: 173.1285 $[M + H]^+$, found 173.1286, HRMS (ESI) m/z calcd. for $[C_8H_{16}N_2O_2Na]^+$: 195.1104 $[M + Na]^+$, found 195.1105.

_N_-Acetyl-L-leucine _t_-butyl amide (4r)

Prepared according to GP10 using 1.75 g (10.1 mmol) of _N_-acetyl-L-leucine, recrystallised from hot ethanol. Colourless needles, yield: 0.17 g (7%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.83 (d, J = 8.5 Hz, 1H), 7.48 (s, 1H), 4.25 (td, J = 8.6, 6.3 Hz, 1H), 1.82 (s, 3H), 1.60–1.46 (m, 1H), 1.44–1.28 (m, 2H), 1.23 (s, 9H), 0.87 (d, J = 6.6 Hz, 3H), 0.83 ppm (d, J = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.6, 168.8, 51.2, 50.0, 41.5, 28.5, 24.3, 23.0, 22.5, 21.9 ppm. HRMS (ESI) m/z calcd. for $[C_{12}H_{25}N_2O_2]^+$: 229.1911 $[M + H]^+$, found 229.1916, HRMS (ESI) m/z calcd. for $[C_{12}H_{24}N_2O_2Na]^+$: 251.1730 $[M + Na]^+$, found 251.1736

_N_-Acetyl-L-phenylalanine methyl ester (24c)

*(a) N-Acetyl-L-phenylalanine (**24a**):* Prepared according to GP1 using 11.2 g (67.9 mmol) of L-phenylalanine. A white solid, yield: 12.7 g (90%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.65 (s, 1H), 8.18 (d, J = 8.1 Hz, 1H), 7.31–7.25 (m, 2H), 7.24–7.17 (m, 3H), 4.40 (ddd, J = 9.6, 8.1, 4.9 Hz, 1H), 3.03 (dd, J = 13.8, 4.9 Hz, 1H), 2.83 (dd, J = 13.8, 9.6 Hz, 1H), 1.77 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.2, 169.2, 137.7, 129.0, 128.2, 126.4, 53.5, 36.8, 22.3 ppm. HRMS (ESI) m/z calcd. for $[C_{11}H_{14}NO_3]^+$: 208.0978 [M + H]$^+$, found 208.0975, HRMS (ESI) m/z calcd. for $[C_{22}H_{27}N_2O_6]^+$: 415.1869 [2M + H]$^+$, found 415.1893.

*(b) N-Acetyl-L-phenylalanine methyl ester (**24c**):* Prepared according to GP3. White crystals, yield: 1.84 g (70%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.33 (d, J = 7.8 Hz, 1H), 7.33–7.23 (m, 2H), 7.26–7.16 (m, 3H), 4.44 (ddd, J = 9.1, 7.7, 5.6 Hz, 1H), 3.59 (s, 3H), 3.00 (dd, J = 13.8, 5.6 Hz, 1H), 2.87 (dd, J = 13.7, 9.3 Hz, 1H), 1.79 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 169.3, 137.3, 129.0, 128.2, 126.5, 53.6, 51.8, 36.7, 22.2 ppm. HRMS (ESI) m/z calcd. for $[C_{12}H_{16}NO_3]^+$: 222.1130 [M + H]$^+$, found 222.1137, HRMS (ESI) m/z calcd. for $[C_{12}H_{15}NO_3Na]^+$: 244.0950 [M + Na]$^+$, found 244.0961.

_N_-Acetyl-L-phenylalanine-(β,β-d_2) methyl ester (26c)

*(a) L-phenylalanine-(β,β-d_2) (**26**):*[38] L-Phenylalanine (1.65 g, 10.0 mmol) and 10% wt. palladium on carbon (0.17 g) were suspended in deuterated water (20 mL) under hydrogen atmosphere at 110 °C for 6 hours. The reaction mixture was then filtered through a PTFE membrane filter (0.45 μm, Merck) and the filtrate was concentrated under reduced pressure. Water (10 mL) was added to the residue and the resulting solution was lyophilised to give the desired product as a colourless solid (1.14 g, 68%). ^{1}H NMR (400 MHz, D_2O): δ 7.49–7.35 (m, 3H), 7.37–7.28 (m,

2H), 3.99 ppm (s, 1H). ^{13}C {^{1}H} NMR (101 MHz, D_2O/1% MeOH): δ 174.5, 135.6, 130.0, 129.7, 128.3, 56.5 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_9H_{10}D_2NO_2]^+$: 168.0994 $[M + H]^+$, found 168.0989.

*(b) N-Acetyl-L-phenylalanine-(β,β-d_2) (**26a**):* Prepared according to GP1 using 1.46 g (8.75 mmol) of L-phenylalanine-(β,β-d_2). A white solid, yield: 1.53 g (84%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.64 (s, 1H), 8.17 (d, *J* = 8.0 Hz, 1H), 7.32–7.23 (m, 2H), 7.26–7.15 (m, 3H), 4.38 (d, *J* = 8.0 Hz, 1H), 1.77 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.2, 169.2, 137.6, 129.0, 128.2, 126.4, 53.4, 22.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{11}H_{12}D_2NO_3]^+$: 210.1099 $[M + H]^+$, found 210.1095, HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{23}D_4N_2O_6]^+$: 419.2120 $[2M + H]^+$, found 419.2116.

*(c) N-Acetyl-L-phenylalanine-(β,β-d_2) methyl ester (**26c**):* Prepared according to GP3 using 1.51 g (7.23 mmol) of *N*-acetyl-L-phenylalanine-(β,β-d_2). White crystals, yield: 1.38 g (85%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.33 (d, *J* = 7.7 Hz, 1H), 7.33–7.24 (m, 2H), 7.25–7.16 (m, 3H), 4.43 (d, *J* = 7.7 Hz, 1H), 3.59 (s, 3H), 1.79 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 169.3, 137.2, 129.0, 128.2, 126.5, 53.5, 51.8, 22.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{14}D_2NO_3]^+$: 224.1256 $[M + H]^+$, found 224.1248, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{27}D_4N_2O_6]^+$: 447.2433 $[M + Na]^+$, found 447.2423.

_N_-Acetylglycyl-L-phenylalanine methyl ester (27c)

*(a) L-Phenylalanine methyl ester hydrochloride salt (**24b**)*: Prepared according to GP2. A white solid, yield: 16.8 g (99%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.74 (s, 3H), 7.38–7.28 (m, 2H), 7.32–7.20 (m, 3H), 4.23 (dd, *J* = 7.4, 5.9 Hz, 1H), 3.65 (s, 3H), 3.21 (dd, *J* = 14.0, 5.7 Hz, 1H), 3.10 ppm (dd, *J* = 14.0, 7.5 Hz, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.3, 134.7, 129.4, 128.6, 127.2, 53.2, 52.5, 35.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{10}H_{14}NO_2]^+$: 180.1019 $[M - Cl]^+$, found 180.1010.

*(b) N-Acetylglycyl-L-phenylalanine methyl ester (**27c**):* Prepared according to GP4, purified by column chromatography (99:1 EtOAc/AcOH, Hanessian's stain). White crystals, yield: 1.72 g (61%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.30 (d, J = 7.7 Hz, 1H), 8.03 (t, J = 5.9 Hz, 1H), 7.33–7.23 (m, 2H), 7.26–7.16 (m, 3H), 4.47 (td, J = 8.2, 5.7 Hz, 1H), 3.70 (dd, J = 16.7, 5.9 Hz, 1H), 3.63 (dd, J = 16.7, 5.9 Hz, 1H), 3.59 (s, 3H), 3.02 (dd, J = 13.7, 5.7 Hz, 1H), 2.91 (dd, J = 13.7, 8.8 Hz, 1H), 1.83 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.9, 169.5, 169.1, 137.0, 129.1, 128.3, 126.6, 53.6, 51.9, 41.6, 36.7, 22.4 ppm. HRMS (ESI) m/z calcd. for $[C_{14}H_{19}N_2O_4]^+$: 279.1345 $[M + H]^+$, found 279.1341, HRMS (ESI) m/z calcd. for $[C_{14}H_{18}N_2O_4Na]^+$: 301.1164 $[M + Na]^+$, found 301.1161.

_N_-Acetyl-L-valyl-L-phenylalanine methyl ester (28c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white solid, yield: 1.44 g (45%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.39 (d, J = 7.3 Hz, 1H), 7.79 (d, J = 9.0 Hz, 1H), 7.31–7.22 (m, 2H), 7.25–7.15 (m, 3H), 4.45 (dt, J = 8.5, 6.5 Hz, 1H), 4.18 (dd, J = 9.1, 7.0 Hz, 1H), 3.55 (s, 3H), 3.01 (dd, J = 13.9, 6.1 Hz, 1H), 2.93 (dd, J = 13.9, 8.8 Hz, 1H), 1.91 (hept, J = 7.0 Hz, 1H), 1.83 (s, 3H), 0.82 (d, J = 7.0 Hz, 3H), 0.79 ppm (d, J = 7.0 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.8, 171.3, 169.0, 137.1, 129.0, 128.2, 126.5, 57.2, 53.5, 51.7, 36.5, 30.6, 22.4, 19.1, 18.1 ppm. HRMS (ESI) m/z calcd. for $[C_{17}H_{25}N_2O_4]^+$: 321.1814 $[M + H]^+$, found 321.1814, HRMS (ESI) m/z calcd. for $[C_{17}H_{24}N_2O_4Na]^+$: 343.1633 $[M + Na]^+$, found 343.1638.

_N_-Acetyl-L-leucyl-L-phenylalanine methyl ester (29c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain) and recrystallised from hot ethyl acetate. A white solid, yield: 1.27 g (38%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.30 (d, J = 7.4 Hz, 1H), 7.90

(d, *J* = 8.4 Hz, 1H), 7.29–7.24 (m, 2H), 7.24–7.18 (m, 3H), 4.48–4.40 (m, 1H), 4.32 (td, *J* = 8.7, 6.0 Hz, 1H), 3.56 (s, 3H), 3.02 (dd, *J* = 13.9, 5.9 Hz, 1H), 2.94 (dd, *J* = 13.9, 8.6 Hz, 1H), 1.80 (s, 3H), 1.61–1.49 (m, 1H), 1.44–1.28 (m, 2H), 0.87 (d, *J* = 6.6 Hz, 3H), 0.82 ppm (d, *J* = 6.5 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.3, 171.8, 168.9, 137.1, 129.0, 128.2, 126.5, 53.5, 51.8, 50.5, 40.9, 36.4, 24.1, 22.9, 22.4, 21.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{27}N_2O_4]^+$: 335.1965 $[M + H]^+$, found 335.1976, HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{26}N_2O_4Na]^+$: 357.1785 $[M + Na]^+$, found 357.1790.

_N_-Acetyl-L-phenylalanylglycine methyl ester (30c)

Prepared according to GP4, recrystallised from hot ethyl acetate. A white solid, yield: 1.07 g (39%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.45 (t, *J* = 5.9 Hz, 1H), 8.13 (d, *J* = 8.5 Hz, 1H), 7.35–7.13 (m, 5H), 4.53 (td, *J* = 9.4, 4.3 Hz, 1H), 3.85 (d, *J* = 5.9 Hz, 2H), 3.63 (s, 3H), 3.02 (dd, *J* = 13.9, 4.3 Hz, 1H), 2.72 (dd, *J* = 13.8, 10.2 Hz, 1H), 1.75 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.0, 170.2, 169.1, 138.1, 129.1, 128.0, 126.2, 53.7, 51.7, 40.6, 37.6, 22.5 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{19}N_2O_4]^+$: 279.1345 $[M + H]^+$, found 279.1351, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{18}N_2O_4Na]^+$: 301.1164 $[M + Na]^+$, found 301.1169.

_N_-Acetyl-L-phenylalanyl-L-valine methyl ester (31c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white solid, yield: 1.03 g (32%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.24 (d, *J* = 8.1 Hz, 1H), 8.07 (d, *J* = 8.3 Hz, 1H), 7.28–7.24 (m, 4H), 7.24–7.13 (m, 1H), 4.62 (ddd, *J* = 10.0, 8.4, 4.4 Hz, 1H), 4.17 (dd, *J* = 8.1, 6.3 Hz, 1H), 3.62 (s, 3H), 2.96 (dd, *J* = 13.9, 4.4 Hz, 1H), 2.71 (dd, *J* = 13.9, 10.1 Hz, 1H), 2.04 (h, *J* = 6.8 Hz, 1H), 1.74 (s, 3H), 0.90 (d, *J* = 6.9 Hz, 3H), 0.87 ppm (d, *J* = 6.8 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.9, 171.8, 169.1, 137.9, 129.2, 128.0, 126.2,

57.4, 53.5, 51.7, 37.5, 29.9, 22.4, 18.9, 18.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{25}N_2O_4]^+$: 321.1814 $[M + H]^+$, found 321.1809, HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{24}N_2O_4Na]^+$: 343.1633 $[M + Na]^+$, found 343.1628.

_N_-Acetyl-L-phenylalanyl-L-leucine methyl ester (32c)

Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white solid, yield: 2.68 g (81%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.37 (d, *J* = 7.7 Hz, 1H), 8.06 (d, *J* = 8.4 Hz, 1H), 7.26 (d, *J* = 4.4 Hz, 4H), 7.19 (p, *J* = 4.5 Hz, 1H), 4.54 (ddd, *J* = 10.1, 8.5, 4.3 Hz, 1H), 4.29 (ddd, *J* = 9.6, 7.7, 5.3 Hz, 1H), 3.61 (s, 3H), 2.99 (dd, *J* = 13.9, 4.3 Hz, 1H), 2.71 (dd, *J* = 13.9, 10.0 Hz, 1H), 1.74 (s, 3H), 1.68–1.57 (m, 1H), 1.60–1.44 (m, 2H), 0.90 (d, *J* = 6.4 Hz, 3H), 0.84 ppm (d, *J* = 6.3 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 171.7, 169.0, 138.0, 129.1, 128.0, 126.2, 53.6, 51.8, 50.3, 39.7, 37.6, 24.2, 22.7, 22.4, 21.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{27}N_2O_4]^+$: 335.1965 $[M + H]^+$, found 335.1975, HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{26}N_2O_4Na]^+$: 357.1785 $[M + Na]^+$, found 357.1784.

_N_-Acetyl-L-phenylalanyl-L-glutamic acid 1,5-dimethyl ester (33c)

*(a) L-Glutamic acid 1,5-dimethyl ester hydrochloride salt (**6b'**)*: Prepared according to GP2 using 5.66 g (38.4 mmol) of L-glutamic acid and 9.0 mL (123 mmol) of thionyl chloride. White crystals, yield: 7.52 g (92%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.86–8.60 (m, 3H), 4.37–4.31 (m, 1H), 3.81 (s, 3H), 3.65 (s, 3H), 2.81–2.67 (m, 1H), 2.62 (dt, *J* = 17.0, 7.1 Hz, 1H), 2.41 ppm (q, *J* = 7.3 Hz, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.2, 169.5, 52.8, 51.2, 51.1, 28.9, 25.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_7H_{14}NO_4]^+$: 176.0923 $[M - Cl]^+$, found 176.0918.

*(b) N-Acetyl-L-phenylalanyl-L-glutamic acid 1,5-dimethyl ester (**33c**):* Prepared according to GP4, purified by column chromatography (1:9 Pet. ether/EtOAc, PMA stain). A white solid, yield: 1.93 g (53%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.42 (d,

J = 7.6 Hz, 1H), 8.09 (d, *J* = 8.4 Hz, 1H), 7.30–7.22 (m, 4H), 7.22–7.15 (m, 1H), 4.52 (ddd, *J* = 10.0, 8.4, 4.4 Hz, 1H), 4.29 (ddd, *J* = 9.1, 7.5, 5.3 Hz, 1H), 3.61 (s, 3H), 3.60 (s, 3H), 2.98 (dd, *J* = 13.9, 4.4 Hz, 1H), 2.72 (dd, *J* = 13.9, 10.0 Hz, 1H), 2.37 (td, *J* = 7.4, 6.8 Hz, 2H), 2.07–1.96 (m, 1H), 1.91–1.79 (m, 1H), 1.74 ppm (s, 3H). ^{13}C {^{1}H} NMR (126 MHz, DMSO-d_6): δ 172.6, 171.9, 171.7, 169.1, 137.9, 129.1, 128.0, 126.2, 53.7, 51.9, 51.4, 51.1, 37.5, 29.6, 26.0, 22.4 ppm. HRMS (ESI) *m/z* calcd. for $[C_{18}H_{25}N_2O_6]^+$: 365.1713 $[M + H]^+$, found 365.1710, HRMS (ESI) *m/z* calcd. for $[C_{18}H_{24}N_2O_6Na]^+$: 387.1532 $[M + Na]^+$, found 387.1527.

_N_-Acetyl-L-phenylalanyl-L-phenylalanine methyl ester (34c)

Prepared according to GP4, recrystallised from hot ethyl acetate. A white solid, yield: 2.33 g (64%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.44 (d, *J* = 7.5 Hz, 1H), 8.03 (d, *J* = 8.5 Hz, 1H), 7.30–7.16 (m, 10H), 4.57–4.51 (m, 1H), 4.51–4.45 (m, 1H), 3.58 (s, 3H), 3.04 (dd, *J* = 13.8, 5.9 Hz, 1H), 2.95 (dd, *J* = 13.0, 5.3 Hz, 1H), 2.94 (dd, *J* = 14.5, 7.7 Hz, 1H), 2.67 (dd, *J* = 13.9, 10.0 Hz, 1H), 1.71 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.71, 171.54, 168.92, 137.89, 137.02, 129.10, 129.06, 128.24, 127.96, 126.55, 126.18, 53.57, 53.45, 51.82, 37.49, 36.55, 22.39 ppm. HRMS (ESI) *m/z* calcd. for $[C_{21}H_{25}N_2O_4]^+$: 369.1814 $[M + H]^+$, found 369.1804, HRMS (ESI) *m/z* calcd. for $[C_{21}H_{24}N_2O_4Na]^+$: 391.1634 $[M + Na]^+$, found 391.1604.

_N_-Acetyl-L-phenylalanine amide (24q)[39]

L-Phenylalanine methyl ester hydrochloride salt (2.40 g, 11.1 mmol) was dissolved in 28% aqueous ammonia (25 mL) and stirred overnight at room temperature. The solvent was then removed under reduced pressure to give L-phenylalanine amide as an off-white solid. L-Phenylalanine amide (0.56 g, 3.39 mmol) was dissolved in acetonitrile (5 mL) and acetic anhydride (1.0 mL, 10.6 mmol) was added dropwise. The resulting mixture was stirred for 1 day at room temperature and filtered to give

the desired product as a white solid (0.06 g, 9%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.01 (d, J = 8.5 Hz, 1H), 7.42 (s, 1H), 7.29–7.21 (m, 4H), 7.22–7.13 (m, 1H), 7.01 (s, 1H), 4.41 (td, J = 9.2, 4.7 Hz, 1H), 2.98 (dd, J = 13.7, 4.7 Hz, 1H), 2.72 (dd, J = 13.7, 9.7 Hz, 1H), 1.75 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 173.2, 169.0, 138.2, 129.1, 128.0, 126.2, 53.8, 37.7, 22.5 ppm. HRMS (ESI) m/z calcd. for $[C_{11}H_{15}N_2O_2]^+$: 207.1134 $[M + H]^+$, found 207.1130, HRMS (ESI) m/z calcd. for $[C_{11}H_{14}N_2O_4Na]^+$: 229.0953 $[M + Na]^+$, found 229.0949.

_N_-Acetyl-L-phenylalanine methyl amide (24s)[40]

_N_-Acetyl-L-phenylalanine (2.10 g, 10.1 mmol) was suspended in a 1:1 dimethylformamide/tetrahydrofuran mixture (14 mL) and cooled to –10 °C. _N_-Methylmorpholine (1.2 mL, 10.9 mmol), followed by pivaloyl chloride (1.25 mL, 10.2 mmol), was added dropwise. After 10 minutes, a solution of methylamine hydrochloride (3.36 g, 49.8 mmol) and triethylamine (7.0 mL, 50.2 mmol) in water (7 mL) was added. The resulting mixture was stirred for 1.5 hours and the solvent was removed under reduced pressure. The residue was dissolved in chloroform (30 mL) and washed with water (2 x 30 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow oil. The crude product was purified by column chromatography (99:1 DCM/MeOH, PMA stain) to give the desired product as a white solid (0.39 g, 18%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.10 (d, J = 8.5 Hz, 1H), 7.88 (q, J = 4.6 Hz, 1H), 7.30–7.21 (m, 2H), 7.24–7.13, (m, 3H), 4.40 (ddd, J = 9.5, 8.4, 5.0 Hz, 1H), 2.95 (dd, J = 13.7, 5.0 Hz, 1H), 2.72 (dd, J = 13.7, 9.6 Hz, 1H), 2.56 (d, J = 4.5 Hz, 3H), 1.75 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.6, 169.0, 138.2, 129.0, 128.0, 126.2, 54.1, 37.8, 25.5, 22.5 ppm. HRMS (ESI) m/z calcd. for $[C_{12}H_{17}N_2O_2]^+$: 221.1290 $[M + H]^+$, found 221.1283, HRMS (ESI) m/z calcd. for $[C_{24}H_{33}N_4O_4]^+$: 441.2502 $[2M + H]^+$, found 441.2495.

_N_-Acetyl-L-phenylalanine _t_-butyl amide (24r)

Prepared according to GP10 using 3.50 g (16.9 mmol) of *N*-acetyl-L-phenylalanine, recrystallised from hot ethanol. White crystals, yield: 0.19 g (4%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.96 (d, *J* = 8.4 Hz, 1H), 7.49 (s, 1H), 7.29–7.18 (m, 4H), 7.22–7.13, (m, 1H), 4.46 (td, *J* = 8.7, 6.1 Hz, 1H), 2.85 (dd, *J* = 13.5, 5.7 Hz, 1H), 2.72 (dd, *J* = 13.5, 8.9 Hz, 1H), 1.75 (s, 3H), 1.19 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 170.5, 168.8, 137.9, 129.3, 127.9, 126.1, 54.1, 50.0, 38.3, 28.4, 22.5 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{23}N_2O_2]^+$: 263.1760 $[M + H]^+$, found 263.1752, HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{22}N_2O_2K]^+$: 301.1318 $[M + K]^+$, found 301.1310.

_N_-Acetyl-L-valyl-L-valyl-L-phenylalanine methyl ester (39c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35).

*(a) N-Boc-L-valyl-L-phenylalanine methyl ester (**28e**):* Prepared according to GP4, purified by column chromatography (7:3 Pet. ether/EtOAc, ninhydrin stain). A white powder, yield: 3.15 g (83%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.25 (d, *J* = 7.6 Hz, 1H), 7.30–7.22 (m, 2H), 7.25–7.15 (m, 3H), 6.56 (d, *J* = 9.1 Hz, 1H), 4.49 (ddd, *J* = 9.4, 7.7, 5.8 Hz, 1H), 3.78 (dd, *J* = 9.0, 7.3 Hz, 1H), 3.57 (s, 3H), 3.02 (dd, *J* = 13.9, 5.7 Hz, 1H), 2.92 (dd, *J* = 13.9, 8.9 Hz, 1H), 1.93–1.77 (m, 1H), 1.37 (s, 9H), 0.77 (d, *J* = 6.6 Hz, 3H), 0.76 ppm (d, *J* = 6.7 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.8, 171.4, 155.2, 137.1, 129.0, 128.2, 126.5, 78.0, 59.5, 53.4, 51.7, 36.7, 30.5, 28.2, 19.0, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{31}N_2O_5]^+$: 379.2233 $[M + H]^+$, found 379.2222, HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{30}N_2O_5Na]^+$: 401.2052 $[M + Na]^+$, found 401.2052.

*(b) L-Valyl-L-phenylalanine methyl ester trifluoroacetate salt (**28b**):* Prepared according to GP5 using 2.95 g (7.78 mmol) of *N*-Boc-L-valyl-L-phenylalanine

methyl ester. A white powder, yield: 2.85 g (93%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.90 (d, *J* = 7.1 Hz, 1H), 8.10 (s, 3H), 7.35–7.26 (m, 2H), 7.28–7.19 (m, 3H), 4.55 (dt, *J* = 7.4, 6.0 Hz, 1H), 3.65 (d, *J* = 5.1 Hz, 1H), 3.59 (s, 3H), 3.07 (dd, *J* = 14.0, 5.9 Hz, 1H), 2.98 (dd, *J* = 14.0, 8.5 Hz, 1H), 2.18–2.04 (m, 1H), 0.95 (d, *J* = 6.9 Hz, 3H) 0.91 ppm (d, *J* = 6.9 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.3, 168.2, 136.8, 129.0, 128.4, 126.7, 57.0, 53.9, 51.9, 36.4, 29.9, 18.3, 17.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{23}N_2O_3]^+$: 279.1709 [M – CF_3CO_2]$^+$, found 279.1719.

*(c) N-Acetyl-L-valyl-L-valyl-L-phenylalanine methyl ester (**39c**):* Prepared according to GP4 using 0.73 g (4.57 mmol) of *N*-acetyl-L-valine and 1.82 g (4.64 mmol) of L-valyl-L-phenylalanine methyl ester trifluoroacetate salt, recrystallised from hot methanol. A white solid, yield: 0.14 g (7%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.34 (d, *J* = 7.4 Hz, 1H), 7.87 (d, *J* = 8.8 Hz, 1H), 7.66 (d, *J* = 9.0 Hz, 1H), 7.25 (t, *J* = 7.4 Hz, 2H), 7.20 (d, *J* = 7.2 Hz, 3H), 4.49 (ddd, *J* = 8.7, 7.2, 5.7 Hz, 1H), 4.16 (dd, *J* = 8.9, 6.9 Hz, 2H), 3.55 (s, 3H), 3.02 (dd, *J* = 14.0, 5.8 Hz, 1H), 2.92 (dd, *J* = 14.0, 8.9 Hz, 1H), 1.90 (dt, *J* = 13.8, 6.9 Hz, 2H), 1.85 (s, 3H), 0.80 (d, *J* = 7.7 Hz, 6H), 0.77 ppm (d, *J* = 6.9 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.7, 170.9, 169.2, 137.0, 128.9, 128.2, 126.5, 57.7, 57.2, 53.3, 51.7, 36.5, 30.8, 30.2, 22.5, 19.2, 19.0, 18.2, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{34}N_3O_5]^+$: 420.2499 [M + H]$^+$, found 420.2496, HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{33}N_3O_5Na]^+$: 442.2318 [M + Na]$^+$, found 442.2314.

_N_-Acetyl-L-valyl-L-phenylalanyl-L-valine methyl ester (40c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35).

*(a) N-Boc-L-phenylalanyl-L-valine methyl ester (**31e**):* Prepared according to GP4, purified by column chromatography (7:3 Pet. ether/EtOAc, ninhydrin stain). A white powder, yield: 3.27 g (87%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.08 (d, *J* =

8.3 Hz, 1H), 7.30–7.24 (m, 4H), 7.23–7.15 (m, 1H), 6.93 (d, *J* = 8.6 Hz, 1H), 4.25 (ddd, *J* = 23.0, 9.3, 5.2 Hz, 2H), 3.63 (s, 3H), 2.94 (dd, *J* = 13.9, 4.3 Hz, 1H), 2.73 (dd, *J* = 13.8, 10.4 Hz, 1H), 2.05 (h, *J* = 6.7 Hz, 1H), 1.30 (s, 9H), 0.89 ppm (dd, *J* = 9.0, 6.8 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.1, 171.9, 155.2, 138.1, 129.2, 128.0, 126.1, 78.0, 57.2, 55.4, 51.7, 37.2, 30.1, 28.1, 18.9, 18.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{31}N_2O_5]^+$: 379.2233 $[M + H]^+$, found 379.2228, HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{30}N_2O_5Na]^+$: 401.2052 $[M + Na]^+$, found 401.2048.

*(b) L-Phenylalanyl-L-valine methyl ester trifluoroacetate salt (**31b**):* Prepared according to GP5 using 2.49 g (6.57 mmol) of *N*-Boc-L-phenylalanyl-L-valine methyl ester. A white powder, yield: 2.28 g (88%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.73 (d, *J* = 8.1 Hz, 1H), 8.25 (s, 3H), 7.32 (dd, *J* = 7.8, 6.1 Hz, 2H), 7.32–7.21 (m, 3H), 4.21 (dd, *J* = 8.1, 6.3 Hz, 1H), 4.16 (t, *J* = 6.8 Hz, 1H), 3.63 (s, 3H), 3.08 (dd, *J* = 14.0, 6.1 Hz, 1H), 2.97 (dd, *J* = 14.0, 7.4 Hz, 1H), 2.11–1.97 (m, 1H), 0.90 ppm (t, *J* = 7.0 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.2, 168.3, 134.7, 129.5, 128.5, 127.1, 57.6, 53.1, 51.9, 37.0, 30.1, 18.8, 18.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{23}N_2O_3]^+$: 279.1709 $[M – CF_3CO_2]^+$, found 279.1703, HRMS (ESI) *m*/*z* calcd. for $[C_{30}H_{45}N_4O_6]^+$: 557.3339 $[2M – C_4HO_4F_6]^+$, found 557.3332.

*(c) N-Acetyl-L-valyl-L-phenylalanyl-L-valine methyl ester (**40c**):* Prepared according to GP4 using 0.40 g (2.50 mmol) of *N*-acetyl-L-valine and 0.99 g (2.53 mmol) of L-phenylalanyl-L-valine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.57 g (54%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.13 (d, *J* = 7.7 Hz, 1H), 8.09 (d, *J* = 8.1 Hz, 1H), 7.85 (d, *J* = 8.7 Hz, 1H), 7.29–7.18 (m, 4H), 7.22–7.13 (m, 1H), 4.60 (ddd, *J* = 9.7, 8.1, 4.7 Hz, 1H), 4.16 (dd, *J* = 8.1, 6.4 Hz, 1H), 4.09 (dd, *J* = 8.9, 7.0 Hz, 1H), 3.60 (s, 3H), 2.99 (dd, *J* = 14.0, 4.7 Hz, 1H), 2.80 (dd, *J* = 13.9, 9.7 Hz, 1H), 2.04 (hept, *J* = 6.9 Hz, 1H), 1.94–1.84 (m, 1H), 1.83 (s, 3H), 0.88 (d, *J* = 7.0 Hz, 3H), 0.86 (d, *J* = 7.0 Hz, 3H), 0.74 (d, *J* = 2.7 Hz, 3H), 0.72

ppm (d, J = 2.7 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.7, 171.4, 170.9, 169.1, 137.7, 129.2, 127.9, 126.2, 57.8, 57.4, 53.6, 51.7, 37.3, 30.3, 29.9, 22.5, 19.1, 18.9, 18.2, 18.1 ppm. HRMS (ESI) m/z calcd. for $[C_{22}H_{34}N_3O_5]^+$: 420.2499 $[M + H]^+$, found 420.2496, HRMS (ESI) m/z calcd. for $[C_{22}H_{33}N_3O_5Na]^+$: 442.2318 $[M + Na]^+$, found 442.2315.

_N_-Acetyl-L-phenylalanyl-L-valyl-L-valine methyl ester (41c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35).

*(a) N-Boc-L-valyl-L-valine methyl ester (**9e**):* Prepared according to GP4, purified by column chromatography (3:1 Pet. ether/EtOAc, ninhydrin stain). A white powder, yield: 2.94 g (89%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.97 (d, J = 7.9 Hz, 1H), 6.69 (d, J = 9.1 Hz, 1H), 4.18 (dd, J = 8.0, 6.2 Hz, 1H), 3.86 (dd, J = 9.2, 7.0 Hz, 1H), 3.61 (s, 3H), 2.04 (hept, J = 6.8 Hz, 1H), 1.91 (hept, J = 6.8 Hz, 1H), 1.38 (s, 9H), 0.91–0.80 ppm (m, 12H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.83, 171.78, 155.36, 77.96, 59.49, 57.24, 51.58, 30.32, 29.86, 28.15, 19.12, 18.86, 18.21, 18.17 ppm. HRMS (ESI) m/z calcd. for $[C_{16}H_{31}N_2O_5]^+$: 331.2233 $[M + H]^+$, found 331.2227, HRMS (ESI) m/z calcd. for $[C_{16}H_{30}N_2O_5Na]^+$: 353.2052 $[M + Na]^+$, found 353.2046.

*(b) L-Valyl-L-valine methyl ester trifluoroacetate salt (**9b**):* Prepared according to GP5 using 2.17 g (6.58 mmol) of *N*-Boc-L-valyl-L-valine methyl ester. A white powder, yield: 2.08 g (92%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.61 (d, J = 7.4 Hz, 1H), 8.14 (d, J = 4.4 Hz, 3H), 4.20 (dd, J = 7.4, 5.9 Hz, 1H), 3.73 (t, J = 4.8 Hz, 1H), 3.64 (s, 3H), 2.17–2.00 (m, 2H), 1.00–0.88 ppm (m, 12H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.36, 168.35, 57.77, 56.99, 51.77, 29.94, 29.67, 18.84, 18.20, 18.15, 17.45 ppm. HRMS (ESI) m/z calcd. for $[C_{11}H_{23}N_2O_3]^+$: 231.1709 $[M - CF_3CO_2]^+$, found 231.1706, HRMS (ESI) m/z calcd. for $[C_{22}H_{45}N_4O_6]^+$: 461.3339 $[2M - C_4HO_4F_6]^+$, found 461.3338.

*(c) N-Acetyl-L-phenylalanyl-L-valyl-L-valine methyl ester (**41c**):* Prepared according to GP4 using 1.28 g (6.19 mmol) of *N*-acetyl-L-phenylalanine and 2.08 g (6.05 mmol) of L-valyl-L-valine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 1.75 g (69%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.18 (d, *J* = 7.7 Hz, 1H), 8.13 (d, *J* = 8.4 Hz, 1H), 7.95 (d, *J* = 8.9 Hz, 1H), 7.27–7.19 (m, 4H), 7.21–7.12 (m, 1H), 4.56 (ddd, *J* = 10.1, 8.5, 4.2 Hz, 1H), 4.28 (dd, *J* = 8.9, 6.9 Hz, 1H), 4.14 (dd, *J* = 7.5, 6.4 Hz, 1H), 3.60 (s, 3H), 2.98 (dd, *J* = 13.9, 4.2 Hz, 1H), 2.72 (dd, *J* = 14.0, 9.9 Hz, 1H), 2.11–1.99 (m, 1H), 2.03–1.91 (m, 1H), 1.74 (s, 3H), 0.89 (dd, *J* = 9.5, 5.5 Hz, 6H), 0.85 ppm (dd, *J* = 7.3, 5.6 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.76, 171.30, 171.26, 169.15, 138.07, 129.17, 127.96, 126.16, 57.54, 57.33, 53.88, 51.57, 37.32, 30.79, 29.66, 22.43, 19.08, 18.91, 18.30, 18.13 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{34}N_3O_5]^+$: 420.2499 [M + H]$^+$, found 420.2494, HRMS (ESI) *m*/*z* calcd. for $[C_{22}H_{33}N_3O_5Na]^+$: 442.2318 [M + Na]$^+$, found 442.2313.

N-Acetyl-L-leucyl-L-leucyl-L-phenylalanine methyl ester (42c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35).

*(a) N-Boc-L-leucyl-L-phenylalanine methyl ester (**29e**):* Prepared according to GP4 using 4.26 g (18.4 mmol) of *N*-Boc-L-leucine and 4.01 g (18.6 mmol) of L-phenylalanine methyl ester hydrochloride salt, purified by column chromatography (7:3 Pet. ether/EtOAc, ninhydrin stain). A white powder, yield: 6.12 g (85%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.12 (d, *J* = 7.7 Hz, 1H), 7.30–7.22 (m, 2H), 7.22–7.18 (m, 3H), 6.80 (d, *J* = 8.5 Hz, 1H), 4.48 (td, *J* = 8.2, 5.7 Hz, 1H), 3.96 (td, *J* = 9.0, 5.7 Hz, 1H), 3.57 (s, 3H), 3.02 (dd, *J* = 13.8, 5.8 Hz, 1H), 2.94 (dd, *J* = 13.9, 8.7 Hz, 1H), 1.61–1.44 (m, 1H), 1.37 (s, 9H), 1.31 (d, *J* = 7.5 Hz, 2H), 0.85 (d, *J* = 6.6 Hz, 3H), 0.81 ppm (d, *J* = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.5, 171.8, 155.1, 137.0, 129.1, 128.2, 126.5, 78.0, 53.3, 52.6, 51.8, 40.8, 36.6, 28.2, 24.1, 22.8, 21.6 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{33}N_2O_5]^+$: 393.2384 [M + H]$^+$,

found 393.2383, HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{32}N_2O_5Na]^+$: 415.2204 $[M + Na]^+$, found 415.2203.

*(b) L-Leucyl-L-phenylalanine methyl ester trifluoroacetate salt (**29b**):* Prepared according to GP5 using 4.70 g (12.0 mmol) of *N*-Boc-L-leucyl-L-phenylalanine methyl ester. A white powder, yield: 3.79 g (78%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.96 (d, *J* = 7.3 Hz, 1H), 8.14 (s, 3H), 7.35–7.26 (m, 2H), 7.28–7.19 (m, 3H), 4.61–4.49 (m, 1H), 3.77 (dd, *J* = 8.5, 5.3 Hz, 1H), 3.60 (s, 3H), 3.07 (dd, *J* = 14.0, 5.9 Hz, 1H), 2.99 (dd, *J* = 14.0, 8.6 Hz, 1H), 1.73–1.58 (m, 1H), 1.53 (dd, *J* = 7.7, 5.9 Hz, 2H), 0.90 (d, *J* = 5.1 Hz, 3H), 0.88 ppm (d, *J* = 4.7 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.3, 169.4, 136.8, 129.0, 128.4, 126.7, 53.9, 52.0, 50.6, 40.2, 36.3, 23.4, 22.8, 21.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{16}H_{25}N_2O_3]^+$: 293.1865 $[M – CF_3CO_2]^+$, found 293.1860.

*(c) N-Acetyl-L-leucyl-L-leucyl-L-phenylalanine methyl ester (**42c**):* Prepared according to GP4 using 0.79 g (4.57 mmol) of *N*-acetyl-L-leucine and 1.86 g (4.58 mmol) of L-leucyl-L-phenylalanine methyl ester trifluoroacetate salt, purified by column chromatography (19:1 DCM/MeOH, PMA stain). A white solid, yield: 0.87 g (43%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.21 (d, *J* = 7.4 Hz, 1H), 7.95 (d, *J* = 8.1 Hz, 1H), 7.83 (d, *J* = 8.4 Hz, 1H), 7.27 (dd, *J* = 8.0, 6.1 Hz, 2H), 7.20 (td, *J* = 6.6, 1.7 Hz, 3H), 4.46 (td, *J* = 8.2, 6.1 Hz, 1H), 4.28 (dq, *J* = 15.0, 7.8 Hz, 2H), 3.56 (s, 3H), 3.01 (dd, *J* = 13.9, 6.0 Hz, 1H), 2.94 (dd, *J* = 14.0, 8.5 Hz, 1H), 1.82 (s, 3H), 1.55 (dh, *J* = 13.5, 6.8 Hz, 2H), 1.38 (q, *J* = 7.1 Hz, 4H), 0.86 (d, *J* = 6.6 Hz, 6H), 0.82 ppm (dd, *J* = 6.6, 3.1 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.0, 171.9, 171.7, 169.1, 137.0, 129.0, 128.2, 126.5, 53.4, 51.8, 50.9, 50.6, 40.9, 40.7, 36.5, 24.1, 24.0, 23.0, 22.9, 22.5, 21.7, 21.6 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{38}N_3O_5]^+$: 448.2812 $[M + H]^+$, found 448.2809, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{37}N_3O_5K]^+$: 486.2370 $[M + K]^+$, found 486.2366.

_N_-Acetyl-L-leucyl-L-phenylalanyl-L-leucine methyl ester (43c)

Synthesised from the N-terminus (Method B, Table 2.2-6, page 35).

*(a) N-Acetyl-L-leucyl-L-phenylalanine (**29a**):* Prepared according to GP6 using 0.67 g (2.02 mmol) of *N*-acetyl-L-leucyl-L-phenylalanine methyl ester. A white powder, yield: 0.51 g (79%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.66 (s, 1H), 8.08 (d, *J* = 7.8 Hz, 1H), 7.90 (d, *J* = 8.5 Hz, 1H), 7.30–7.21 (m, 2H), 7.24–7.15 (m, 3H), 4.40 (td, *J* = 8.2, 5.3 Hz, 1H), 4.31 (td, *J* = 8.8, 5.9 Hz, 1H), 3.04 (dd, *J* = 13.9, 5.2 Hz, 1H), 2.91 (dd, *J* = 13.9, 8.8 Hz, 1H), 1.80 (s, 3H), 1.55 (dp, *J* = 13.4, 6.7 Hz, 1H), 1.43–1.30 (m, 2H), 0.86 (d, *J* = 6.6 Hz, 3H), 0.82 ppm (d, *J* = 6.6 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.7, 172.1, 168.9, 137.5, 129.1, 128.1, 126.4, 53.3, 50.6, 40.8, 36.5, 24.1, 23.0, 22.5, 21.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{25}N_2O_4]^+$: 321.1814 $[M + H]^+$, found 321.1811, HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{24}N_2O_4Na]^+$: 343.1634 $[M + Na]^+$, found 343.1630.

*(b) N-Acetyl-L-leucyl-L-phenylalanyl-L-leucine methyl ester (**43c**):* Prepared according to GP4 using 0.49 g (1.51 mmol) of *N*-acetyl-L-leucyl-L-phenylalanine and 0.29 g (1.59 mmol) of L-leucine methyl ester hydrochloride salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.30 g (43%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.18 (d, *J* = 7.6 Hz, 1H), 7.93 (dd, *J* = 8.3, 5.0 Hz, 2H), 7.29–7.11 (m, 5H), 4.51 (td, *J* = 8.9, 4.9 Hz, 1H), 4.29 (td, *J* = 9.1, 5.2 Hz, 1H), 4.19 (q, *J* = 7.6 Hz, 1H), 3.60 (s, 3H), 3.03 (dd, *J* = 14.1, 4.8 Hz, 1H), 2.80 (dd, *J* = 13.9, 9.3 Hz, 1H), 1.80 (s, 3H), 1.67–1.52 (m, 2H), 1.54–1.42 (m, 2H), 1.30 (t, *J* = 7.3 Hz, 2H), 0.89 (d, *J* = 6.3 Hz, 3H), 0.83 (d, *J* = 6.4 Hz, 6H), 0.79 ppm (d, *J* = 6.5 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.7, 171.9, 171.1, 169.2, 137.7, 129.2, 128.0, 126.2, 53.3, 51.8, 51.1, 50.2, 40.7, 39.7, 37.1, 24.1, 24.0, 22.9, 22.8, 22.4, 21.7, 21.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{38}N_3O_5]^+$: 448.2812 $[M + H]^+$, found 448.2809, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{37}N_3O_5Na]^+$: 470.2631 $[M + Na]^+$, found 470.2628.

_N_-Acetyl-L-phenylalanyl-L-leucyl-L-leucine methyl ester (44c)

Synthesised from the N-terminus (Method B, Table 2.2-6, page 35).

*(a) N-Acetyl-L-phenylalanyl-L-leucine (**32a**):* Prepared according to GP6. A white powder, yield: 0.47 g (79%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.55 (s, 1H), 8.23 (d, *J* = 7.9 Hz, 1H), 8.04 (d, *J* = 8.5 Hz, 1H), 7.29–7.20 (m, 4H), 7.22–7.14 (m, 1H), 4.54 (ddd, *J* = 10.2, 8.6, 4.0 Hz, 1H), 4.22 (ddd, *J* = 9.2, 7.8, 5.6 Hz, 1H), 3.00 (dd, *J* = 14.0, 4.0 Hz, 1H), 2.70 (dd, *J* = 13.8, 10.2 Hz, 1H), 1.73 (s, 3H), 1.70–1.56 (m, 1H), 1.56–1.49 (m, 2H), 0.90 (d, *J* = 6.5 Hz, 3H), 0.85 ppm (d, *J* = 6.4 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 174.8, 172.4, 171.8, 137.8, 129.8, 128.9, 127.1, 54.9, 51.4, 40.6, 37.9, 25.0, 22.7, 22.3, 21.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{25}N_2O_4]^+$: 321.1814 $[M + H]^+$, found 321.1810, HRMS (ESI) *m*/*z* calcd. for $[C_{17}H_{24}N_2O_4Na]^+$: 343.1634 $[M + Na]^+$, found 343.1630.

*(b) N-Acetyl-L-phenylalanyl-L-leucyl-L-leucine methyl ester (**44c**):* Prepared according to GP4 using 0.46 g (1.43 mmol) of *N*-acetyl-L-phenylalanyl-L-leucine and 0.39 g (2.12 mmol) of L-leucine methyl ester hydrochloride salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.36 g (56%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.20 (d, *J* = 7.7 Hz, 1H), 8.04 (d, *J* = 8.0 Hz, 2H), 7.24 (d, *J* = 4.3 Hz, 4H), 7.20–7.15 (m, 1H), 4.51 (td, *J* = 9.8, 4.2 Hz, 1H), 4.39–4.23 (m, 2H), 3.60 (s, 3H), 2.97 (dd, *J* = 13.9, 4.1 Hz, 1H), 2.70 (dd, *J* = 13.9, 9.9 Hz, 1H), 1.74 (s, 3H), 1.68–1.52 (m, 3H), 1.54–1.41 (m, 3H), 0.90 (dd, *J* = 6.3, 1.6 Hz, 6H), 0.85 ppm (dd, *J* = 8.5, 6.4 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.9, 173.0, 172.0, 171.5, 138.4, 130.2, 129.3, 127.6, 55.9, 52.6, 52.4, 51.7, 41.5, 41.0, 38.1, 25.4, 25.3, 23.3, 23.2, 23.0, 21.8, 21.7 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{38}N_3O_5]^+$: 448.2812 $[M + H]^+$, found 448.2809, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{37}N_3O_5Na]^+$: 470.2631 $[M + Na]^+$, found 470.2629.

_N_-Acetyl-L-phenylalanylglycyl-L-phenylalanine methyl ester (45c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35).

*(a) N-Boc-glycyl-L-phenylalanine methyl ester (**27e**):* Prepared according to GP4, purified by column chromatography (2:3 Pet. ether/EtOAc, ninhydrin stain). A viscous light-yellow oil, yield: 3.31 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.18 (d, J = 7.7 Hz, 1H), 7.33–7.11 (m, 5H), 6.88 (t, J = 6.3 Hz, 1H), 4.46 (q, J = 7.6 Hz, 1H), 3.57 (s, 3H), 3.55–3.41 (m, 2H), 3.00 (dd, J = 13.8, 5.8 Hz, 1H), 2.90 (dd, J = 13.8, 8.5 Hz, 1H), 1.35 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.9, 169.4, 155.7, 137.0, 129.1, 128.3, 126.6, 78.0, 53.5, 51.8, 42.8, 36.8, 28.2 ppm.

*(b) Glycyl-L-phenylalanine methyl ester trifluoroacetate salt (**27b**):* Prepared according to GP5. An off-white solid, yield: 3.12 g (91%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.91 (d, J = 7.7 Hz, 1H), 8.05 (s, 3H), 7.34–7.25 (m, 2H), 7.27–7.19 (m, 3H), 4.57 (td, J = 8.4, 5.6 Hz, 1H), 3.62 (s, 3H), 3.62–3.50 (m, 2H), 3.07 (dd, J = 13.8, 5.5 Hz, 1H), 2.92 ppm (dd, J = 13.8, 8.9 Hz, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.5, 166.1, 136.8, 129.1, 128.4, 126.7, 53.8, 52.1, 40.0, 36.7 ppm. HRMS (ESI) m/z calcd. for $[C_{12}H_{17}N_2O_3]^+$: 237.1239 $[M - CF_3CO_2]^+$, found 237.1243.

*(c) N-Acetyl-L-phenylalanylglycyl-L-phenylalanine methyl ester (**45c**):* Prepared according to GP4, recrystallised from hot ethyl acetate. A white solid, yield: 1.80 g (43%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.27 (t, J = 5.6 Hz, 1H), 8.25 (d, J = 8.0 Hz, 1H), 8.14 (d, J = 8.0 Hz, 1H), 7.32–7.14 (m, 10H), 4.51–4.43 (m, 2H), 3.76 (dd, J = 16.8, 6.0 Hz, 1H), 3.62 (dd, J = 17.0, 5.7 Hz, 1H), 3.59 (s, 3H), 3.02 (dd, J = 13.8, 5.8 Hz, 1H), 2.98 (dd, J = 13.8, 4.2 Hz, 1H), 2.92 (dd, J = 13.8, 8.7 Hz, 1H), 2.72 (dd, J = 13.9, 10.1 Hz, 1H), 1.74 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.80, 171.72, 169.41, 168.74, 138.06, 137.04, 129.08, 129.05, 128.27, 128.02, 126.59, 126.21, 54.19, 53.63, 51.87, 41.60, 37.38, 36.77, 22.46 ppm. HRMS (ESI) m/z calcd. for

$[C_{23}H_{28}N_3O_5]^+$: 426.2029 [M + H]$^+$, found 426.2045, HRMS (ESI) *m*/*z* calcd. for $[C_{23}H_{27}N_3O_5Na]^+$: 448.1848 [M + Na]$^+$, found 448.1865.

_N_-Acetyl-L-phenylalanyl-L-leucyl-L-phenylalanine methyl ester (46c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35). Prepared according to GP4 using 0.87 g (4.19 mmol) of *N*-acetyl-L-phenylalanine and 1.72 g (4.22 mmol) of L-leucyl-L-phenylalanine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.67 g (33%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.29 (d, *J* = 7.4 Hz, 1H), 8.04 (d, *J* = 8.4 Hz, 1H), 7.99 (d, *J* = 8.3 Hz, 1H), 7.31–7.13 (m, 10H), 4.54–4.42 (m, 2H), 4.43 (td, *J* = 8.3, 6.4 Hz, 1H), 3.56 (s, 3H), 3.03 (dd, *J* = 13.9, 6.0 Hz, 1H), 3.00–2.88 (m, 2H), 2.67 (dd, *J* = 13.9, 10.2 Hz, 1H), 1.73 (s, 3H), 1.61–1.48 (m, 1H), 1.47–1.34 (m, 2H), 0.88 (d, *J* = 6.6 Hz, 3H), 0.83 ppm (d, *J* = 6.5 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.9, 171.7, 171.1, 169.1, 138.0, 137.1, 129.1, 129.0, 128.2, 128.0, 126.5, 126.1, 53.7, 53.4, 51.8, 50.8, 41.0, 37.4, 36.5, 24.0, 22.9, 22.4, 21.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{27}H_{36}N_3O_5]^+$: 482.2655 [M + H]$^+$, found 482.2656, HRMS (ESI) *m*/*z* calcd. for $[C_{27}H_{35}N_3O_5Na]^+$: 504.2474 [M + Na]$^+$, found 504.2469.

_N_-Acetyl-L-phenylalanyl-L-phenylalanyl-L-valine methyl ester (47c)

Synthesised from the C-terminus (Method A, Table 2.2-6, page 35). Prepared according to GP4 using 0.54 g (2.62 mmol) of *N*-acetyl-L-phenylalanine and 1.05 g (2.68 mmol) of L-phenylalanyl-L-valine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.78 g (65%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.20 (d, *J* = 8.1 Hz, 1H), 8.10 (d, *J* = 8.1 Hz, 1H), 8.00 (d, *J* = 8.4 Hz, 1H), 7.28–7.24 (m, 4H), 7.25–7.11 (m, 6H), 4.62 (td, *J* = 8.8, 4.8 Hz, 1H), 4.46 (td, *J* = 9.2, 4.1 Hz, 1H), 4.19 (dd, *J* = 8.3, 6.2 Hz, 1H), 3.63 (s, 3H), 3.02 (dd, *J* = 14.0, 4.7 Hz, 1H), 2.92 (dd, *J* = 13.9, 4.2 Hz, 1H), 2.82 (dd, *J* = 14.0, 9.2 Hz, 1H), 2.64

(dd, *J* = 13.9, 9.9 Hz, 1H), 2.04 (h, *J* = 6.7 Hz, 1H), 1.70 (s, 3H), 0.89 ppm (t, *J* = 7.4 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, CD_3OD): δ 173.29, 173.28, 173.16, 173.04, 138.43, 138.10, 130.45, 130.18, 129.39, 129.38, 127.73, 127.69, 59.28, 55.89, 55.63, 52.49, 38.89, 38.62, 31.91, 22.36, 19.44, 18.62 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{26}H_{34}N_3O_5]^+$: 468.2499 $[M + H]^+$, found 482.2655, HRMS (ESI) *m*/*z* calcd. for $[C_{26}H_{33}N_3O_5Na]^+$: 490.2318 $[M + Na]^+$, found 490.2305.

_N_-Acetyl-L-phenylalanyl-L-phenylalanyl-L-leucine methyl ester (48c)

Synthesised from the N-terminus (Method B, Table 2.2-6, page 35).

*(a) N-Acetyl-L-phenylalanyl-L-phenylalanine (**34a**):* Prepared according to GP6 using 0.84 g (2.29 mmol) of *N*-acetyl-L-phenylalanyl-L-phenylalanine methyl ester. A white powder, yield: 0.81 g (100%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.77 (s, 1H), 8.29 (d, *J* = 7.8 Hz, 1H), 8.10 (d, *J* = 8.6 Hz, 1H), 7.32–7.16 (m, 9H), 7.20–7.12 (m, 1H), 4.51 (ddd, *J* = 10.1, 8.5, 4.2 Hz, 1H), 4.44 (td, *J* = 8.4, 5.1 Hz, 1H), 3.08 (dd, *J* = 13.9, 5.1 Hz, 1H), 2.97 (dd, *J* = 10.0, 4.0 Hz, 1H), 2.92 (d, *J* = 14.4 Hz, 1H), 2.67 (dd, *J* = 13.9, 10.2 Hz, 1H), 1.71 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.7, 171.5, 169.0, 138.0, 137.5, 129.2, 129.1, 128.2, 127.9, 126.4, 126.1, 53.7, 53.5, 37.4, 36.6, 22.4 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{23}N_2O_4]^+$: 355.1658 $[M + H]^+$, found 355.1654, HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{22}N_2O_4Na]^+$: 377.1477 $[M + Na]^+$, found 377.1474.

*(b) N-Acetyl-L-phenylalanyl-L-phenylalanyl-L-leucine methyl ester (**48c**):* Prepared according to GP4 using 0.77 g (2.18 mmol) of *N*-acetyl-L-phenylalanyl-L-phenylalanine and 0.40 g (2.20 mmol) of L-leucine methyl ester hydrochloride salt, recrystallised from hot ethyl acetate. A white solid, yield: 0.78 g (74%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.30 (d, *J* = 7.7 Hz, 1H), 8.09 (d, *J* = 8.2 Hz, 1H), 8.00 (d, *J* = 8.3 Hz, 1H), 7.31–7.23 (m, 4H), 7.25–7.11 (m, 6H), 4.54 (td, *J* = 8.8, 4.7 Hz,

1H), 4.45 (ddd, *J* = 9.9, 8.2, 4.2 Hz, 1H), 4.31 (ddd, *J* = 9.7, 7.7, 5.1 Hz, 1H), 3.61 (s, 3H), 3.04 (dd, *J* = 14.0, 4.7 Hz, 1H), 2.91 (dd, *J* = 13.9, 4.2 Hz, 1H), 2.81 (dd, *J* = 13.9, 9.2 Hz, 1H), 2.64 (dd, *J* = 13.9, 9.9 Hz, 1H), 1.71 (s, 3H), 1.69–1.42 (m, 3H), 0.90 (d, *J* = 6.4 Hz, 3H), 0.85 ppm (d, *J* = 6.3 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.53, 172.16, 171.87, 171.68, 137.27, 137.12, 129.63, 129.39, 128.63, 128.61, 126.95, 126.93, 55.09, 54.21, 52.36, 51.15, 39.97, 37.41, 37.24, 24.58, 22.38, 21.96, 21.02 ppm. HRMS (ESI) *m/z* calcd. for $[C_{27}H_{36}N_3O_5]^+$: 482.2655 $[M + H]^+$, found 482.2655, HRMS (ESI) *m/z* calcd. for $[C_{27}H_{35}N_3O_5Na]^+$: 504.2474 $[M + Na]^+$, found 504.2472.

_N_-Acetyl-L-valyl-L-valylglycyl-L-phenylalanine methyl ester (49c)

*(a) N-Acetyl-L-valyl-L-valine (**9c**):* Prepared according to GP6. A white powder, yield: 0.44 g (93%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 12.52 (s, 1H), 7.93 (d, *J* = 8.1 Hz, 1H), 7.84 (d, *J* = 8.9 Hz, 1H), 4.27 (dd, *J* = 9.0, 6.9 Hz, 1H), 4.08 (dd, *J* = 8.1, 5.9 Hz, 1H), 2.09–2.00 (m, 1H), 1.98–1.89 (m, 1H), 1.86 (s, 3H), 0.92–0.80 ppm (m, 12H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.82, 171.50, 169.14, 57.32, 57.26, 30.56, 29.64, 22.48, 19.17, 19.07, 18.21, 18.13 ppm.

*(b) N-Acetyl-L-valyl-L-valylglycyl-L-phenylalanine methyl ester (**49c**):* Prepared according to GP4 using 1.16 g (4.50 mmol) of *N*-acetyl-L-valyl-L-valine and 1.79 g (5.09 mmol) of glycyl-L-phenylalanine methyl ester trifluoroacetate salt, recrystallised from hot ethyl acetate. A white solid, yield: 1.26 g (59%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.29 (d, *J* = 7.7 Hz, 1H), 8.10 (t, *J* = 5.8 Hz, 1H), 7.87 (d, *J* = 8.8 Hz, 1H), 7.71 (d, *J* = 8.4 Hz, 1H), 7.30–7.25 (m, 2H), 7.24–7.16 (m, 3H), 4.47 (dd, *J* = 8.2, 6.0 Hz, 1H), 4.18 (dd, *J* = 8.8, 7.0 Hz, 1H), 4.13 (dd, *J* = 8.5, 6.7 Hz, 1H), 3.76–3.64 (m, 2H), 3.58 (s, 3H), 3.01 (dd, *J* = 13.7, 6.0 Hz, 1H), 2.91 (dd, *J* = 13.8, 8.5 Hz, 1H), 1.98–1.89 (m, 2H), 1.86 (s, 3H), 0.85–0.79 ppm (m, 12H). ^{13}C {^{1}H} NMR (101

MHz, DMSO-d_6): δ 171.79, 171.16, 171.02, 169.24, 168.63, 136.97, 129.04, 128.26, 126.58, 57.79, 57.69, 53.60, 51.83, 41.41, 36.84, 30.45, 30.08, 22.48, 19.29, 19.15, 18.20, 18.15 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{37}N_4O_6]^+$: 477.2708 $[M + H]^+$, found 477.2717, HRMS (ESI) *m*/*z* calcd. for $[C_{24}H_{36}N_4O_6Na]^+$: 499.2527 $[M + Na]^+$, found 499.2525.

GP11. General procedure for the O-acetylation of tyrosine residues[25]

The N-acetylated tyrosine-containing amino acid/peptide methyl ester (4.41 mmol) and DMAP (5 mg, 0.04 mmol) were dissolved in pyridine (20 mL). Acetic anhydride (8.0 mL, 84.8 mmol) was added dropwise and the resulting mixture was stirred for 2 hours at room temperature. The reaction was quenched by the addition of ethanol (10 mL) and the solvent was removed under reduced pressure to give a brown residue. Toluene (3 x 10 mL) was added and removed under reduced pressure to get rid of residual pyridine. The crude product was then purified by column chromatography or recrystallisation from hot ethyl acetate to give the O-acetylated tyrosine-containing amino acid/peptide derivative.

***N*-Acetyl-*O*-acetyl-L-tyrosine methyl ester (25c')**

*(a) N-Acetyl-L-tyrosine methyl ester (**25c**):* Provided by Dr. Luke Gamon. ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.21 (s, 1H), 8.26 (d, *J* = 7.6 Hz, 1H), 7.02–6.95 (m, 2H), 6.69–6.62 (m, 2H), 4.35 (ddd, *J* = 9.0, 7.7, 5.8 Hz, 1H), 3.57 (s, 3H), 2.86 (dd, *J* = 13.8, 5.8 Hz, 1H), 2.74 (dd, *J* = 13.8, 9.0 Hz, 1H), 1.79 ppm (s, 3H). HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{16}NO_4]^+$: 238.1074 $[M + H]^+$, found 238.1119, HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{15}NO_4Na]^+$: 260.0893 $[M + Na]^+$, found 260.0941.

*(b) N-Acetyl-O-acetyl-L-tyrosine methyl ester (**25c'**):* Prepared according to GP11, purified by column chromatography (1:9 Pet. ether/EtOAc, PMA stain). A white solid, yield: 0.94 g (77%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.35 (d, *J* = 7.7 Hz, 1H),

7.28–7.20 (m, 2H), 7.08–6.99 (m, 2H), 4.44 (ddd, *J* = 9.2, 7.7, 5.5 Hz, 1H), 3.59 (s, 3H), 3.01 (dd, *J* = 13.9, 5.5 Hz, 1H), 2.88 (dd, *J* = 13.9, 9.3 Hz, 1H), 2.25 (s, 3H), 1.79 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.1, 169.4, 169.2, 149.1, 134.7, 130.0, 121.5, 53.5, 51.8, 36.0, 22.2, 20.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{18}NO_5]^+$: 280.1185 $[M + H]^+$, found 280.1189, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{17}NO_5Na]^+$: 302.1004 $[M + Na]^+$, found 302.1010.

_N_-Acetyl-L-phenylalanyl-_O_-acetyl-L-tyrosine methyl ester (35c')

*(a) L-Tyrosine methyl ester hydrochloride salt (**25b**)*: Prepared according to GP2 using 6.94 g (38.3 mmol) of L-tyrosine. A white solid, yield: 8.60 g (97%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.49 (s, 1H), 8.66 (s, 3H), 7.05–6.91 (m, 2H), 6.81–6.69 (m, 2H), 4.13 (dd, *J* = 7.1, 5.7 Hz, 1H), 3.66 (s, 3H), 3.08 (dd, *J* = 14.1, 5.7 Hz, 1H), 2.99 ppm (dd, *J* = 14.1, 7.1 Hz, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 169.4, 156.7, 130.4, 124.3, 115.4, 53.5, 52.5, 35.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{10}H_{14}NO_3]^+$: 196.0974 $[M - Cl]^+$, found 196.0970, HRMS (ESI) *m*/*z* calcd. for $[C_{20}H_{27}N_2O_6]^+$: 391.1869 $[2M - 2Cl - H]^+$, found 391.1866.

*(b) N-Acetyl-L-phenylalanyl-L-tyrosine methyl ester (**35c'**):* Prepared according to GP4, purified by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white powder, yield: 2.07 g (54%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.23 (s, 1H), 8.37 (d, *J* = 7.5 Hz, 1H), 8.05 (d, *J* = 8.6 Hz, 1H), 7.30–7.19 (m, 4H), 7.20–7.13 (m, 1H), 7.05–6.94 (m, 2H), 6.70–6.60 (m, 2H), 4.54 (ddd, *J* = 10.1, 8.5, 4.4 Hz, 1H), 4.39 (td, *J* = 7.9, 6.0 Hz, 1H), 3.57 (s, 3H), 2.95 (dd, *J* = 13.9, 4.5 Hz, 1H), 2.91 (dd, *J* = 13.8, 5.8 Hz, 1H), 2.82 (dd, *J* = 13.9, 8.4 Hz, 1H), 2.67 (dd, *J* = 13.8, 10.0 Hz, 1H), 1.72 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.8, 171.5, 169.0, 156.0, 137.9, 130.0, 129.1, 128.0, 127.0, 126.1, 115.0, 53.9, 53.5, 51.8, 37.5, 35.6, 22.4 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{25}N_2O_5]^+$: 385.1764 $[M + H]^+$, found 385.1757, HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{24}N_2O_5Na]^+$: 407.1583 $[M + Na]^+$, found 407.1577.

*(c) N-Acetyl-L-phenylalanyl-O-acetyl-L-tyrosine methyl ester (**35c'**):* Prepared according to GP11 using 0.86 g (2.24 mmol) of *N*-acetyl-L-phenylalanyl-L-tyrosine methyl ester, recrystallised from hot ethyl acetate. A white solid, yield: 0.45 g (47%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.46 (d, *J* = 7.5 Hz, 1H), 8.04 (d, *J* = 8.6 Hz, 1H), 7.29–7.20 (m, 6H), 7.22–7.15 (m, 1H), 7.06–6.99 (m, 2H), 4.54 (ddd, *J* = 10.0, 8.5, 4.5 Hz, 1H), 4.48 (ddd, *J* = 8.7, 7.4, 5.8 Hz, 1H), 3.58 (s, 3H), 3.04 (dd, *J* = 14.0, 5.8 Hz, 1H), 2.98–2.95 (m, 1H), 2.95–2.92 (m, 1H), 2.68 (dd, *J* = 13.9, 10.0 Hz, 1H), 2.24 (s, 3H), 1.72 ppm (s, 3H). ^{13}C {^{1}H} NMR (126 MHz, DMSO-d_6): δ 171.63, 171.58, 169.13, 168.96, 149.17, 137.90, 134.50, 130.07, 129.10, 127.96, 126.18, 121.50, 53.49 (2C), 51.85, 37.46, 35.82, 22.39, 20.84 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{23}H_{27}N_2O_6]^+$: 427.1869 $[M + H]^+$, found 427.1868, HRMS (ESI) *m*/*z* calcd. for $[C_{23}H_{26}N_2O_6Na]^+$: 449.1689 $[M + Na]^+$, found 449.1684.

_N_-Acetyl-_O_-acetyl-L-tyrosyl-L-leucine methyl ester (36c')

*(a) N-Acetyl-L-tyrosyl-L-leucine methyl ester (**36c**):* Prepared according to GP4, purified by column chromatography (1:9 Pet. ether/EtOAc, PMA stain). A white powder, yield: 2.12 g (61%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.14 (s, 1H), 8.31 (d, *J* = 7.6 Hz, 1H), 7.98 (d, *J* = 8.4 Hz, 1H), 7.06–6.99 (m, 2H), 6.67–6.58 (m, 2H), 4.45 (ddd, *J* = 9.8, 8.5, 4.4 Hz, 1H), 4.28 (ddd, *J* = 9.5, 7.6, 5.2 Hz, 1H), 3.60 (s, 3H), 2.86 (dd, *J* = 13.9, 4.4 Hz, 1H), 2.58 (dd, *J* = 13.9, 9.9 Hz, 1H), 1.74 (s, 3H), 1.67–1.43 (m, 3H), 0.89 (d, *J* = 6.4 Hz, 3H), 0.84 ppm (d, *J* = 6.3 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, CD_3CN): δ 173.8, 172.2, 170.8, 156.6, 131.4, 129.4, 115.9, 55.3, 52.6, 51.7, 41.2, 37.6, 25.4, 23.1, 22.9, 21.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{27}N_2O_5]^+$: 351.1920 $[M + H]^+$, found 351.1914, HRMS (ESI) *m*/*z* calcd. for $[C_{18}H_{26}N_2O_5Na]^+$: 373.1739 $[M + Na]^+$, found 373.1734.

*(b) N-Acetyl-O-acetyl-L-tyrosyl-L-leucine methyl ester (**36c'**):* Prepared according to GP11 using 1.01 g (2.88 mmol) of N-acetyl-L-tyrosyl-L-leucine methyl ester, purified

by column chromatography (1:4 Pet. ether/EtOAc, PMA stain). A white solid, yield: 1.04 g (92%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.38 (d, J = 7.6 Hz, 1H), 8.08 (d, J = 8.4 Hz, 1H), 7.28 (d, J = 8.4 Hz, 2H), 7.02 (d, J = 8.4 Hz, 2H), 4.54 (ddd, J = 9.9, 8.5, 4.3 Hz, 1H), 4.29 (ddd, J = 9.6, 7.6, 5.3 Hz, 1H), 3.61 (s, 3H), 2.98 (dd, J = 14.0, 4.3 Hz, 1H), 2.72 (dd, J = 13.9, 9.9 Hz, 1H), 2.24 (s, 3H), 1.75 (s, 3H), 1.68–1.57 (m, 1H), 1.61–1.45 (m, 2H), 0.90 (d, J = 6.4 Hz, 3H), 0.84 ppm (d, J = 6.4 Hz, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.8, 171.6, 169.2, 169.1, 148.9, 135.4, 130.0, 121.3, 53.5, 51.8, 50.3, 39.7, 36.9, 24.2, 22.7, 22.4, 21.3, 20.9 ppm. HRMS (ESI) m/z calcd. for $[C_{20}H_{29}N_2O_6]^+$: 393.2026 $[M + H]^+$, found 393.2022, HRMS (ESI) m/z calcd. for $[C_{20}H_{28}N_2O_6Na]^+$: 415.1845 $[M + Na]^+$, found 415.1841.

_N_-Acetyl-_O_-acetyl-L-tyrosyl-L-phenylalanine methyl ester (37c')

*(a) N-Acetyl-L-tyrosyl-L-phenylalanine methyl ester (**37c**):* Prepared according to GP4. A white powder, yield: 1.86 g (48%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.16 (s, 1H), 8.39 (d, J = 7.6 Hz, 1H), 7.96 (d, J = 8.5 Hz, 1H), 7.27 (t, J = 7.4 Hz, 2H), 7.20 (d, J = 7.4 Hz, 3H), 7.00 (d, J = 8.0 Hz, 2H), 6.63 (d, J = 8.0 Hz, 2H), 4.51–4.38 (m, 2H), 3.57 (s, 3H), 3.03 (dd, J = 13.9, 5.9 Hz, 1H), 2.93 (dd, J = 13.9, 8.6 Hz, 1H), 2.82 (dd, J = 14.0, 4.6 Hz, 1H), 2.55 (dd, J = 14.0, 9.9 Hz, 1H), 1.72 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.72, 171.67, 168.90, 155.72, 137.05, 130.02, 129.07, 128.24, 127.90, 126.55, 114.79, 53.86, 53.57, 51.81, 36.75, 36.69, 22.43 ppm. HRMS (ESI) m/z calcd. for $[C_{21}H_{25}N_2O_5]^+$: 385.1764 $[M + H]^+$, found 385.1764, HRMS (ESI) m/z calcd. for $[C_{21}H_{24}N_2O_5Na]^+$: 407.1583 $[M + Na]^+$, found 407.1577.

*(b) N-Acetyl-O-acetyl-L-tyrosyl-L-phenylalanine methyl ester (**37c'**):* Prepared according to GP11 using 1.10 g (2.87 mmol) of *N*-acetyl-L-tyrosyl-L-phenylalanine methyl ester, recrystallised from hot ethyl acetate. A white solid, yield: 0.50 g (41%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.46 (d, J = 7.5 Hz, 1H), 8.05 (d, J = 8.5 Hz, 1H), 7.33–7.23 (m, 3H), 7.26–7.18 (m, 4H), 7.03–6.98 (m, 2H), 4.57–4.51 (m,

1H), 4.52–4.43 (m, 1H), 3.58 (s, 3H), 3.04 (dd, *J* = 13.9, 5.9 Hz, 1H), 2.98–2.95 (m, 1H), 2.95–2.90 (m, 1H), 2.68 (dd, *J* = 13.7, 10.1 Hz, 1H), 2.24 (s, 3H), 1.73 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.7, 171.5, 169.2, 169.0, 148.9, 137.0, 135.4, 130.0, 129.1, 128.2, 126.6, 121.3, 53.6, 53.4, 51.8, 36.8, 36.6, 22.4, 20.9 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{23}H_{27}N_2O_6]^+$: 427.1869 $[M + H]^+$, found 427.1866, HRMS (ESI) *m*/*z* calcd. for $[C_{23}H_{26}N_2O_6Na]^+$: 449.1689 $[M + Na]^+$, found 449.1682.

_N_-Acetyl-_O_-acetyl-L-tyrosyl-_O_-acetyl-L-tyrosine methyl ester (38c')

*(a) N-Acetyl-L-tyrosyl-L-tyrosine methyl ester (**38c**):* Prepared according to GP4. A white powder, yield: 2.38 g (60%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.21 (s, 1H), 9.14 (s, 1H), 8.29 (d, *J* = 7.5 Hz, 1H), 7.96 (d, *J* = 8.6 Hz, 1H), 7.03–6.95 (m, 4H), 6.75–6.52 (m, 4H), 4.44 (ddd, *J* = 9.8, 8.5, 4.6 Hz, 1H), 4.38 (ddd, *J* = 8.3, 7.5, 6.0 Hz, 1H), 3.57 (s, 3H), 2.90 (dd, *J* = 13.8, 6.1 Hz, 1H), 2.87–2.77 (m, 2H), 2.55 (dd, *J* = 13.9, 9.9 Hz, 1H), 1.73 ppm (s, 3H). ^{13}C {^{1}H} NMR (126 MHz, DMSO-d_6): δ 171.83, 171.60, 168.91, 155.99, 155.68, 130.02, 130.00, 127.95, 126.98, 115.03, 114.77, 53.90, 53.84, 51.75, 36.73, 35.90, 22.44 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{25}N_2O_6]^+$: 401.1713 $[M + H]^+$, found 401.1710, HRMS (ESI) *m*/*z* calcd. for $[C_{21}H_{24}N_2O_6Na]^+$: 423.1532 $[M + Na]^+$, found 423.1527.

*(b) N-Acetyl-O-acetyl-L-tyrosyl-O-acetyl-L-tyrosine methyl ester (**38c'**):* Prepared according to GP11 using 0.89 g (2.22 mmol) of *N*-acetyl-L-tyrosyl-L-tyrosine methyl ester, recrystallised from hot ethyl acetate. A white solid, yield: 0.61 g (57%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.48 (d, *J* = 7.5 Hz, 1H), 8.06 (d, *J* = 8.5 Hz, 1H), 7.25 (dd, *J* = 8.4, 4.3 Hz, 4H), 7.02 (t, *J* = 8.0 Hz, 4H), 4.54 (dt, *J* = 9.1, 4.6 Hz, 1H), 4.49 (td, *J* = 8.4, 6.1 Hz, 1H), 3.59 (s, 3H), 3.05 (dd, *J* = 14.0, 5.8 Hz, 1H), 2.95 (ddd, *J* = 14.1, 6.7, 4.2 Hz, 2H), 2.69 (dd, *J* = 13.9, 9.9 Hz, 1H), 2.24 (s, 6H), 1.73 ppm (s, 3H). ^{13}C {^{1}H} NMR (126 MHz, DMSO-d_6): δ 171.63, 171.50, 169.16, 169.13, 169.04, 149.18, 148.93, 135.36, 134.48, 130.08, 130.03, 121.51, 121.25, 53.49, 53.45, 51.86, 36.78, 35.82,

22.41, 20.85 (2C) ppm. HRMS (ESI) *m/z* calcd. for $[C_{25}H_{29}N_2O_8]^+$: 485.1924 [M + H]$^+$, found 485.1935, HRMS (ESI) *m/z* calcd. for $[C_{25}H_{28}N_2O_8Na]^+$: 507.1743 [M + Na]$^+$, found 507.1747.

_N_-Acetyl-N_ε-acetyl-L-tryptophan methyl ester (50c')

*(a) N-Acetyl-L-tryptophan methyl ester (**50c**):* Prepared according to GP3. A white powder, yield: 2.65 g (84%). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.85 (s, 1H), 8.30 (d, *J* = 7.5 Hz, 1H), 7.49 (d, *J* = 7.5 Hz, 1H), 7.37–7.32 (m, 1H), 7.14 (d, *J* = 2.4 Hz, 1H), 7.07 (ddd, *J* = 8.2, 7.0, 1.2 Hz, 1H), 7.01–6.95 (m, 1H), 4.49 (td, *J* = 7.9, 5.7 Hz, 1H), 3.57 (s, 3H), 3.13 (dd, *J* = 14.7, 5.4 Hz, 1H), 3.02 (dd, *J* = 14.3, 8.4 Hz, 1H), 1.81 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.6, 169.3, 136.1, 127.1, 123.6, 121.0, 118.4, 118.0, 111.4, 109.5, 53.1, 51.8, 27.1, 22.3 ppm. HRMS (ESI) *m/z* calcd. for $[C_{14}H_{17}N_2O_3]^+$: 261.1239 [M + H]$^+$, found 261.1234, HRMS (ESI) *m/z* calcd. for $[C_{14}H_{16}N_2O_3Na]^+$: 283.1059 [M + Na]$^+$, found 283.1053.

*(b) N-Acetyl-N_ε-acetyl-L-tryptophan methyl ester (**50c'**):*[38] *N*-Acetyl-L-tryptophan methyl ester (1.30 g, 5.01 mmol) and tetrabutylammonium hydrogen sulfate (0.18 g, 0.52 mmol) were suspended in dichloromethane (50 mL) under inert atmosphere. Freshly powdered sodium hydroxide (1.01 g, 25.3 mmol) was added and the resulting mixture was stirred for 15 minutes at room temperature. Acetyl chloride (1.2 mL, 16.9 mmol) was added dropwise over a period of 15 minutes. The reaction mixture was stirred for 4 hours at room temperature under inert atmosphere and water (70 mL) was added. The aqueous layer was extracted with dichloromethane (2 x 50 mL) and the combined organic layers were washed with brine (100 mL) and dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure and the crude product was purified by recrystallisation from hot ethanol to give the desired product as white crystals (0.70 g, 46%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.38 (d, *J* = 7.6 Hz, 1H), 8.30 (d, *J* = 7.8 Hz, 1H), 7.65

(s, 1H), 7.58 (dd, *J* = 8.0, 1.3 Hz, 1H), 7.33 (dd, *J* = 8.4, 7.0 Hz, 1H), 7.28 (td, *J* = 7.3, 1.3 Hz, 1H), 4.59 (td, *J* = 8.0, 5.8 Hz, 1H), 3.60 (s, 3H), 3.13 (dd, *J* = 14.7, 5.8 Hz, 1H), 3.04 (dd, *J* = 14.7, 8.3 Hz, 1H), 2.61 (s, 3H), 1.83 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 172.1, 169.4, 169.1, 135.0, 130.1, 125.0, 124.8, 123.3, 118.8, 116.8, 115.9, 52.1, 51.9, 26.6, 23.8, 22.3 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{16}H_{19}N_2O_4]^+$: 303.1345 [M + H]$^+$, found 303.1339, HRMS (ESI) *m*/*z* calcd. for $[C_{32}H_{37}N_4O_8]^+$: 605.2611 [M + Na]$^+$, found 605.2610.

_N_-Acetyl-L-proline methyl ester (51c)

*(a) N-Acetyl-L-proline (**51a**):*[42] L-Proline (8.07 g, 70.1 mmol) was dissolved in water (8 mL). The solution was heated to 55 °C, followed by dropwise addition of acetic anhydride (10 mL, 105 mmol). The resulting mixture was heated to 70 °C and stirred for 1 hour. Half of the solvent was then removed under reduced pressure. The solution was cooled to room temperature and the resulting solid was filtered. The crude product was recrystallised from hot water and washed with cold water and ethanol (2 x 5 mL) to give the desired product as white shards (6.99 g, 63%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 12.53 (s, 1H), 4.18 (dd, *J* = 8.8, 3.7 Hz, 1H), 3.57–3.42 (m, 2H), 2.22–1.98 (m, 2H), 1.96 (s, 3H), 1.95–1.85 ppm (m, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 173.5, 168.2, 58.1, 47.2, 29.1, 24.4, 22.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_7H_{12}NO_3]^+$: 158.0817 [M + H]$^+$, found 158.0813, HRMS (ESI) *m*/*z* calcd. for $[C_{14}H_{23}N_2O_6]^+$: 315.1556 [2M + H]$^+$, found 315.1550.

*(b) N-Acetyl-L-proline methyl ester (**51c**):* Prepared according to GP3 using 0.99 g (6.28 mmol) of *N*-acetyl-L-proline. A colourless liquid, yield: 0.88 g (82%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 4.25 (dd, *J* = 8.8, 4.2 Hz, 1H), 3.60 (s, 3H), 3.58–3.46 (m, 2H), 2.22–2.08 (m, 1H), 1.97 (s, 3H), 1.96–1.86 (m, 2H), 1.85–1.78 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.6, 168.3, 58.0,

51.7, 47.2, 29.0, 24.4, 22.0 ppm. HRMS (ESI) *m/z* calcd. for $[C_8H_{14}NO_3]^+$: 172.0974 $[M + H]^+$, found 172.0969, HRMS (ESI) *m/z* calcd. for $[C_8H_{13}NO_3Na]^+$: 194.0793 $[M + Na]^+$, found 194.0789.

GP12. General procedure for the proline-containing peptide coupling[43]

Method (a):[43a] The N-acetylated amino acid (5.73 mmol) was suspended in dichloromethane (15 mL) and the suspension was cooled to 0 °C, followed by addition of HOBt·H_2O (1.32 g, 8.59 mmol) and EDC·HCl (1.65 g, 8.59 mmol). After 10 minutes, the hydrochloride salt of amino acid methyl ester (6.30 mmol) was added, followed by dropwise addition of diisopropylethylamine (3.0 mL, 17.2 mmol). The resulting mixture was stirred overnight at room temperature and then treated with saturated aqueous sodium bicarbonate (50 mL). The aqueous layer was extracted with dichloromethane (4 x 50 mL) and the combined organic layers were dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure and the crude product was purified by column chromatography to give the proline-containing dipeptide derivative.

Method (b):[43b] The N-acetyl amino acid (10.1 mmol) was suspended in dichloromethane (20 mL) and cooled to –20 °C. *N*-Methylmorpholine (1.2 mL, 10.9 mmol), followed by isobutylchloroformate (1.5 mL, 11.6 mmol), was added dropwise. After 10 minutes, the hydrochloride salt of amino acid methyl ester (10.1 mmol) was added, followed by *N*-methylmorpholine (1.2 mL, 10.9 mmol). The resulting mixture was stirred for 1 hour at room temperature and then washed sequentially with 0.2 M hydrochloric acid (2 x 20 mL) and brine (10 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow oil. Diethyl ether (20 mL) was added to the residue and the resulting precipitate was filtered and washed with ice-cold diethyl ether (2 x 10 mL) to give the proline-containing dipeptide derivative.

Method (c):[43c] The N-acetylated amino acid (3.09 mmol) and HOBt·H_2O (0.48 g, 3.12 mmol) were suspended in dichloromethane (10 mL) under inert atmosphere and cooled to 0 °C. Diisopropylethylamine (0.65 mL, 3.73 mmol), followed by HBTU (1.15 g, 3.04 mmol), was then added and the resulting mixture was stirred for 1 hour at room temperature. This solution was then added to an ice-cold solution of the hydrochloride salt of amino acid methyl ester hydrochloride salt (3.28 mmol) and diisopropylethylamine (0.65 mL, 3.73 mmol) in dichloromethane (10 mL). The reaction mixture was stirred at 0 °C for 3 hours and quenched by the addition of 10% aqueous citric acid (15 mL). The mixture was filtered and the solid was washed with dichloromethane (30 mL). The filtrate was washed sequentially with 10% aqueous citric acid (15 mL), saturated aqueous sodium bicarbonate (15 mL) and brine (15 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow oil. The crude product was purified by column chromatography to give the proline-containing dipeptide derivative.

***N*-Acetylglycyl-L-proline methyl ester (52c)**

Prepared according to GP12 method (a), purified by column chromatography (97:3 $CHCl_3$/MeOH, $KMnO_4$ stain). A colourless liquid, yield: 0.57 g (43%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.00 (t, *J* = 5.5 Hz, 1H), 4.29 (dd, *J* = 7.9, 3.3 Hz, 1H), 3.97 (dd, *J* = 17.3, 4.8 Hz, 1H), 3.82 (dd, *J* = 17.3, 4.8 Hz, 1H), 3.61 (s, 3H), 3.53 (q, *J* = 7.2 Hz, 2H), 1.92 (p, *J* = 7.8 Hz, 2H), 1.85 (s, 3H), 1.87–1.79 ppm (m, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.8, 169.8, 167.7, 58.9, 52.2, 46.0, 41.4, 29.0, 24.9, 22.8 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{10}H_{17}N_2O_4]^+$: 229.1183 $[M + H]^+$, found 229.1183, HRMS (ESI) *m*/*z* calcd. for $[C_{10}H_{16}N_2O_4Na]^+$: 251.1002 $[M + Na]^+$, found 251.1001.

_N_-Acetyl-L-valyl-L-proline methyl ester (53c)

Prepared according to GP12 method (a), purified by column chromatography (99:1 $CHCl_3$/MeOH, $KMnO_4$ stain). A colourless liquid, yield: 1.03 g (65%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.11 (d, J = 8.2 Hz, 1H), 4.41 (dd, J = 7.6, 2.4 Hz, 1H), 4.13 (ddd, J = 8.0, 6.4, 1.4 Hz, 1H), 3.62 (s, 3H), 3.53–3.42 (m, 2H), 2.07–1.95 (m, 1H), 1.95 (s, 3H), 1.93–1.81 (m, 3H), 1.79–1.72 (m, 1H), 0.87 ppm (d, J = 5.5 Hz, 6H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.1, 172.0, 168.3, 58.5, 57.3, 51.7, 47.5, 30.0, 29.2, 24.2, 22.3, 19.0, 18.2 ppm. HRMS (ESI) m/z calcd. for $[C_{13}H_{23}N_2O_4]^+$: 271.1652 $[M + H]^+$, found 271.1653, HRMS (ESI) m/z calcd. for $[C_{13}H_{22}N_2O_4Na]^+$: 293.1472 $[M + Na]^+$, found 293.1471.

_N_-Acetyl-L-phenylalanyl-L-proline methyl ester (54c)

Prepared according to GP12 method (a) using 2.12 g (10.2 mmol) of _N_-acetyl-L-phenylalanine and 1.87 g (11.3 mmol) of L-proline methyl ester hydrochloride salt, purified by column chromatography (49:1 $CHCl_3$/MeOH, $KMnO_4$ stain). A colourless liquid, yield: 0.67 g (21%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.29 (d, J = 8.3 Hz, 1H), 7.29–7.26 (m, 4H), 7.23–7.16 (m, 1H), 4.68 (td, J = 8.6, 5.3 Hz, 1H), 4.31 (dd, J = 8.7, 4.7 Hz, 1H), 3.69 (dt, J = 9.8, 6.9 Hz, 1H), 3.61 (s, 3H), 3.44 (dt, J = 9.8, 6.6 Hz, 1H), 2.93 (dd, J = 13.8, 5.3 Hz, 1H), 2.74 (dd, J = 13.8, 9.0 Hz, 1H), 2.18–2.10 (m, 1H), 1.93–1.85 (m, 2H), 1.84–1.77 (m, 1H), 1.75 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.2, 170.0, 168.9, 137.6, 129.2, 128.1, 126.4, 58.6, 51.9, 51.8, 46.5, 36.9, 28.6, 24.6, 22.3 ppm. HRMS (ESI) m/z calcd. for $[C_{17}H_{23}N_2O_4]^+$: 319.1658 $[M + H]^+$, found 319.1653, HRMS (ESI) m/z calcd. for $[C_{17}H_{22}N_2O_4Na]^+$: 341.1477 $[M + Na]^+$, found 341.1471.

_N_-Acetyl-L-prolyl-L-phenylalanine methyl ester (55c)

Prepared according to GP12 method (b). A white solid, yield: 2.12 g (66%). ^{1}H NMR (500 MHz, DMSO-d_6) major rotamer: δ 8.47 (d, J = 8.5 Hz, 1H), 7.23–7.17 (m, 5H), 4.61 (ddd, J = 10.9, 8.5, 4.6 Hz, 1H), 4.21 (dd, J = 8.9, 2.5 Hz, 1H), 3.65 (s, 3H), 3.37 (ddd, J = 11.6, 9.3, 3.0 Hz, 1H), 3.27 (ddd, J = 11.3, 8.5, 6.8 Hz, 1H), 3.14 (dd, J = 13.8, 4.6 Hz, 1H), 2.93 (dd, J = 13.7, 2.2 Hz, 1H), 2.11–2.04 (m, 1H), 1.72–1.63 (m, 3H), 1.61 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 171.84, 171.79, 168.64, 137.46, 129.03, 128.12, 126.41, 60.30, 52.88, 52.00, 46.16, 36.03, 31.49, 23.97, 22.21 ppm. HRMS (ESI) m/z calcd. for $[C_{17}H_{23}N_2O_4]^+$: 319.1658 $[M + H]^+$, found 319.1658, HRMS (ESI) m/z calcd. for $[C_{17}H_{22}N_2O_4Na]^+$: 341.1477 $[M + Na]^+$, found 341.1468.

_N_-Acetyl-L-prolyl-L-proline methyl ester (56c)

Prepared according to GP12 method (c), purified by column chromatography (49:1 $CHCl_3$/MeOH, $KMnO_4$ stain). A colourless liquid (0.16 g, 19%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 4.54 (dd, J = 8.8, 3.8 Hz, 1H), 4.28 (dd, J = 8.7, 4.5 Hz, 1H), 3.69 (t, J = 5.0 Hz, 1H), 3.60 (s, 3H), 3.57–3.45 (m, 3H), 2.22–2.06 (m, 2H), 1.97–1.89 (m, 6H), 1.90–1.76 ppm (m, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.3, 170.1, 167.8, 58.3, 57.0, 51.8, 47.5, 46.2, 28.4, 28.2, 24.6, 24.1, 22.1 ppm. HRMS (ESI) m/z calcd. for $[C_{13}H_{21}N_2O_4]^+$: 269.1496 $[M + H]^+$, found 269.1496, HRMS (ESI) m/z calcd. for $[C_{13}H_{20}N_2O_4Na]^+$: 291.1316 $[M + Na]^+$, found 291.1315.

_N_-Acetylglycyl-L-prolyl-L-phenylalanine methyl ester (57c)

*(a) N-Boc-L-proline (**51d**):*[44] L-Proline (2.83 g, 24.6 mmol) was suspended in dichloromethane (50 mL) and triethylamine (3.9 mL, 28.0 mmol) was added dropwise. Boc anhydride (7.81 g, 35.8 mmol) was then added in portions and the resulting mixture was stirred for 2.5 hours at room temperature. The solution was

then washed sequentially with saturated aqueous citric acid (20 mL), water (25 mL) and brine (25 mL). The organic layer was dried over anhydrous magnesium sulfate and concentrated under reduced pressure to give a yellow oil. Diethyl ether (20 mL) was added to the residue and the resulting yellow precipitate was filtered and recrystallised from dichloromethane/petroleum ether mixtures to give the desired product as colourless prismatic crystals (3.56 g, 67%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 12.45 (s, 1H), 4.05 (dd, J = 8.8, 4.0 Hz, 1H), 3.35–3.30 (m, 2H), 2.21–2.13 (m, 1H), 1.84–1.78 (m, 3H), 1.34 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 174.3, 153.1, 78.6, 58.6, 46.1, 30.3, 27.9, 23.1 ppm.

*(b) N-Boc-L-prolyl-L-phenylalanine methyl ester (**55e**):* Prepared according to GP12 method (b). A white powder, yield: 2.64 g (74%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.21 (d, J = 8.0 Hz, 1H), 7.31–7.21 (m, 5H), 4.53 (ddd, J = 9.9, 7.9, 5.1 Hz, 1H), 4.02 (dd, J = 8.8, 3.4 Hz, 1H), 3.61 (s, 3H), 3.40–3.33 (m, 1H), 3.27–3.18 (m, 1H), 3.06 (dd, J = 13.9, 5.1 Hz, 1H), 2.93 (dd, J = 13.9, 9.8 Hz, 1H), 1.75–1.56 (m, 4H), 1.21 ppm (s, 9H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 172.6, 172.0, 153.3, 137.3, 128.9, 128.2, 126.5, 78.4, 59.5, 53.3, 51.8, 46.4, 36.4, 30.7, 27.8, 22.8 ppm. HRMS (ESI) m/z calcd. for $[C_{20}H_{29}N_2O_5]^+$: 377.2077 $[M + H]^+$, found 377.2069, HRMS (ESI) m/z calcd. for $[C_{20}H_{28}N_2O_5Na]^+$: 399.1896 $[M + Na]^+$, found 399.1887.

*(c) L-Prolyl-L-phenylalanine methyl ester trifluoroacetate salt (**55b**):* Prepared according to GP5 using 1.41 g (3.75 mmol) of *N*-Boc-L-proyl-L-phenylalanine methyl ester. A white powder, yield: 1.45 g (99%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.42 (s, 1H), 9.00 (d, J = 7.6 Hz, 1H), 8.54 (s, 1H), 7.34–7.26 (m, 2H), 7.28–7.19 (m, 3H), 4.61–4.50 (m, 1H), 4.20–4.10 (m, 1H), 3.63 (s, 3H), 3.19 (td, J = 12.5, 6.0 Hz, 2H), 3.11 (dd, J = 14.0, 5.3 Hz, 1H), 2.96 (dd, J = 13.8, 10.0 Hz, 1H), 2.36–2.22 (m, 1H), 1.95–1.74 ppm (m, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.3, 168.4, 136.9,

129.0, 128.4, 126.7, 58.7, 54.0, 52.1, 45.7, 36.2, 29.5, 23.4 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{21}N_2O_3]^+$: 277.1552 $[M - CF_3CO_2]^+$, found 277.1550.

*(d) N-Acetylglycyl-L-prolyl-L-phenylalanine methyl ester (**57c**):* Prepared according to GP12 method (a) using 0.19 g (1.62 mmol) of *N*-acetylglycine and 0.69 g (1.76 mmol) of L-prolyl-L-phenylalanine methyl ester trifluoroacetate salt, purified by column chromatography (24:1 $CHCl_3$/MeOH, $KMnO_4$ stain). A colourless liquid, yield: 0.12 g (19%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.19 (d, *J* = 7.7 Hz, 1H), 7.99 (t, *J* = 5.4 Hz, 1H), 7.29–7.26 (m, 2H), 7.23–7.20 (m, 3H), 4.46 (q, *J* = 7.2, 6.5 Hz, 1H), 4.34 (dd, *J* = 8.9, 2.8 Hz, 1H), 3.89 (d, *J* = 5.5 Hz, 2H), 3.57 (s, 3H), 3.50–3.38 (m, 2H), 3.02 (dd, *J* = 13.8, 5.7 Hz, 1H), 2.93 (dd, *J* = 13.9, 9.0 Hz, 1H), 1.99–1.90 (m, 1H), 1.87 (s, 3H), 1.72–1.65 (m, 2H), 1.64–1.51 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 171.8, 171.6, 169.5, 167.3, 137.1, 129.1, 128.2, 126.6, 59.2, 53.5, 51.9, 45.8, 41.3, 36.6, 28.9, 23.9, 22.4 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{19}H_{26}N_3O_5]^+$: 376.1867 $[M + H]^+$, found 376.1869, HRMS (ESI) *m*/*z* calcd. for $[C_{19}H_{25}N_3O_5Na]^+$: 398.1687 $[M + Na]^+$, found 398.1686.

L-Pyroglutamic acid methyl ester (58b)

Prepared according to GP3. A colourless liquid, yield: 0.88 g (49%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.98 (s, 1H), 4.18 (ddd, *J* = 9.0, 4.2, 1.0 Hz, 1H), 3.67 (s, 3H), 2.41–2.28 (m, 1H), 2.19–2.09 (m, 2H), 2.04–1.91 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 177.0, 173.3, 54.6, 52.0, 28.9, 24.5 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_6H_{10}NO_3]^+$: 144.0661 $[M + H]^+$, found 144.0656, HRMS (ESI) *m*/*z* calcd. for $[C_{12}H_{19}N_2O_6]^+$: 287.1243 $[2M + H]^+$, found 287.1237.

_N_-Methyl-L-pyroglutamic acid methyl ester (58t)[45]

L-Pyroglutamic acid (2.60 g, 20.1 mmol) was added to a suspension of sodium hydride (60% dispersion in mineral oil, 2.19 g, 54.8 mmol) in dimethylformamide (60 mL) under inert atmosphere at 0 °C. The resulting mixture was stirred for 1 hour at room temperature and cooled again to 0 °C, followed by dropwise addition of methyl iodide (3.45 mL, 55.4 mmol). The resulting mixture was then stirred for 16 hours at room temperature. The solvent was removed under reduced pressure and water (75 mL) was added. The aqueous layer was extracted with dichloromethane (3 x 100 mL). The combined organic extracts were washed sequentially with water (75 mL), 5% aqueous lithium chloride (50 mL) and brine (50 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to give a yellow liquid. The crude product was purified by column chromatography (49:1 $CHCl_3$/MeOH, $KMnO_4$ stain) to give the desired product as a colourless liquid (0.58 g, 18%). ^{1}H NMR (400 MHz, DMSO-d_6): δ 4.25 (dd, J = 8.5, 3.2 Hz, 1H), 3.70 (s, 3H), 2.69 (s, 3H), 2.37–2.25 (m, 1H), 2.28–2.17 (m, 2H), 1.99–1.87 ppm (m, 1H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 174.3, 172.4, 60.6, 52.2, 28.8, 28.3, 22.1 ppm. HRMS (ESI) m/z calcd. for $[C_7H_{12}NO_3]^+$: 158.0812 $[M + H]^+$, found 158.0813, HRMS (ESI) m/z calcd. for $[C_7H_{11}NO_3Na]^+$: 180.0632 $[M + Na]^+$, found 180.0631.

_N_-Acetyl-_N_-methylglycine methyl ester (59c)

*(a) N-Methylglycine methyl ester hydrochloride salt (**59b**):* Prepared according to GP2 using 1.43 g (16.0 mmol) of *N*-methylglycine. A white solid, yield: 1.92 g (86%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.55 (s, 2H), 3.93 (s, 2H), 3.73 (s, 3H), 2.54 ppm (s, 3H). ^{13}C {^{1}H} NMR (126 MHz, DMSO-d_6): δ 167.1, 52.5, 47.8, 32.5 ppm. HRMS (ESI) m/z calcd. for $[C_4H_{10}NO_2]^+$: 104.0712 $[M – Cl]^+$, found 104.0709.

*(b) N-Acetyl-N-methylglycine methyl ester (**59c**):* Prepared according to the procedure for *N*-acetyl-L-serine methyl ester using 1.03 g (7.41 mmol) of *N*-methylglycine methyl ester hydrochloride salt and 0.55 mL (7.74 mmol) of acetyl chloride, purified by column chromatography (EtOAc, $KMnO_4$ stain). A colourless liquid, yield: 0.59 g (55%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 4.05 (s, 2H), 3.63 (s, 3H), 3.01 (s, 3H), 2.02 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 170.5, 169.9, 51.6, 48.7, 36.8, 21.1 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_6H_{12}NO_3]^+$: 146.0812 $[M + H]^+$, found 146.0812, HRMS (ESI) *m*/*z* calcd. for $[C_6H_{11}NO_3Na]^+$: 168.0632 $[M + Na]^+$, found 168.0631.

_N_-Acetyl-L-phenylalanyl-_N_-methylglycine methyl ester (60c)

Prepared according to GP12 method (a), purified by column chromatography (EtOAc, PMA stain). A colourless gel, yield: 1.35 g (80%). ^{1}H NMR (400 MHz, DMSO-d_6) major rotamer: δ 8.27 (d, *J* = 8.6 Hz, 1H), 7.27–7.23 (m, 4H), 7.18–7.16 (m, 1H), 4.94 (td, *J* = 8.7, 5.5 Hz, 1H), 4.13 (d, *J* = 17.2 Hz, 1H), 3.99 (d, *J* = 17.1 Hz, 1H), 3.63 (s, 3H), 3.03 (s, 3H), 2.93 (dd, *J* = 13.7, 5.5 Hz, 1H), 2.74 (dd, *J* = 13.8, 8.9 Hz, 1H), 1.76 ppm (s, 3H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6) major rotamer: δ 171.8, 169.5, 168.7, 137.5, 129.2, 128.1, 126.4, 51.7, 49.7, 49.2, 37.2, 36.0, 22.2 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{21}N_2O_4]^+$: 293.1501 $[M+H]^+$, found 293.1496, HRMS (ESI) *m*/*z* calcd. for $[C_{15}H_{20}N_2O_4Na]^+$: 315.1321 $[M+Na]^+$, found 315.1313.

Laser flash photolysis studies

Time-dependent decay plots of $NO_3^\bullet$ signals at 630 nm for its reaction with each compound are fitted according to the exponential decay fitting available on Igor Pro 6.03 (WaveMetrics, Inc.; Lake Oswego, OR, USA). The fitting used for the calculations follows the equation: $y = y_o + Ae^{-invTau*x}$; where y is the optical density, OD (proportional to $[NO_3^\bullet]$) and x is the time (unit in ns). Each coefficient value (i.e., y_o, A, and $invTau$) is obtained from the fitting software along with one standard deviation. The value of k_{obs} corresponds to the value of $invTau$ (unit in ns^{-1}).

Second-order rate coefficients, k, for the reactions of $NO_3^\bullet$ with all substrates were obtained from the slopes of k_{obs} vs [substrate] plots. Typical plots are shown in Appendices, page 226–228. The intercept is due to the reaction of $NO_3^\bullet$ with the solvent (see page 53–54).

Error determination: each result is reported as the average of three different runs. For example, if three different runs produced this set of data: $(x_1 \pm \sigma_1)$, $(x_2 \pm \sigma_2)$ and $(x_3 \pm \sigma_3)$, then the result would be taken as $\frac{x_1 + x_2 + x_3}{3}$ and σ is the overall standard deviation obtained following the statistical equation below:

$$\sigma = \sqrt{\frac{{\sigma_1}^2 + {\sigma_2}^2 + {\sigma_3}^2}{3} + \frac{2}{9}({x_1}^2 + {x_2}^2 + {x_3}^2 - x_1 \cdot x_2 - x_2 \cdot x_3 - x_3 \cdot x_1)}, \text{ where}$$

x_1, x_2, x_3 are the result from each respective run and $\sigma_1, \sigma_2, \sigma_3$ are the standard deviation of each result. Overall errors are given as 2σ statistical uncertainties. Typical errors measured in these experiments are around 5–15%.

Reaction of $NO_2^\bullet$ with peptides

The liquid $NO_2^\bullet$ (0.5 mL, condensed at –20 °C) was injected into a solution of the peptide (1 mmol) in acetonitrile (100 mL) at 10 °C. The reaction mixture was stirred for 20 minutes and saturated aqueous sodium bicarbonate (50 mL) was added. Acetonitrile was then removed under reduced pressure and the crude product was extracted by ethyl acetate (3 x 100 mL). The combined organic layers were dried over magnesium sulfate and the solvent was removed under reduced pressure to give a yellow oil or powder.

***N*-Nitro-L-prolyl-L-*p*-nitrophenylalanine methyl ester (127B)**

Isolated according to the procedure above with a stream of ozonised oxygen (Scheme 5.3-2, page 115), purified by column chromatography (2:3 Pet. ether/EtOAc, UV light). White needles. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.74 (d, J = 8.0 Hz, 1H), 8.15 (d, J = 8.7 Hz, 2H), 7.52 (d, J = 8.6 Hz, 2H), 4.67–4.59 (m, 2H), 3.84–3.57 (m, 2H), 3.63 (s, 3H), 3.23 (dd, J = 13.9, 5.3 Hz, 1H), 3.10 (dd, J = 13.9, 9.6 Hz, 1H), 2.29–2.20 (m, 1H), 1.96–1.88 (m, 1H), 1.84–1.74 ppm (m, 2H). ^{13}C {^{1}H} NMR (101 MHz, DMSO-d_6): δ 171.2, 169.3, 146.4, 145.4, 130.5, 123.2, 62.5, 52.8, 52.1, 50.6, 36.0, 29.2, 22.0 ppm. HRMS (ESI) m/z calcd. for $[C_{15}H_{19}N_4O_7]^+$: 367.1248 $[M + H]^+$, found 367.1249, HRMS (ESI) m/z calcd. for $[C_{15}H_{18}N_4O_7K]^+$: 405.0807 $[M + K]^+$, found 405.0809.

Reaction of $NO_2^{\bullet}$ with cholesterol (161A)

The liquid $NO_2^{\bullet}$ (0.5 mL, condensed at –20 °C) was injected into a solution of cholesterol (0.39 g, 1.01 mmol) in dichloromethane (100 mL) at room temperature and the resulting mixture was stirred for 2 hours. Saturated aqueous sodium bicarbonate (50 mL) was then added and the organic layer was separated. The aqueous layer was extracted with dichloromethane (50 mL) and the combined organic layers were dried over sodium sulfate and concentrated under reduced pressure to give a yellow viscous oil. The crude product was then purified by column chromatography (3:7 Pet. ether/EtOAc, PMA stain) to give 5-nitrato-6-nitroimino cholesterol (**163A**) as a colourless powder (0.13 g, 25 %).

^{1}H NMR (400 MHz, $CDCl_3$): δ 3.95 (tt, *J* = 11.4, 4.8 Hz, 1H, *H*-1), 2.47 (ddd, *J* = 15.1, 4.7, 2.0 Hz, 1H, *H*-4), 2.40 (dd, *J* = 13.2, 4.1 Hz, 1H, *H*-4), 2.09 – 1.92 (m, 4H), 1.91 – 1.57 (m, 6H), 1.56 – 1.43 (m, 4H), 1.40 – 1.18 (m, 8H), 1.19 – 1.10 (m, 4H), 1.08 (s, 3H), 1.04 – 0.97 (m, 1H), 0.89 (d, *J* = 6.5 Hz, 3H), 0.86 (dd, *J* = 6.6, 1.8 Hz, 6H), 0.66 ppm (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$): δ 176.13 (*C*-6), 93.14 (*C*-5), 66.48 (*C*-1), 56.08, 55.99, 46.56, 45.53, 43.49, 39.59, 39.36, 38.98, 36.17, 35.78, 33.30, 32.60, 31.13, 30.07, 28.15, 28.08, 23.92, 22.94, 22.70, 21.83, 18.72, 15.02, 12.24 ppm. HRMS (ESI) *m*/*z* calcd. for $[C_{27}H_{46}N_3O_6]^+$: 508.3387 $[M + H]^+$, found 508.3401, HRMS (ESI) *m*/*z* calcd. for $[C_{27}H_{49}N_4O_6]^+$: 525.3652 $[M + NH_4]^+$, found 525.3637. IR (neat): ν_{max} 831, 1290, 1310, 1574, 1638, 2869, 2948 cm^{-1}.

References

1. a) Firket, J. *Trans. Faraday Soc.* **1936,** *32,* 1192–1197; b) Logan, W. P. D. *Lancet* **1953,** *261,* 336–338; c) Greenburg, L.; Jacob, M. B.; Drolette, B. M.; Field, F.; Braverman, M. M. *Public Health Rep.* **1962,** *77,* 7–16; d) Schwartz, J. *Environ. Res.* **1994,** *64,* 36–52; e) Helfand, W. H.; Lazarus, J.; Theerman, P. *Am. J. Public Health* **2001,** *91,* 553; f) Nemery, B.; Hoet, P. H.; Nemmar, A. *Lancet* **2001,** *357,* 704–708; g) Bell, M. L.; Davis, D. L. *Eniron. Health Perspect.* **2001,** *109* (suppl 3), 389–394; h) Brunekreef, B.; Holgate, S. T. *Lancet* **2002,** *360,* 1233–1242.
2. Stone, V. *Am. J. Respir. Crit. Care Med.* **2000,** *162* (suppl 1), 44–47.
3. a) Burden of disease from the joint effects of Household and Ambient Air Pollution for 2012, 2014. World Health Organization Web site. http://www.who.int/phe/health_topics/outdoorair/databases/FINAL_HAP_AAP_BoD_24March2014.pdf (accessed September 13, 2018); b) World Health Organization Home Page. http://www.who.int/mediacentre/news/releases/2014/air-pollution/en (accessed September 13, 2018).
4. a) National Environmental Protection (Ambient Air Quality) Measure Review, 2011. National Environmental Protection Council Web site. http://www.nepc.gov.au/system/files/resources/3405e986-afe9-bdb4-5d2c-383f3ea1e911/files/aaq-review-report-2011.pdf (accessed September 13, 2018); b) India Celebrating Web site. https://www.indiacelebrating.com/environmental-issues/air-pollution/ (accessed September 13, 2018).
5. a) Pick, H. Z. *Elektrochem.* **1920,** *26,* 182–196; b) Koppenol, W. H. *FEBS Lett.* **1982,** *140,* 169–172.
6. a) Ehrlich, R. *Environ. Health Perspect.* **1980,** *35,* 89–100; b) Bayram, H.; Sapsford, R. J.; Abdellaziz, M. M.; Khair, O. A. *J. Allergy Clin. Immun.* **2001,** *107,* 287–294; c) Sunyer, J.; Basagña, X.; Belmonte, J.; Antó, J. M. *Thorax* **2002,** *57,* 687–693; d) Chauhan, A. J.; Johnston, S. L. *Brit. Med. Bull.* **2003,** *68,* 95–112.
7. a) Jarvis, D. J.; Adamkiewicz, G.; Heroux, M.-E.; Rapp, R.; Kelly, F. J. Nitrogen Dioxide. *WHO Guidelines for Indoor Air Quality: Selected Pollutants*; WHO Regional Office for Europe: Denmark, 2010; pp 201–287; b) National Pollutant Inventory. http://www.npi.gov.au/npidata (accessed September 14, 2018).
8. a) Schindler, T. L.; Gray, E. T.; Sharghi, K.; Duncan, B.; Lamsal, L. NASA Images Show Human Fingerprint on Global Air Quality, 2015. NASA Website. https://svs.gsfc.nasa.gov/12094 (accessed September 14, 2018); b) Cooper, O. NOAA: Surface Ozone Around the Globe, 2016. InsideClimate News Web site. https://insideclimatenews.org/news/19032018/global-warming-arctic-air-pollution-short-lived-climate-pollutants-methane-black-carbon-hfcs-slcp (accessed September 14, 2018).
9. a) Ehrlich, R.; Findlay, J. C.; Fenters, J. D.; Gardner, D. E. *Environ. Res.* **1977,** *14,* 223–231; b) Ehrlich, R.; Findlay, J. C.; Gardner, D. E. *J. Toxicol. Env. Health* **1979,** *5,* 631–

642; c) Ichinose, T.; Sagai, M. *Toxicology* **1989,** *59,* 259–270; d) Gelzleichter, T. R.; Witschi, H.; Last, J. A. *Toxicol. Appl. Pharm.* **1992,** *112,* 73–80.

10. a) Franze, T.; Weller, M. G.; Niessner, R.; Pöschl, U. *Analyst* **2003,** *128,* 824–831; b) Franze, T.; Weller, M. G.; Niessner, R.; Pöschl, U. *Environ. Sci. Technol.* **2005,** *39,* 1673–1678; c) Shiraiwa, M.; Selzle, K.; Yang, H.; Sosedova, Y.; Ammann, M.; Pöschl, U. *Environ. Sci. Technol.* **2012,** *46,* 6672–6680.
11. Wayne, R. P.; Barnes, I.; Biggs, P.; Burrows, J. P.; Canosa-Mas, C. E.; Hjorth, J.; Le Bras, G.; Moortgat, G. K.; Perner, D.; Restelli, G.; Sidebottom, H. *Atmos. Environ.* **1991,** *25A,* 1–203.
12. a) Orlando, J. J.; Tyndall, G. S.; Moortgat, G. K.; Calvert, J. G. *J. Phys. Chem.* **1993,** *97,* 10996–11000; b) Monks, P. S. *Chem. Soc. Rev.* **2005,** *34,* 376–395.
13. a) Geyer, A.; Alicke, B.; Konrad, S.; Schmitz, T.; Stutz, J.; Platt, U. *J. Geophys. Res.* **2001,** *106,* 8013–8025; b) Platt, U.; Alicke, B.; Dubois, R.; Geyer, A.; Hofzumahaus, A.; Holland, F.; Martinez, M.; Mihelcic, D.; Klüpfel, T.; Lohrmann, B.; Pätz, W; Perner, D.; Rohrer, F.; Schäfer, J.; Stutz, J. *J. Atmos. Chem.* **2002,** *42,* 359–394; c) Asaf, D.; Tas, E.; Pedersen, D.; Peleg, M.; Luria, M. *Environ. Sci. Technol.* **2010,** *44,* 5901–5907.
14. Neta, P.; Huie, R. E.; Ross, A. B. *J. Phys. Chem. Ref. Data* **1988,** *17,* 1027–1284.
15. Geyer, A.; Alicke, B.; Ackermann, R.; Martinez, M.; Harder, H.; Brune, W.; di Carlo, P.; Williams, E.; Jobson, T.; Hall, S.; Shetter, R.; Stutz, J. *J. Geophys. Res.* **2003,** *108,* 4368–4378.
16. Kampa, M.; Castanas, E. *Environ. Pollut.* **2008,** *151,* 362–367.
17. Kelly, F. J. *Occup. Environ. Med.* **2003,** *60,* 612–616.
18. a) Cross, C. E.; van der Vliet, A.; O'Neill, C. A.; Louise, S.; Halliwell, B. *Environ. Health Perspect.* **1994,** *102* (suppl. 10), 185–191; b) Kelly, F. J.; Cotgrove, M.; Mudway, I. S. *Cent. Eur. J. Pub. Health* **1996,** *4* (suppl.), 11–14; c) Mudway, I. S.; Kelly, F. J. *Toxicol. Appl. Pharm.* **1998,** *148,* 91–100; d) Kelly, F. J. *Food Chem. Toxicol.* **1999,** *37,* 963–966.
19. a) Pope, C.A., III *Environ. Health Perspect.* **1999,** *107,* 613–614; b) Kelly, F. J.; Mudway, I.; Blomberg, A.; Frew, A.; Sandström, T. *Lancet* **1999,** *354,* 482–483; c) Kelly, F. J.; Dunster, C.; Mudway, I. *Eur. Respir. J.* **2003,** *21* (suppl. 40), 70s–75s; d) Bernstein, J. A. *J. Allergy Clin. Immunol.* **2004,** *114,* 1116–1123.
20. Garrett, R. H.; Grisham, C. M. Molecular Components of Cells. In *Biochemistry*, 4th ed.; Cengage Learning: Boston, MA, 2008; p 23
21. Tait, M. J.; Franks, F. *Nature* **1971,** *230,* 91–94.
22. a) Böhm, G. M.; Saldiva, P. H. N.; Pasqualucci, C. A. G.; Massad, E.; Martins, M. A.; Zin, W. A.; Cardoso, W. V.; Criado P. M. P.; Komatsuzaki, M.; Sakae, R. S.; Negri, E. M.; Lemos, M.; Capelozzi, V. M.; Crestana, C.; Silva, R. *Environ. Res.* **1989,** *49,* 208–216; b) Carter, J. D.; Ghio, A. J.; Samet, J. M.; Devlin, R. B. *Toxicol. Appl. Pharmacol.* **1997,** *146,* 180–188; c) Somers, C. M.; Yauk, C. L.; White, P. A.; Parfett, C. L. J.; Quinn, J. S. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 15904–15907.
23. Davies, M. J. *Biochim. Biophys. Acta* **2005,** *1703,* 93–109.
24. Pulfer, M. K.; Harrison, K.; Murphy, R. C. *J. Am. Soc. Mass Spectrom.* **2004,** *15,* 430–437.

25. Gamon, L. F.; White, J. M.; Wille, U. *Org. Biomol. Chem.* **2014,** *12,* 8280–8287.
26. Kroon, M. C.; Spronsen, J.; Peters, C. J.; Sheldon, R. A.; Witkamp, G. J. *Green Chem.* **2006,** *8,* 246–249.
27. Salomé, C.; Salomé-Grosjean, E.; Park, K. D.; Morieux, P.; Swendiman, R.; DeMarco, E.; Stables, J. P.; Kohn, H. *J. Med. Chem.* **2010,** *53* (Suppl. S101), 1288–1305.
28. Yamada, S.; Sudo, A.; Goto, M.; Endo, T. *RSC Adv.* **2014,** *4* (Suppl. S2), 29890–29896.
29. Chaudhary, S.; Kumar, H.; Verma, H.; Rajpoot, A. *Int. J. Pharm. Tech. Res.* **2012,** *4,* 194–200.
30. Gordon-Wylie, S. W.; Teplin, S.; Morris, J. C.; Trombley, M. I.; McCarthy, S. M.; Cleaver, W. M.; Clark, G. R. *Cryst. Growth Des.* **2004,** *4,* 789–797.
31. Gillie, A. D.; Reddy, R. J.; Davies, P. W. *Adv. Synth. Catal.* **2016,** *358* (Suppl. S2), 226–239.
32. Meng, S.; Xu, Z.; Hong, G.; Zhao, L.; Zhao, Z.; Guo, J.; Ji, H.; Liu, T. *Eur. J. Med. Chem.* **2015,** *92* (Suppl. S6), 35–48.
33. Berlin, K. D.; Burpo, D. H.; Pagilagan, R. U.; Bude, D. *Chem. Commun. (London)* **1967,** No. 20, 1060–1061.
34. Kehler, J.; Breuer, E. *Synthesis* **1998,** No. 10, 1419–1420.
35. Antonello, S.; Formaggio, F.; Moretto, A.; Toniolo, C.; Maran, F. *J. Am. Chem. Soc.* **2001,** *123,* 9577–9584.
36. Piwowarska, N. A.; Banala, S.; Overkleeft, H. S.; Süssmuth, R. D. *Chem. Commun.* **2013,** *49* (Suppl. S5), 10703–10705.
37. Ekkati, A. R.; Campandi, A. A.; Abouelatta, A. I.; Shamoun, M.; Kalapugama, S.; Kelley, M.; Kokando, J. J. *Amino Acids* **2010,** *38,* 747–751.
38. Maegawa, T.; Akashi, A.; Esaki, H.; Aoki, F.; Sajiki, H. *Synlett.* **2005,** No. 5, 845–847.
39. Tsukano, C.; Yokouchi, S.; Girard, A.-L.; Kuribayashi, T.; Sakamoto, S.; Enomoto, T.; Takemoto, Y. *Org. Biomol. Chem.* **2012,** *10,* 6074–6086.
40. Chapkanov, A. G.; Dzimbova, T. A.; Ivanova, B. B. *Bulg. Chem. Commun.* **2012,** *44,* 228–232.
41. Gamon, L. F. Oxidative and chemical modification of amino acids by nitrogen dioxide, ozone and the reactive paracetamol metabolite NAPQI. PhD Thesis, The University of Melbourne, April 2016.
42. Kuvaeva, Z. I.; Lopatik,D. V.; Nikolaeva, T. A.; Knizhnikova, A. N.; Naidenov, V. E.; Markovich, M. M. *Pharm. Chem. J.* **2010,** *44,* 307–309.
43. (a) Steinmetz, H.; Li, J.; Fu, C.; Zaburannyi, N.; Kunze, B.; Harmrolfs, K.; Schmitt, V.; Herrmann, J.; Reichenbach, H.; Höfle, G.; Kalesse, M.; Müller, R. *Angew. Chem. Int. Ed.* (English) **2016,** *55* (Suppl. S35), 10113–10117; (b) Beaulieu, P. L.; Lavallée, P.; Abraham, A.; Anderson, P. C.; Boucher, C.; Bousquet, Y.; Duceppe, J.-S.; Gillard, J.; Gorys, V.; Grand-Maître, C.; Grenier, L.; Guindon, Y.; Guse, I.; Plamondon, L.; Soucy, F.; Valois, S.; Wernic, D.; Yoakim, C. *J. Org. Chem.* **1997,** *62,* 3440–3448; (c) Kubyshkin, V.; Durkin, P.; Budisa, N. *New J. Chem.* **2016,** *40* (Suppl. S4), 5209–5220.
44. Wu, X.; He, C.; Wu, X.; Qu, S.; Duan, C. *Chem. Commun.* **2011,** *47* (Suppl. S2), 8415–8417.

45. Gao, M.; Wang, M.; Green, M. A.; Hutchins, G. D.; Zheng, Q. H. *Bioorg. Med. Chem.* **2015,** *25,* 1965–1970.
46. a) Buxton, G. V.; Greenstock, C. L.; Helman, W. P.; Ross, A. B. *J. Phys. Chem. Ref. Data* **1988,** *17*, 513–886; b) Hawkins, C. L.; Davies, M. J. *J. Chem. Soc., Perkins Trans. 2* **1998,** 2617–2622; c) Štefanić, I.; Bonifačić, M.; Asmus, K.; Armstrong, D. A. *J. Phys. Chem. A* **2001,** *105*, 8681–8690; d) Połczyński, P.; Jurczakowski, R.; Grochala, W. *Chem. Commun.* **2013,** *49,* 7480–7482.
47. a) Chan, B.; O'Reilly, R. J.; Easton, C. J.; Radom, L. *J. Org. Chem.* **2012,** *77,* 9807–9812; b) Bosio, G.; Criado, S.; Massad, W.; Nieto, F. J. R.; Gonzalez, M. C.; García, N. A.; Mártire, D. O. *Photochem. Photobiol. Sci.* **2005,** *4*, 840–846; c) Salamone, B.; Basili, F.; Bietti, M. *J. Org. Chem.* **2015,** *80*, 3643–3650; d) Bonifačić, M.; Štefanić, I.; Hug, G. L.; Armstrong, D. A.; Asmus, K. *J. Am. Chem. Soc.* **1998,** *120,* 9930–9940.
48. a) Schwarzberg, M.; Sperling, J.; Elad, D. *J. Am. Chem. Soc.* **1973,** *95*, 6418–6426; b) Easton, C. J. *Chem. Rev.* **1997,** *97*, 53–82; c) Ricci, M.; Blakskjær, P.; Skrydstrup, T. *J. Am. Chem. Soc.* **2000,** *122*, 12413–12421.
49. a) Viehe, H. G.; Merényi, R.; Stella, L.; Janousek, Z. *Angew. Chem. Int. Ed. Engl.* **1979,** *18*, 917–932; b) Viehe, H. G.; Janousek, Z.; Merenyi, R. *Acc. Chem. Res.* **1985,** *18*, 148–154.
50. Croft, A. K.; Easton, C. J.; Radom, L. *J. Am. Chem. Soc.* **2003,** *125*, 4119–4124
51. Zipse, H. Radical Stability—A Theoretical Perspective. In *Radical in Synthesis I. Topics in Current Chemistry;* Gansäuer, A., Eds.; Springer: Berlin, Heidelberg, 2006; Vol. 263, pp 163–189.
52. Hioe, J.; Savasci, G.; Brand, H.; Zipse, H. *Chem. Eur. J.* **2011,** *17*, 3781–3789.
53. Garrison, W. M. *Chem. Rev.* **1987,** *87*, 381–398.
54. a) Easton, C. J.; Hay, M. P. *J. Chem. Soc., Chem. Commun.* **1986,** 55–57; b) Burgess, V. A.; Easton, C. J.; Hay, M. P. *J. Am. Chem. Soc.* **1989,** *111*, 1047–1052.
55. a) Easton, C. J.; Hutton, C. A.; Rositano, G.; Tan, E. W. *J. Org. Chem.* **1991,** *56*, 5614–5618; b) Croft, A. K.; Easton, C. J. Kociuba, K.; Radom, L. *Tetrahedron: Asymmetry* **2003,** *14*, 2919–2926.
56. Venkatachalapathy, B.; Ramamurthy, P. *J. Photochem. Photobiol. A* **1996,** *93*, 1–5.
57. a) Ito, O.; Akiho, S.; Iino, M. *J. Org. Chem.* **1989,** *54*, 2436–2440; b) Alfassi, Z.B.; Padmaja, S.; Neta, P.; Huie, R. E. *J. Phys. Chem.* **1993,** *97*, 3780–3782; c) Giacco, T. D.; Baciocchi, E.; Steenken, S. *J. Phys. Chem.* **1993,** *97*, 5451–5456.
58. a) Krüger, O.; Wille, U. *Org. Lett.* **2001,** *3,* 1455–1458; b) Cataldo, F. *Int. J. Biol. Macromol.* **2006,** *38,* 248–254; c) Sigmund, D. C. E.; Wille, U. *Chem. Commun.* **2008,** No. 18, 2121–2123; d) Wille, U.; Goeschen, C. *Aust. J. Chem.* **2011,** *64,* 833–842.
59. Biggs, P.; Canosa-Mas, C. E.; Monks, P. S.; Wayne, R. P.; Benter, Th.; Schindler, R. N. *Int. J. Chem. Kinet.* **1993,** *25,* 805–817.
60. a) Markham, G. D.; Parkin, D. W.; Mentch, F.; Schramm, V. L. *J. Biol. Chem.* **1987,** *262,* 5609–5615; b) Geimer, J.; Beckert, D.; Jenichen, A. *Chem. Phys. Lett.* **1997,** *280,* 353–358; c) Adam, W.; Krebs, O.; Orfanopoulos, M.; Startakis, M.; Vougiouskalakis, G. C. *J. Org. Chem.* **2003,** *68,* 2420–2425; d) Bonifačić, M.; Armstrong, D. A.; Štefanić, I.; Asmus, K. *J. Phys. Chem. B* **2003,** *107,* 7268–7276.

61. Seebach, D.; Beck, A. K.; Studer, A. Some Effects of Lithium Salts, of Strong Bases, and of the Cosolvent DMPU in Peptide Chemistry, and Elsewhere. In *Modern Synthetic Methods;* Ernst, B., Leumann, C., Eds.; Wiley-VCH: New York, 1995; p 90.
62. Hwu, J. R.; Jain, M. L.; Tsay, S.; Hakimelahi, G. H. *Tetrahedron Lett.* **1996,** *37,* 2035–2038.
63. a) Robinson, M. G.; Weiss, J. J.; Wheeler, C. M. *Biochim. Biophys. Acta* **1966,** *124,* 176–180; b) Dizdaroglu, M.; Gajewski, E. *Cancer Res.* **1989,** *49,* 3463–3467; c) Xiao, H.; Cai, G.; Liu, M. *Spectroscopy* **2007,** *21,* 91–103; d) Piscopo, M.; Trifuoggi, M.; Scarano, C.; Gori, C.; Giarra, A.; Febbraio, G. *Sci. Rep.* **2018,** *8* (7414), 1–10.
64. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, F. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09*, Revision C.01, Gaussian, Inc., Wallingford CT, **2010**.
65. Chan, B.; Karton, A.; Easton, C. J.; Radom, L. *J. Chem. Theory Comput.* **2016,** *12*, 1606–1613.
66. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008,** *120*, 215–241.
67. a) Mandal, D.; Sahu, C.; Bagchi, S.; Das, A. K. *J. Phys. Chem. A* **2013,** *117*, 3739–3750; b) Peverati, R.; Truhlar, G. *Phil. Trans. Royal Soc. A* **2014,** *372* (2011), 1–52.
68. Müller, J.-F.; Liu, Z.; Nguyen, V. S.; Stavrakou, T.; Harvey, J. N.; Peeters, J. *Nature Commun.* **2016,** *7* (No. 13213), 1–11.
69. a) Hammes-Schiffer, S.; Stuchebrukhov, A. A. *Chem. Rev.* **2010,** *110,* 6939–6960; b) Choi, G. J.; Knowles, R. R. *J. Am. Chem. Soc.* **2015,** *137,* 9226–9229; c) Soudackov, A. V.; Hammes-Schiffer, S. *Faraday Discuss.* **2016,** *195,* 171–189.
70. a) Hawkins, C. L.; Davies, M. J. *Biochem. J.* **1998,** *332,* 617–625; b) Hawkins, C. L.; Davies, M. K. *J. Chem. Soc., Perkins Trans. 2* **1998,** 1937–1946; c) Pattison, D. I.; Davies, M. J.; Asmus, K.-D. *J. Chem. Soc., Perkins Trans. 2* **2002,** 1461–1467; d) Ito, T.; Morimoto, S.; Fujita, S.; Kobayashi, K.; Tagawa, S.; Nishimoto, S. *Chem. Phys. Lett.* **2008,** *462,* 116–120.
71. a) Chan, B.; Moran., D.; Easton, C. J.; Radom, L. *Chem. Asian J.* **2017,** *12,* 1485–1489; b) Chan, B.; Easton, C. J.; Radom, L. *J. Phys. Chem. A* **2018,** *122,* 1741–1746.
72. Harriman, A. *J. Phys. Chem.* **1987,** *91,* 6102–6104.
73. a) Berlett, B. S.; Stadtman, E. R. *J. Biol. Chem.* **1997,** *272,* 20313–20316; b) Stadtman, E. R.; Levine, R. L. *Amino Acids* **2003,** *25,* 207–218; c) Friguet, B. *FEBS Lett.* **2006,** *580,* 2910–2916.

74. a) Maskos, Z.; Rush, J. D.; Koppenol, W. H. *Arch. Biochem. Biophys.* **1992,** *296,* 521–529; b) Huggins, T. G.; Wells-Knecht, M. C.; Detorie, N. A.; Baynes, J. W.; Thorpe, S. R. *J. Biol. Chem.* **1993,** *268,* 12341–12347; c) Dean, R. T.; Gieseg, S.; Davies, M. J. *Trends Biochem. Sci.* **1993,** *18,* 437–441.
75. a) Armstrong, R. C.; Swallow, A. J. *Radiat. Res.* **1969,** *40,* 563–579; b) Fang, X.; Jin, F.; Jin, H.; Sonntag, C. *J. Chem. Soc., Perkins Trans. 2* **1998,** 259–263; c) Winchester, R. V.; Lynn, K. R. *Int. J. Radiat. Biol.* **1970,** *17,* 541–548.
76. Uchida, K.; Kawakishi, S. *FEBS Lett.* **1993,** *332,* 208–210.
77. a) Okumura, K.; Murachi, T. *J. Biochem.* **1975,** *77,* 913–918; b) Edgcomb, S. P.; Murphy, K. P. *Proteins: Struct. Funct. Bioinf.* **2002,** *49* (No. 1), 1–6.
78. a) Goeschen, C.; Wibowo, N.; White, J. M.; Wille, U. *Org. Biomol. Chem.* **2011,** *9,* 3380–3385; b) Gamon, L. F.; Wille, U. *Acc. Chem. Res.* **2016,** *49,* 2136–2145.
79. Merkel, P. B.; Luo, P.; Dinnocenzo, J. P.; Farid, S. *J. Org. Chem.* **2009,** *74,* 5163–5173.
80. Ji, J. A.; Zhang, B.; Cheng, W.; Wang, Y. J. *J. Pharm. Sci.* **2009,** *98,* 4485–4500.
81. Baciocchi, E.; Del Giacco, T.; Murgia, S. M.; Sebastiani, G. V. *J. Chem. Soc., Chem. Commun.* **1987,** 1246–1248.
82. Croft, A. K.; Easton, C. J. *Aust. J. Chem.* **2004,** *57,* 651–654.
83. a) Wynne-Jones, W. F. K. *J. Chem. Phys.* **1934,** *2,* 381–385; b) Westheimer, F. H. *Chem. Rev.* **1961,** *61,* 265–273; c) Geimer, J.; Beckert, D.; Jenichen, A. *Chem. Phys. Lett.* **1997,** *280,* 353–358.
84. a) Glass, R. S.; Hug, G. L.; Schöneich, C.; Wilson, G. S.; Kuznetsova, L.; Lee, T. M.; Ammam, M.; Lorance, E.; Nauser, T.; Nichol, G. S.; Yamamoto, T. *J. Am. Chem. Soc.* **2009,** *131,* 13791–13805; b) Glass, R. S.; Schöneich, C.; Wilson, G. S.; Nauser, T.; Yamamoto, T.; Lorance, E.; Nichol, G. S.; Ammam, M. *Org. Lett.* **2011,** *13,* 2837–2839.
85. a) Gray, H. G.; Winkler, J. R. *Chem. Phys. Lett.* **2009,** *483,* 1–9; b) Williamson, H. R.; Dow, B. A.; Davidson, V. L. *Bioorg. Chem.* **2014,** *57,* 213–221; c) Shah, A.; Adhikari, B.; Martic, S.; Munir, A.; Shahzad, S.; Ahmad, K.; Kraatz, H.-B. *Chem. Soc. Rev.* **2015,** *44,* 1015–1027.
86. a) Giese, B.; Graber, M.; Cordes, M. *Curr. Opin. Chem. Biol.* **2008,** *12,* 755–759; b) Bixon, M.; Giese, B.; Wessely, S.; Langenbacher, T.; Michel-Beyerle, M. E.; Jortner, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 11713–11716.
87. a) Wang, M.; Gao, J.; Müller, P.; Giese, B. *Angew. Chem. Int. Ed.* **2009,** *48,* 4232–4234; b) Giese, B.; Wang, M.; Gao, J.; Stoltz, M.; Müller, P.; Graber, M. *J. Org. Chem.* **2009,** *74,* 3621–3625
88. Feliciano, G. T.; Steidl, R. J.; Reguera, G. *Phys. Chem. Chem. Phys.* **2015,** *17,* 22217–22226.
89. Seim, K. L.; Obermeyer, A. C.; Francis, M. B. *J. Am. Chem. Soc.* **2011,** *133,* 16970–16976.
90. (a) Morgan, A. A.; Rubenstein, E. *PLoS One* **2013,** *8* (No. 1), 1–9; (b) Meinikov, S.; Mailliot, J.; Rigger, L.; Neuner, S.; Shin, B. S.; Yusupova, G.; Dever, T. E.; Micura, R.; Yusupov, M. *EMBOS Rep.* **2016,** *17,* 1776–1784.
91. Moens, L.; Evans, R. J.; Looker, M. J.; Nimlos, M. R. *Fuel* **2004,** *83,* 1433–1443.

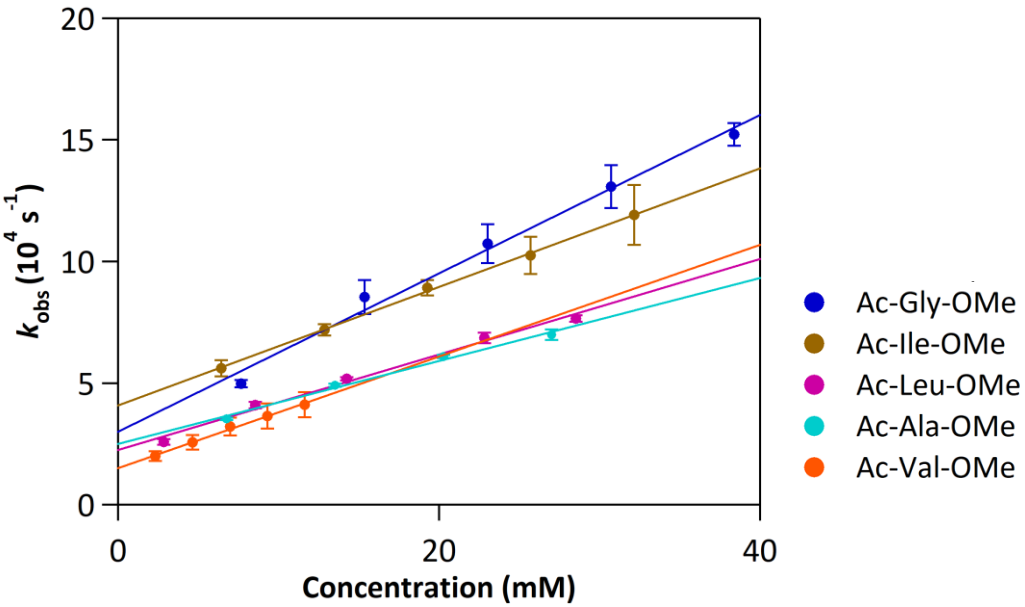
92. (a) Rustgi, S.; Joshi, A.; Riesz, P.; Friedberg, F. *Int. J. Radiat. Biol.* **1977,** *32,* 533–552; (b) Rustgi, S.; Joshi, A.; Moss, H.; Riesz, P. *Int. J. Radiat. Biol.* **1977,** *31,* 415–440.
93. Dean, R. T.; Wolff, S. P.; McElligott, M. A. *Free Radic. Res. Commun.* **1989,** *7,* 97–103.
94. (a) Kaul, S.; Sharma, S. S.; Mehta, I. K. *Amino Acids* **2008,** *34,* 315–320; (b) Roy, D.; Basu, N.; Bhunia, A.; Banerjee, S. K. *Biol. Plant.* **1993,** *35,* 69–72; (c) Hayat, S.; Hayat, Q.; Alyemeni, M. N.; Wani, A. S.; Pichtel, J.; Ahmad, A. *Plant Signal. Behav.* **2012,** *7,* 1456–1466.
95. (a) Signorelli, S.; Coitiño, E. L.; Borsani, O.; Monza, J. *J. Phys. Chem. B* **2014,** *118,* 37–47; (b) Signorelli, S.; Dans, P. D.; Coitiño, E. L.; Borsani, O.; Monza, J. *PLoS One* **2015,** *10* (No. 3), 1–14.
96. (a) Barnett, N. M.; Naylor, A. W. *Plant Physiol.* **1966,** *11,* 1222–1230; (b) Ashraf, M.; Foolad, M.R. *Environ. Exp. Bot.* **2007,** *59,* 206–261M; (c) Szabados, L.; Savouré, A. *Trends Plant Sci.* **2010,** *15,* 89–97.
97. Büssis, D.; Heineke, D. *J. Exp. Bot.* **1998,** *49,* 1361–1370.
98. (a) Noctor, G.; Foyer, C. H. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998,** *49,* 249–279; (b) Mittler, R. *Trends Plant Sci.* **2002,** *7,* 405–410.
99. Szoke, A.; Miao, G. H.; Hong, Z.; Verma, D. P. S. *Plant Physiol.* **1992,** *99,* 1642–1649.
100. Shelp, B. J.; Bown, A. W.; McLean, M. D. *Trends Plant Sci.* **1999,** *4,* 446–452.
101. Christgen, S. L.; Becker, D. F. *Antioxid. Redox Signal.* **2018,** *0* (No. 0), 1–27.
102. Warner, T. F.; Azen, E. A. *Am. Rev. Respir. Dis.* **1984,** *130,* 115–118.
103. a) Ehrlich, R. *Bateriol. Rev.* **1966,** *30,* 604–614; b) Sega, K.; Fugas, M. *J. Expo. Anal. Environ. Epidemiol.* **1991,** *1,* 227–234; c) Berglund, M. *Scand. J. Work Environ. Health* **1993,** *19,* 14–20.
104. Halliwell, B.; Hu, M. L.; Louie, S.; Duvall, T. R.; Tarkington, B. K.; Motchnik, P.; Cross, C. E. *FEBS Lett.* **1992,** *313,* 62–66.
105. a) Taylor, O. C.; Eaton, F. M. *Plant Physiol.* **1966,** *41,* 132–135; b) Cooney, R. V.; Harwood, P. J.; Custer, L. J.; Franke, A. A. *Environ. Health Perspect.* **1994,** *102,* 460–462.
106. a) Williams, R. A.; Rhoades, R. A.; Adams, W. S. *Arch. Intern. Med.* **1971,** *128,* 101–108; b) Drózdz, M.; Kucharz, E.; Szyja, J. *Environ. Res.* **1977,** *13,* 369–377; c) Schnizlein, C. T.; Bice, D. E.; Rebar, A. H.; Wolff, R. K.; Beethe, R. L. *Environ. Res.* **1980,** *23,* 362–370.
107. a) Little, C.; O'Brien, P. J. *Biochem. Biophys. Res. Comm.* **1968,** *31,* 145–150; b) Sagai, M.; Ichinose, T. *Environ. Health Perspect.* **1987,** *73,* 179–189; d) Sagai, M.; Ichinose, T.; Oda, H.; Kubota, K. *J. Toxicol. Environ. Health A.* **1982,** *9,* 153–164.
108. Walles, S. A. S.; Victorin, K.; Lundborg, M. *Mutat. Res.* **1995,** *328,* 11–19.
109. Krinsky, N. I. *Free Radic. Biol. Med.* **1989,** *7,* 617–635.
110. Hood, D. B.; Gettins, P.; Johnson, D. A. *Arch. Biochem. Biophys.* **1993,** *304,* 17–26.
111. a) Devalia, J. L.; Sapsford, R. J.; Cundell, D. R.; Rusznak, C.; Campbell, A. M.; Davies, R. J. *Eur. Respir. J.* **1993,** *6,* 1308–1316; b) Devalia, J. L.; Rusznak, C.; Herdman, M. J.; Trigg, C. J.; Tarraf, H.; Davies, R. J. *Lancet* **1994,** *344,* 1668–1671; c) Kikugawa, K.; Kato, T.; Okamoto, Y. *Free Radic. Biol. Med.* **1994,** *16,* 373–382; d) Gauderman, W. J.; Avol, E.; Lurmann, F.; Kuenzli, N.; Gilliland, F.; Peters, J.; McConnell, R.

Epidemiology **2005,** *16,* 737–743; e) Petit, P. C.; Fine, D. H.; Vasquez, G. B.; Gamero, L.; Slaughter, M. S.; Dasse, K. A. *ASAIO J.* **2017,** *63,* 7–13.

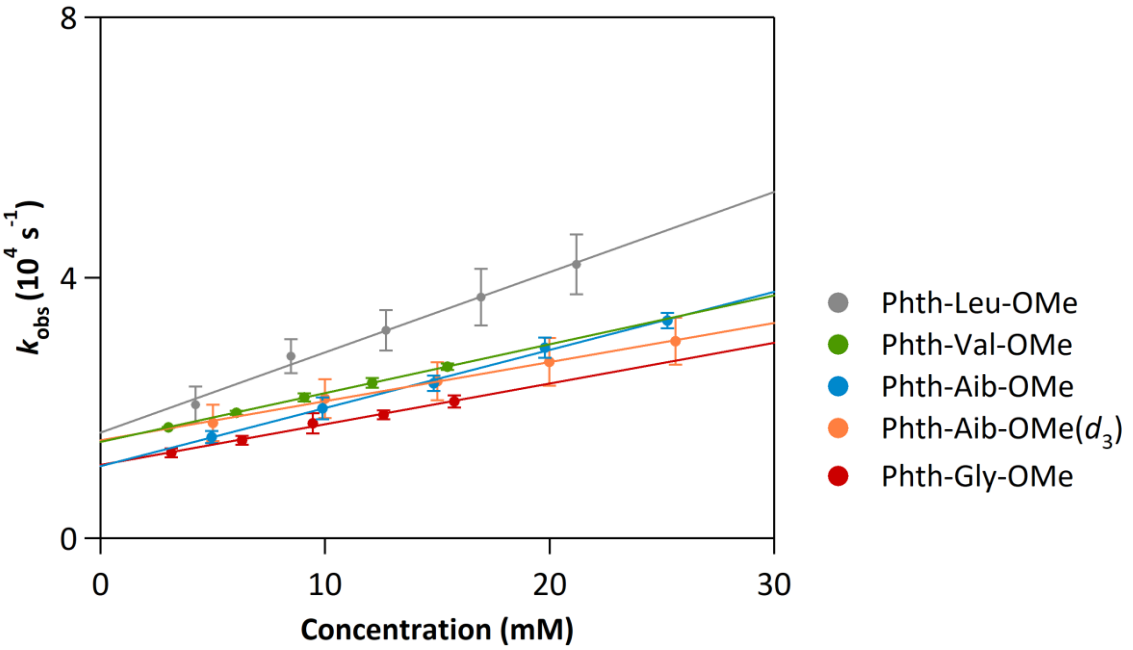
112. Devalia, J. L.; Campbell, A. M.; Sapsford, R. J.; Rusznak, C.; Quint, D.; Godard, P.; Bousquet, J.; Davies, R. J. *Am. J. Respir. Cell Mol. Biol.* **1993,** *9,* 271–278.
113. a) Pryor, W. A.; Lightsey, J. W.; Church, D. F. *J. Am. Chem. Soc.* **1982,** *104,* 6685–6692; b) Elsayed, N. M. *Toxicology* **1994,** *89,* 161–174; c) Pfeiffer, S.; Schmidt, K.; Mayer, B. *J. Biol. Chem.* **2000,** *275,* 6346–6352; d) Balazy, M.; Iesaki, T.; Park, J. L.; Jiang, H.; Kaminski, P. M.; Wolin, M. S. *J. Pharmacol. Exp. Ther.* **2001,** *299,* 611–619; e) Astolfi, P.; Panagiotaki, M.; Rizzoli, C.; Greci, L. *Org. Biomol. Chem.* **2006,** *4,* 3282–3290.
114. a) Gray, P.; Yoffe, A. D. *Chem. Rev.* **1955,** *55,* 1069–1154; b) Huie, R. E. *Toxicology* **1994,** *89,* 193–216; c) Lv, C. L.; Liu, Y. D.; Zhong, R. *J. Phys. Chem. A* **2008,** *112,* 7098–7105; d) Shiri, M.; Zolfigol, M. A.; Kruger, H. G.; Tanbakouchian, Z. *Tetrahedron* **2010,** *66,* 9077–9106.
115. Challis, B. C.; Milligan, J. R.; Mitchell, R. C. *J. Chem. Soc. Perkin Trans. 2* **1991,** No. 10, 1595–1599.
116. Gamon, L. F.; Nathanael, J. G.; Taggert, B. I.; Henry, F. A.; Bogena, J.; Wille, U. *Chem. Eur. J.* **2015,** *21,* 14924–14930.
117. Gold, B.; Deshpande, A.; Linder, W.; Hines, L. *J. Am. Chem. Soc.* **1984,** *106,* 2072–2077.
118. Farahani, H.; Hasan, M. *Pharmacol. Toxicol.* **1990,** *66,* 146–149.
119. Kamel, A. M.; Felmeister, A.; Weiner, N. D. *J. Pharm. Sci.* **1970,** *59,* 1807–1809.
120. a) Kobayashi, T.; Kubota, K. *Chemosphere* **1980,** *9,* 777–784; b) Mirvish, S. S.; Babcook, D. M.; Deshpande, A. D.; Nagel, D. L. *Cancer Lett.* **1986,** *31,* 97–104.
121. Adamopoulos, S.; Boulton, A. J.; Tadayoni, R.; Webb, G. A. *J. Chem. Soc. Perkin Trans. 1* **1987,** 2073–2077.
122. Enami, S.; Hoffmann, M. R.; Colussi, A. J. *J. Phys. Chem. B* **2009,** *113,* 7977–7981.

Appendices

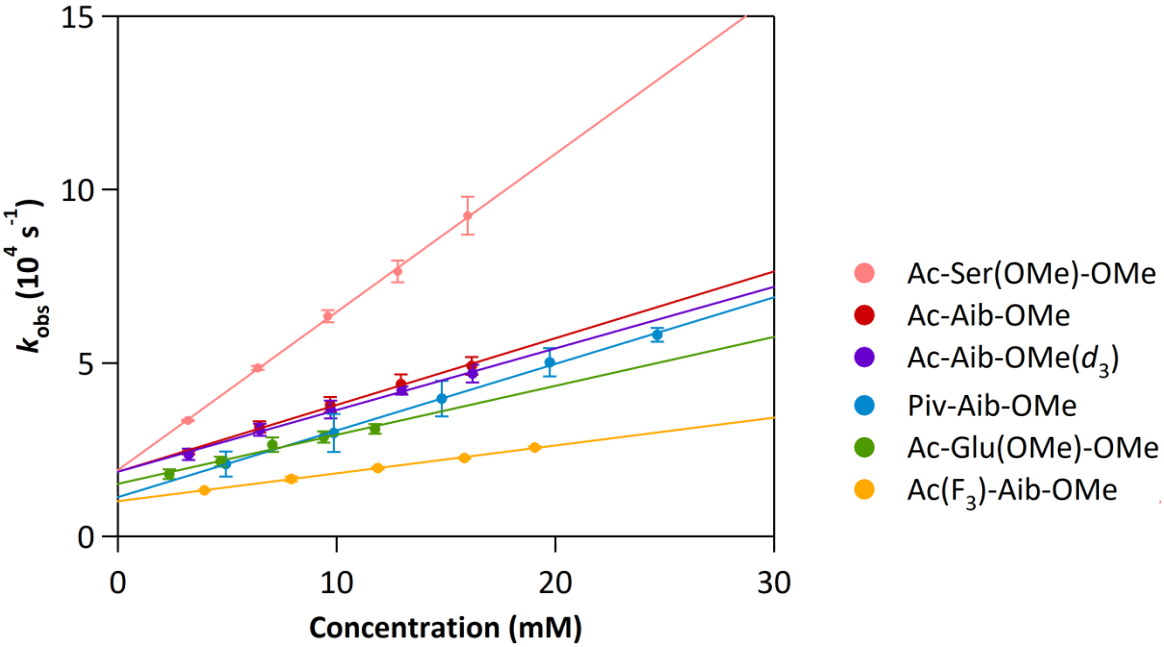
Typical kinetic plots:



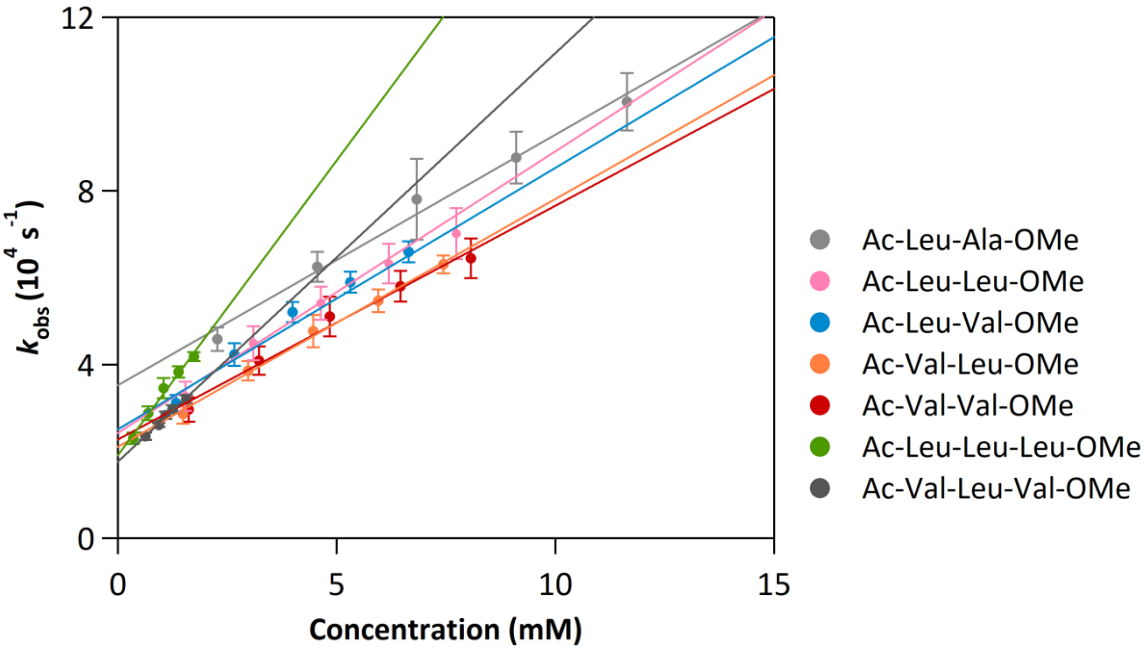
Plot of pseudo-first-order rate coefficient (k_{obs}) versus [amino acid] with aliphatic side chains (see Table 1). Error bars shown are 2σ statistical uncertainties. From the linear regression analysis: Ac-Gly-OMe: intercept = 3.0 x 10^4 s^{-1}, k_{HAT} = 3.3 x 10^6 M^{-1} s^{-1}, R^2 = 0.9892; Ac-Ile-OMe: intercept = 4.1 x 10^4 s^{-1}, k_{HAT} = 2.4 x 10^6 M^{-1} s^{-1}, R^2 = 0.9988; Ac-Leu-OMe: intercept = 2.2 x 10^4 s^{-1}, k_{HAT} = 2.0 x 10^6 M^{-1} s^{-1}, R^2 = 0.9905; Ac-Ala-OMe: intercept = 2.5 x 10^4 s^{-1}, k_{HAT} = 1.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9911; Ac-Val-OMe: intercept = 1.5 x 10^4 s^{-1}, k_{HAT} = 2.3 x 10^6 M^{-1} s^{-1}, R^2 = 0.9938.



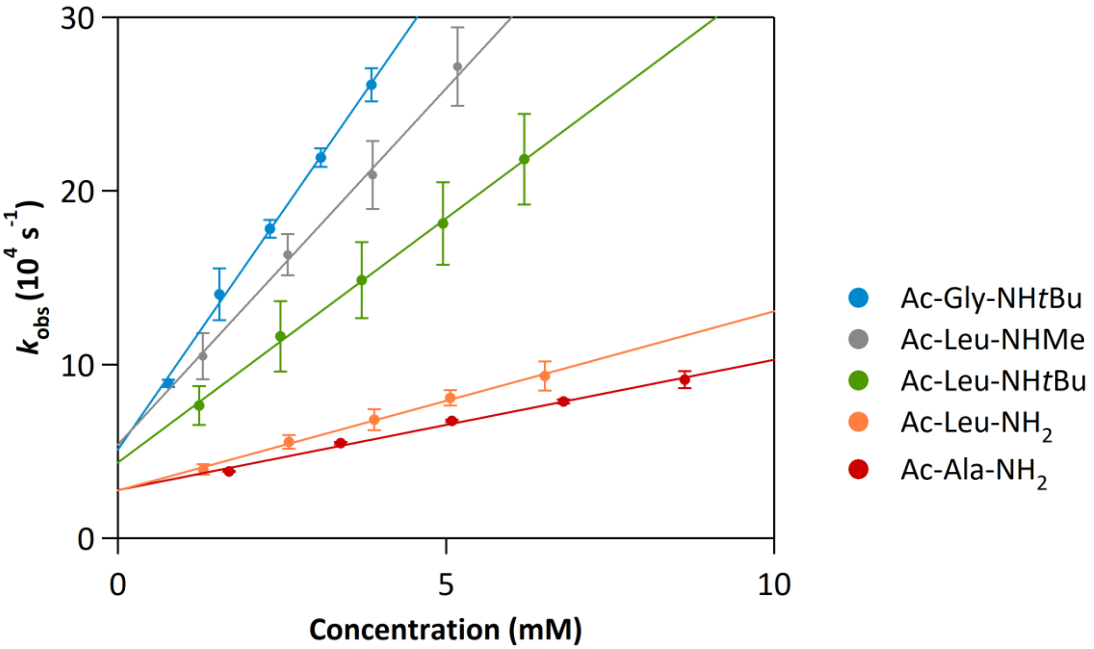
Plot of pseudo-first-order rate coefficient (k_{obs}) versus [amino acid] with phthaloyl protecting group at the N-terminus (see Table 1). Error bars shown are 2σ statistical uncertainties. From the linear regression analysis: Phth-Leu-OMe: intercept = 1.6 x 10^4 s^{-1}, k_{HAT} = 1.2 x 10^6 M^{-1} s^{-1}, R^2 = 0.9900; Phth-Val-OMe: intercept = 1.5 x 10^4 s^{-1}, k_{HAT} = 0.8 x 10^6 M^{-1} s^{-1}, R^2 = 0.9999; Phth-Aib-OMe: intercept = 1.1 x 10^4 s^{-1}, k_{HAT} = 0.9 x 10^6 M^{-1} s^{-1}, R^2 = 0.9971; Ac-Aib-OMe(d_3): intercept = 1.5 x 10^4 s^{-1}, k_{HAT} = 0.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9970; Phth-Gly-OMe: intercept = 1.1 x 10^4 s^{-1}, k_{HAT} = 0.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9919.



Plot of pseudo-first-order rate coefficient (k_{obs}) versus [amino acid] with aliphatic side chains (see Table 1). Error bars shown are 2σ statistical uncertainties. From the linear regression analysis: Ac-Ser(OMe)-OMe: intercept = 1.9 x 10^4 s^{-1}, k_{HAT} = 4.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9991; Ac-Aib-OMe: intercept = 1.9 x 10^4 s^{-1}, k_{HAT} = 1.9 x 10^6 M^{-1} s^{-1}, R^2 = 0.9957; Ac-Aib-OMe(d_3): intercept = 1.9 x 10^4 s^{-1}, k_{HAT} = 1.8 x 10^6 M^{-1} s^{-1}, R^2 = 0.9950; Piv-Aib-OMe: intercept = 1.1 x 10^4 s^{-1}, k_{HAT} = 1.9 x 10^6 M^{-1} s^{-1}, R^2 = 0.9983; Ac-Glu(OMe)-OMe: intercept = 1.5 x 10^4 s^{-1}, k_{HAT} = 1.4 x 10^6 M^{-1} s^{-1}, R^2 = 0.9768; Ac(F_3)-Aib-OMe: intercept = 1.0 x 10^4 s^{-1}, k_{HAT} = 0.8 x 10^6 M^{-1} s^{-1}, R^2 = 0.9995.



Plot of pseudo-first-order rate coefficient (k_{obs}) versus [peptide] with aliphatic side chains (see Table 2). Error bars shown are 2σ statistical uncertainties. From the linear regression analysis: Ac-Leu-Ala-OMe: intercept = 3.5 x 10^4 s^{-1}, k_{HAT} = 5.8 x 10^6 M^{-1} s^{-1}, R^2 = 0.9875; Ac-Leu-Leu-OMe: intercept = 2.5 x 10^4 s^{-1}, k_{HAT} = 6.0 x 10^6 M^{-1} s^{-1}, R^2 = 0.9902; Ac-Leu-Val-OMe: intercept = 2.4 x 10^4 s^{-1}, k_{HAT} = 6.5 x 10^6 M^{-1} s^{-1}, R^2 = 0.9872; Ac-Val-Leu-OMe: intercept = 2.1 x 10^4 s^{-1}, k_{HAT} = 5.7 x 10^6 M^{-1} s^{-1}, R^2 = 0.9961; Ac-Val-Val-OMe: intercept = 2.3 x 10^4 s^{-1}, k_{HAT} = 5.4 x 10^6 M^{-1} s^{-1}, R^2 = 0.9842; Ac-Leu-Leu-Leu-OMe: intercept = 1.9 x 10^4 s^{-1}, k_{HAT} = 13.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9847; Ac-Val-Leu-Val-OMe: intercept = 1.8 x 10^4 s^{-1}, k_{HAT} = 9.4 x 10^6 M^{-1} s^{-1}, R^2 = 0.9880.



Plot of pseudo-first-order rate coefficient (k_{obs}) versus [aliphatic amino acid] with amide groups at the C-terminus (see Table 3). Error bars shown are 2σ statistical uncertainties. From the linear regression analysis: Ac-Gly-NH*t*Bu: intercept = 5.1 x 10^4 s^{-1}, k = 54.6 x 10^6 M^{-1} s^{-1}, R^2 = 0.9976; Ac-Leu-NHMe: intercept = 5.4 x 10^4 s^{-1}, k_{HAT} = 41.0 x 10^6 M^{-1} s^{-1}, R^2 = 0.9976; Ac-Leu-NH*t*Bu: intercept = 4.4 x 10^4 s^{-1}, k = 28.1 x 10^6 M^{-1} s^{-1}, R^2 = 0.9987; Ac-Leu-NH_2: intercept = 2.8 x 10^4 s^{-1}, k = 10.3 x 10^6 M^{-1} s^{-1}, R^2 = 0.9969; Ac-Ala-NH_2: intercept = 2.8 x 10^4 s^{-1}, k = 7.5 x 10^6 M^{-1} s^{-1}, R^2 = 0.9942.

Gaussian Archive Entries for Computational Data

The calculations were performed using the Gaussian software package[64]

Data for Table 3.3-7 (page 70–71)

Reaction of $NO_3^\bullet$ with Ac-Gly-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC193\FOpt\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\27-Aug
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) freq=noraman nosymm geom=che
ckpoint guess=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxcy
cle=500)\\geom&freq\\0,2\C,-1.8288588102,-1.1490304948,1.1028972765\O,
-1.4596285569,-0.0039918606,1.249178097\O,-4.086216397,-4.8565025989,2
.2681812044\O,-4.946119041,-5.398480749,4.1946912033\O,-2.814330334,-5
.1747827244,3.8073027172\N,-4.0248140563,-5.1533752241,3.4828002821\C,
-2.6662428636,-1.9090262446,2.1048346994\H,-2.1259429107,-2.813799582,
2.4073021776\H,-3.5935769119,-2.236909909,1.6214859086\H,-2.550993825,
-0.125338789,3.2282317819\C,-3.7091616962,-1.4895477778,4.259610286\C,
-3.9392705483,-0.517100868,5.3907967423\H,-5.0145558797,-0.3513887619,
5.4960919351\H,-3.5768170366,-0.968218659,6.3180896285\H,-3.4407595369
,0.4419076865,5.2390725774\N,-2.9425542586,-1.0600686133,3.2372839908\
O,-4.1957399992,-2.6239436826,4.2657415511\C,-0.756730238,-1.276992651
6,-0.9803409587\H,0.2087253874,-0.9750087415,-0.5707238832\H,-0.628452
6675,-2.0352531005,-1.7492906417\H,-1.2807459176,-0.4066430335,-1.3793
356543\O,-1.5500379022,-1.89895462,0.0430210788\\Version=AM64L-G09RevB
.01\HF=-756.3366682\S2=0.761062\S2-1=0.\S2A=0.750072\RMSD=0.000e+00\RM
SF=6.677e-06\Dipole=0.985251,1.0421454,-0.9721997\Quadrupole=-4.083133
8,-2.5883077,6.6714415,-14.6898891,8.7578211,14.8768633\PG=C01 [X(C5H9
N2O6)]\\@

Sum of electronic and thermal Free Energies= -756.222748 Hartrees

Entry 1: Reaction at $CH_3C(O)NH\text{-}CH_2\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC011\FTS\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\01-Aug-
2018\0\\#p M062X/6-31+G* nosymm freq=noraman scrf=(cpcm,solvent=aceton
itrile) opt=(ts,noeigentest,z-matrix,calcfc,maxcycle=500)\\geom&freq\\
0,2\C,-0.6537183421,-1.9694419308,-1.44204113\H,-0.1733120988,-1.71496
67386,-2.3895943038\H,-1.1048119945,-2.9601887327,-1.5502985508\C,0.39
91961602,-2.030821514,-0.3465408209\H,-2.2204567294,-1.1137260609,-0.3
166786919\C,-1.6579271365,0.2416180157,-1.7330412262\O,0.3766010181,-1
.3732468078,0.6694843878\O,1.343792126,-2.9172114042,-0.6476461378\C,2
.3943733575,-3.0737288925,0.3193623927\H,1.9792416108,-3.4269821344,1.
2648679636\H,3.0725357845,-3.8117529227,-0.1031319843\H,2.9051325589,-
2.1211902618,0.470211103\O,-0.9633460377,0.5399466335,-2.6983739741\N,
-1.6534311928,-0.9775724745,-1.14558823\N,-0.4161686733,1.279841157,1.
2702093062\O,-1.7493899485,1.1044960007,1.3202930437\O,0.0960978132,1.
5706908953,0.2192242333\O,0.1118185313,1.1220421901,2.3442756777\C,-2.
6049355268,1.2397572328,-1.1200893082\H,-2.4653557736,2.2471126379,-1.
5102086758\H,-3.6499279942,0.9243222206,-1.074258695\H,-2.2594945125,1
.3011988914,0.033380621\\Version=AM64L-G09RevB.01\HF=-756.321511\S2=0.
760615\S2-1=0.\S2A=0.750071\RMSD=7.564e-09\RMSF=4.902e-06\Dipole=-1.43
62642,-1.654266,-0.1689977\Quadrupole=11.2344671,5.4035069,-16.637974,
-5.4013493,2.239327,0.7904462\PG=C01 [X(C5H9N2O6)]\\@

ν_{imag} = -1075.1581 cm^{-1}

Sum of electronic and thermal Free Energies= -756.206849 Hartrees

Product association complex [HNO_3 - $^{\bullet}CH_2C(O)NH$-CH_2-CO_2CH_3]

1\1\GINC-SPARTAN-RC004\FOpt\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\30-Jul
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent=ac
etonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\C,-
0.1102280343,-0.8328594014,-0.254927615\H,0.3406058257,-0.3918855907,0
.6407515267\H,0.1777168047,-0.2214773232,-1.1182917067\C,0.4384843867,
-2.2298103455,-0.4572041049\H,-1.9766146324,-1.792421022,-0.3710308188
\C,-2.276737942,0.0483832844,0.4774090995\O,-0.2515212702,-3.221908699
2,-0.5388842548\O,1.7600305803,-2.2153302382,-0.5476749744\C,2.3921972
195,-3.488763716,-0.7649634386\H,2.0402879836,-3.9199981252,-1.7034370
296\H,3.458073391,-3.2785289667,-0.8110507527\H,2.1645395812,-4.160459
7775,0.0641965387\O,-1.7686084958,1.1328064424,0.8674925583\N,-1.54383
39055,-0.9107880019,-0.1115199869\N,-0.1476796042,2.5187191622,-1.4955
821272\O,0.052833161,2.4762088687,-0.1796042188\O,-0.9914595209,1.7742
41794,-1.9640464627\O,0.5452523218,3.2945468337,-2.1110733396\C,-3.685
3336826,-0.2149058398,0.6676868565\H,-4.1386023455,-1.1417531749,0.340
3756955\H,-4.286142489,0.546151131,1.1453798058\H,-0.6838903332,1.8440
81706,0.2283447498\\Version=AM64L-G09RevB.01\HF=-756.3683042\S2=0.7567
79\S2-1=0.\S2A=0.750036\RMSD=0.000e+00\RMSF=9.448e-06\Dipole=-0.541392
9,-2.9118013,0.8502476\Quadrupole=15.0697377,-9.0788678,-5.9908698,-4.
8141702,-2.3225035,6.2735539\PG=C01 [X(C5H9N2O6)]\\@

Sum of electronic and thermal Free Energies= -756.251318 Hartrees

Entry 2: Reaction at $CH_3C(O)N$*H*-CH_2-CO_2CH_3

Charge transfer complex

1\1\GINC-SPARTAN-RC109\FOpt\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\11-Oct
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cp
cm,solvent=acetonitrile) opt=(z-matrix,calcfc,maxcycle=500)\\geom&freq
\\0,2\C,-0.7032932805,-1.8773694785,0.7886970616\H,-2.6829344944,-2.32
22168103,1.4103110356\C,-2.0634081665,-0.5587677944,2.3914195012\O,-1.
1641201715,0.2503687479,2.4530944866\N,-1.8809335769,-1.7163766182,1.6
052024367\O,-3.3484967659,-3.7622213547,2.8692765586\O,-1.9165522378,-
4.9688079261,3.9599959226\O,-1.3022066176,-3.0741191316,3.1741068164\N
,-2.2369283635,-3.973311812,3.3344200981\C,-3.3765831405,-0.4770343687
,3.107956386\H,-3.4465408226,-1.2934650611,3.8334175376\H,-3.441542578
1,0.4819419129,3.6202579446\H,-4.2068402032,-0.5828424279,2.4040598281
\H,0.1896402971,-1.7002621776,1.3940369733\H,-0.7021298071,-1.14420783
44,-0.0285672713\C,-0.6957251421,-3.2789762019,0.2154284808\O,0.436501
7367,-3.5245627437,-0.4225799337\O,-1.6200208857,-4.0530603525,0.32491
5931\C,0.5478996532,-4.8226612164,-1.0336033433\H,1.5338404815,-4.8423
487838,-1.4918230502\H,0.4583870781,-5.5976094974,-0.2709773854\H,-0.2
324396133,-4.9444357815,-1.7861983486\\Version=AM64L-G09RevB.01\HF=-75
6.3482188\S2=0.759823\S2-1=0.\S2A=0.750077\RMSD=0.000e+00\RMSF=6.033e-
06\Dipole=0.2541649,1.3583288,-2.3015519\Quadrupole=12.2708505,-10.200
7877,-2.0700629,-17.1835087,1.789003,23.2658119\PG=C01 [X(C5H9N2O6)]\\@

Sum of electronic and thermal Free Energies= - 756.228535 Hartrees

Transition state

1\1\GINC-SPARTAN-RC192\FTS\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\09-Oct-
2018\0\\#p M062X/6-31+G* scf=verytight scrf=(cpcm,solvent=acetonitrile
) nosymm opt=(ts,noeigentest,calcfc,maxcycle=500) freq=noraman\\geom&f
req\\0,2\C,-0.8736475014,-1.8528507669,0.54548199\H,-2.6021324187,-2.5
717692779,1.8446963552\C,-1.9412061972,-0.5015385091,2.321768976\O,-1.
0947396712,0.340521977,2.1424485885\N,-1.8627626585,-1.7249282795,1.56

3082721\O,-3.4374088804,-3.5220258474,2.3927197611\O,-3.0648962193,-5.
1217843843,3.7995960567\O,-1.7265805746,-3.423135698,3.7160530804\N,-2
.7266016105,-4.0450623616,3.3283252175\C,-3.0949366866,-0.4358742894,3
.2667078699\H,-2.983788256,-1.223366669,4.0191372019\H,-3.1102738164,0
.5429419989,3.7438487346\H,-4.032100238,-0.6128929086,2.7307498511\H,0
.0229216552,-1.2797642113,0.8035717974\H,-1.2779420487,-1.3893277329,-
0.3725945775\C,-0.5574601137,-3.3124176242,0.2640282016\O,0.3623913405
,-3.3969746835,-0.6837901816\O,-1.0769207474,-4.2432417491,0.828657332
1\C,0.7636337319,-4.7291222441,-1.053617221\H,1.5129147717,-4.59932739
11,-1.8305914609\H,1.1864209022,-5.2396295938,-0.1870240209\H,-0.09839
27628,-5.2807147541,-1.431976273\\Version=AM64L-G09RevB.01\HF=-756.338
8322\S2=0.762621\S2-1=0.\S2A=0.750083\RMSD=4.533e-09\RMSF=9.243e-06\Di
pole=1.1267123,2.6575013,-2.9334877\Quadrupole=13.5712759,-15.222755,1
.6514791,-27.6698253,8.5486942,34.8223132\PG=C01 [X(C5H9N2O6)]\\@

ν_{imag} = -364.5727 cm^{-1}

Sum of electronic and thermal Free Energies= -756.225007 Hartrees

Product association complex [HNO_3 - $CH_3C(O)N^{\bullet}$-CH_2-CO_2CH_3]

1\1\GINC-SPARTAN-RC208\FTS\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\02-Aug-
2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cpc
m,solvent=acetonitrile) opt=(ts,noeigentest,calcfc,maxcycle=500)\\geom
&freq\\0,2\C,-0.1103509647,-1.1743752149,1.7016113795\H,0.5818301004,-
1.1991601765,2.5564824464\H,0.0822996118,-0.22184819,1.1825148409\C,0.
1969532749,-2.3119052316,0.7535397643\H,-2.221966458,-3.2745370172,0.8
528158763\C,-1.7802142947,-0.4648941343,3.2681908297\O,-0.6045311634,-
3.1591509334,0.4018280416\O,1.448387054,-2.2771719224,0.3431002455\C,1
.8562436773,-3.3126787153,-0.570998153\H,1.2496478509,-3.2639382307,-1
.4761698082\H,2.900496856,-3.1057878347,-0.7916909372\H,1.7441880728,-
4.2877536081,-0.0952642094\O,-1.064599125,-0.4841691379,4.2545756647\N
,-1.4673703649,-1.2402558319,2.141708182\O,-3.1890381112,-3.3431457529
,1.1217483987\O,-4.8801829707,-4.593855862,0.6606471539\O,-2.985787377
8,-4.9852207508,-0.3161220659\N,-3.7133872037,-4.3780121723,0.43990774
21\C,-3.0863304954,0.2665454568,3.1664838331\H,-3.0924065614,0.8988073
341,2.2732479064\H,-3.8968399428,-0.4611895061,3.0604786967\H,-3.24034
24244,0.8733506021,4.0588303621\\Version=AM64L-G09RevB.01\HF=-756.3447
476\S2=0.756759\S2-1=0.\S2A=0.750032\RMSD=0.000e+00\RMSF=3.876e-04\Dip
ole=2.3774108,1.9952515,-0.739002\Quadrupole=1.3970411,-4.9826535,3.58
56124,-27.9760334,-1.9543929,3.2549006\PG=C01 [X(C5H9N2O6)]\\@

Sum of electronic and thermal Free Energies= - 756.230464 Hartrees

Entry 3: Reaction at $CH_3C(O)NH$-C*H_2*-CO_2CH_3

Transition state

1\1\GINC-R123\FTS\UM062X\6-31+G(d)\C5H9N2O6(2)\ROOT\03-May-2016\0\\# M
062X/6-31+G* scf=(qc,direct) scrf=(cpcm,solvent=acetonitrile) nosymm f
req=noraman opt=(ts,noeigentest,z-matrix,calcfc,maxcycle=500)\\freq&ge
om\\0,2\C,-0.1345342225,-0.7787498209,0.1362164564\H,0.1168526058,-0.0
424227687,0.9076160971\H,0.3192357453,-0.3373884612,-0.804763578\C,0.5
232191563,-2.119867098,0.3743902428\H,-1.8883053426,-1.8690172808,-0.1
786816121\C,-2.3642251199,0.1555871117,-0.0836941686\C,-3.8216204989,-
0.144820876,-0.3137120905\H,-4.144715907,-1.027139088,0.2436056915\H,-
4.4132277286,0.7205551272,-0.0159374057\H,-3.9837016707,-0.3336645446,
-1.3798457869\O,-0.063941708,-3.1758764014,0.3025933648\O,1.811368521,
-1.9784857559,0.6449908809\C,2.5482285973,-3.1947713831,0.8599741366\H
,2.5045571608,-3.816272387,-0.0358116511\H,3.5693684446,-2.8820705911,
1.0646152234\H,2.1298666798,-3.7340410699,1.7112940747\O,-1.9093872029
,1.2891438877,0.0233840339\N,-1.5312899295,-0.9299427611,-0.0222213159

\N,0.2447410553,2.046960175,-1.6163251358\O,0.8870431478,2.2013916442,
-0.6082902715\O,0.0704864245,0.8301696306,-2.0763239686\O,-0.273283208
4,2.9026267111,-2.3059472132\\Version=ES64L-G09RevD.01\HF=-756.3313472
\S2=0.759882\S2-1=0.\S2A=0.75007\RMSD=0.000e+00\RMSF=7.227e-06\Dipole=
-0.1336758,-2.7990085,0.6682532\Quadrupole=12.8468514,-10.1525425,-2.6
943089,-1.5211665,1.9160131,2.2911078\PG=C01 [X(C5H9N2O6)]\\@

ν_{imag} = -683.4420 cm^{-1}

Sum of electronic and thermal Free Energies= -756.217472 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH$-C•H-CO_2CH_3]

1\1\GINC-SPARTAN-RC147\FOpt\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\27-Aug
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm geom=checkpoint guess
=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxcycle=500) freq
=noraman\\geom&freq\\0,2\C,-0.2392863479,-0.9998258355,0.0719696609\H,
0.3182480421,-0.1824385128,0.5054866605\C,0.3830377462,-2.1639592109,-
0.5151775281\H,-2.0632662085,-1.8353703469,-0.2180166839\C,-2.35807887
06,0.053680328,0.5629682175\O,-0.252352587,-3.1069739922,-0.9703587711
\O,1.7189331162,-2.0905683999,-0.5022270395\C,2.4073612254,-3.21294239
87,-1.067626214\H,2.1340360563,-3.330515509,-2.1182046498\H,3.46651165
18,-2.9837286473,-0.9707159136\H,2.1583950782,-4.1220920103,-0.5168011
324\O,-1.8340323822,1.1201661884,0.9084846899\N,-1.600795018,-0.992623
8344,0.1189099239\N,0.0069127155,2.0993634663,-1.6028144286\O,0.063823
6049,2.423132948,-0.2998838282\O,-0.7963424812,1.2511688956,-1.9352151
23\O,0.7791217876,2.6922733402,-2.3137264417\H,-0.6585166712,1.8798462
651,0.164418388\C,-3.8423902856,-0.1490486492,0.6190141461\H,-4.145390
8808,-1.1664507959,0.3700759998\H,-4.1867100352,0.0983907422,1.6262476
79\H,-4.3123365661,0.5514079693,-0.0772696816\\Version=AM64L-G09RevB.0
1\HF=-756.3872713\S2=0.757203\S2-1=0.\S2A=0.75004\RMSD=0.000e+00\RMSF=
6.835e-06\Dipole=-1.1406397,-1.738337,1.2262514\Quadrupole=15.3141051,
-8.2187098,-7.0953954,-4.3761613,-2.0939795,2.2504445\PG=C01 [X(C5H9N2
O6)]\\@

Sum of electronic and thermal Free Energies= - 756.270750 Hartrees

Entry 4: Reaction at $CH_3C(O)NH$-CH_2-$CO_2C\mathit{H}_3$

Transition state

1\1\GINC-SPARTAN-RC139\FTS\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\29-May-
2018\0\\#p M062X/6-31+G* nosymm scrf=(cpcm,solvent=acetonitrile) freq=
noraman geom=checkpoint guess=check opt=(ts,noeigentest,readfc,maxcycl
e=500)\\geom&freq\\0,2\C,1.8038917842,-1.0554548759,2.4965435793\H,1.4
421142645,-0.0452282873,2.2692572941\H,0.922288263,-1.7044940958,2.557
4104464\C,2.6750225341,-1.523172453,1.3616848714\H,3.5289670363,-1.380
2359406,3.6746406341\C,2.0009083672,-0.703443395,4.8914739466\O,3.8318
915608,-1.8489883889,1.4506974798\O,1.9894476154,-1.5272767824,0.19858
92172\O,0.8270332387,-0.3338788497,4.9476477132\N,2.5623714027,-1.0803
735261,3.7209104806\N,5.1427994056,-0.7835605488,-2.7618678104\O,4.756
3593199,-1.8476047028,-3.1750764498\O,4.5704963354,-0.2656442305,-1.67
21923435\O,6.0177396861,-0.0765673416,-3.2073404132\C,2.8879433761,-0.
7527502394,6.1120501067\H,3.0053097794,0.2631029519,6.4988707331\H,2.3
919088642,-1.3495424642,6.8812600953\H,3.8743200327,-1.1736683125,5.90
87976741\C,2.6698860989,-1.9977503446,-0.9252169628\H,3.4730342841,-1.
2139142124,-1.2305949454\H,3.2317403986,-2.9144852068,-0.7336191559\H,
1.965869352,-2.0584607536,-1.7527151909\\Version=AM64L-G09RevB.01\HF=-
756.321419\S2=0.759139\S2-1=0.\S2A=0.750053\RMSD=9.786e-09\RMSF=2.138e
-06\Dipole=-0.1419148,-0.9522529,0.3090957\Quadrupole=-8.4413993,10.06
67047,-1.6253054,-2.5158168,23.1869195,-3.7718955\PG=C01 [X(C5H9N2O6)]\\@

ν_{imag} = -684.1731 cm^{-1}

Sum of electronic and thermal Free Energies= -756.210673 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH$-CH_2-$CO_2C^{\bullet}H_2$]

1\1\GINC-SPARTAN-RC065\FOpt\UM062X\6-31+G(d)\C5H9N2O6(2)\UWILLE\28-Jul
-2018\0\\#p M062X/6-31+G* nosymm freq=noraman scrf=(cpcm,solvent=aceto
nitrile) opt=(z-matrix,calcfc,maxcycle=500)\\geom&freq\\0,2\C,1.009707
0856,-1.5672185636,2.2932556\H,0.5977230669,-0.9360799689,1.4963195862
\H,0.1682678504,-2.1123326439,2.7376305316\C,1.9652771488,-2.553393730
9,1.6775345969\H,2.6818531953,-0.9903308904,3.4512392784\C,1.077362103
8,0.2248270131,3.9294080854\O,3.1372311044,-2.6593598885,1.949346819\O
,1.3414950817,-3.3104367023,0.7524780737\O,-0.1068903238,0.4855531631,
3.7108992288\N,1.7067769157,-0.7774910253,3.2761082385\N,4.5970402513,
-1.8493460684,-0.3342873291\O,5.0726879599,-2.9467063,-0.5024206577\O,
3.3595015011,-1.6234934865,-0.8614971462\O,5.0666841955,-0.9076860928,
0.2423250772\C,2.0767509053,-4.2553413743,0.0838213721\H,3.0931496491,
-2.4801353661,-1.2651081278\H,2.9884499949,-4.6020076743,0.5514624617\
H,1.4715098871,-4.8574415316,-0.5764754612\C,1.9026401912,1.0067625007
,4.9223255513\H,2.8961840125,0.5834612904,5.079636042\H,2.0053775252,2
.0325223503,4.5574839857\H,1.3673207643,1.0399361543,5.8739877958\\Ver
sion=AM64L-G09RevB.01\HF=-756.3501197\S2=0.754277\S2-1=0.\S2A=0.750015
\RMSD=4.960e-09\RMSF=6.012e-06\Dipole=-0.0827367,-1.5231737,-0.6920317
\Quadrupole=-17.3065512,11.669897,5.6366542,3.0658047,5.2901973,2.0172
271\PG=C01 [X(C5H9N2O6)]\\@

Sum of electronic and thermal Free Energies= -756.234045 Hartrees

Reaction of $NO_3^{\bullet}$ with Mal-Gly-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC163\FOpt\UM062X\6-31+G(d)\C7H7N2O7(2)\UWILLE\26-Aug
-2018\0\\#p M062X/6-31+G* scf=verytight nosymm freq=noraman geom=check
point guess=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxcycl
e=500)\\geom&freq\\0,2\C,0.3669269123,-0.773930344,-0.4017653218\H,1.0
072146581,-0.2246783288,0.2929172998\H,0.6335860042,-0.4763732397,-1.4
191187149\C,0.5884639831,-2.2708631508,-0.2373270204\O,-0.2905959498,-
3.0652717539,0.0059880921\O,1.868369944,-2.5796568537,-0.4039401394\C,
2.2021141203,-3.9725702223,-0.2836944818\H,1.6719115609,-4.5467854861,
-1.0452893229\H,3.2772752433,-4.0284173803,-0.4377178675\H,1.934257105
6,-4.3333276102,0.7107394588\N,-0.3911154237,2.859919293,-0.7286716478
\O,-1.4336279396,2.185627607,-0.2630100362\O,0.5790632129,2.2210497312
,-1.0288215298\O,-0.5854068967,4.0482235952,-0.7803280315\N,-1.0091532
671,-0.4377463847,-0.1577209042\C,-1.5701074345,-0.3934229367,1.124170
1697\C,-2.0224720279,-0.5873313342,-1.1115227621\C,-3.0560117071,-0.34
77646308,0.9323866483\C,-3.3195840574,-0.4609358387,-0.3704688385\H,-3
.7414935861,-0.2465868629,1.763630052\H,-4.2752926278,-0.475554964,-0.
8780703918\O,-0.9348097902,-0.3972079772,2.1521031228\O,-1.8353504168,
-0.7804731875,-2.2895927531\\Version=AM64L-G09RevB.01\HF=-906.5124109\
S2=0.756539\S2-1=0.\S2A=0.750034\RMSD=9.988e-09\RMSF=5.213e-06\Dipole=
-0.1932579,-0.2826524,0.1424972\Quadrupole=13.4768486,-3.208814,-10.26
80346,-8.7097548,-3.309921,1.3043028\PG=C01 [X(C7H7N2O7)]\\@

Sum of electronic and thermal Free Energies= -906.405621 Hartrees

Entry 5: Reaction at $(CHC(O))_2N$-C*H*$_2$-CO_2CH_3

Transition state

1\1\GINC-SPARTAN-RC099\FTS\UM062X\6-31+G(d)\C7H7N2O7(2)\UWILLE\19-Sep-
2018\0\\#p M062X/6-31+G* scf=(qc,direct) scrf=(cpcm,solvent=acetonitri

le) nosymm freq=noraman opt=(ts,noeigentest,calcfc,maxcycle=500)\\geom
&freq\\0,2\C,0.2484436308,-1.2607752492,0.7214496905\H,0.4771293496,-1
.0487659306,1.7705620881\H,0.6346821986,-0.3595792841,0.1135064124\C,1
.0250618798,-2.4702352378,0.2159677677\O,0.5552725331,-3.574730098,0.0
987405318\O,2.2822012966,-2.1330810188,-0.0301080153\C,3.151659981,-3.
2012932915,-0.4461613274\H,3.1925471699,-3.9673781828,0.329823254\H,2.
7853611173,-3.6306318555,-1.3798738646\H,4.1268057866,-2.7416743404,-0
.5876178669\N,-0.1125156272,1.5201035654,-1.1662710208\O,-1.1359035642
,1.2359401063,-0.5807766954\O,0.9546125297,0.7786550763,-0.9628047828\
O,0.0561486007,2.4306187589,-1.9444264813\N,-1.1377523677,-1.362835899
1,0.5394024196\C,-2.1085203534,-0.9710094549,1.4846664375\C,-1.7400568
49,-1.5246321227,-0.7228777781\C,-3.4209808894,-1.0233331445,0.7696591
612\C,-3.2086104981,-1.3491647748,-0.5074885496\H,-4.3551489372,-0.812
2497036,1.273280724\H,-3.9256980218,-1.4656623037,-1.3093307434\O,-1.8
650648841,-0.6658365774,2.6241795873\O,-1.1232140818,-1.7184060374,-1.
7406139486\\Version=AM64L-G09RevB.01\HF=-906.5031355\S2=0.759119\S2-1=
0.\S2A=0.750055\RMSD=0.000e+00\RMSF=8.678e-06\Dipole=-0.3159437,-0.966
0713,0.6447531\Quadrupole=19.9115446,-9.0476647,-10.8638799,-2.9292233
,1.8099098,1.8175583\PG=C01 [X(C7H7N2O7)]\\@

ν_{imag} = -449.8802 cm^{-1}

Sum of electronic and thermal Free Energies= - 906.396300 Hartrees

Product association complex [HNO_3 – $(CHC(O))_2N$-C•H-CO_2CH_3]

1\1\GINC-SPARTAN-RC170\FOpt\UM062X\6-31+G(d)\C7H7N2O7(2)\UWILLE\23-Aug
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman opt=(cal
cfc,maxcycle=500)\\geom&freq\\0,2\C,-0.5409681193,-0.6718522197,1.2783
977179\H,-0.5168177282,-0.3179756305,2.300333136\H,0.4750505876,0.3231
941806,-1.5210272666\C,0.6498257671,-1.2889876042,0.7205743016\O,1.747
0595843,-1.1543615465,1.2244341258\O,0.4089335661,-2.0203336095,-0.381
7817721\C,1.5695708634,-2.6181530239,-0.9712716666\H,2.0317827138,-3.3
122083036,-0.2659472132\H,1.2078663651,-3.1454829631,-1.8525643321\H,2
.2882106173,-1.8426883059,-1.2462381359\N,1.5134447279,1.5193739758,-0
.4919283889\O,0.4847366102,1.7633022142,0.1083221821\O,1.4309060799,0.
5957938293,-1.4892290125\O,2.6015503444,1.9857253442,-0.3238852369\N,-
1.757181604,-0.5975401183,0.6230027427\C,-2.0170182352,-0.3266994131,-
0.7252363345\C,-2.9875153356,-0.6224351948,1.340101843\C,-3.5077251424
,-0.3006094627,-0.8750806799\C,-4.0712914992,-0.4524795796,0.321735786
4\H,-3.9681505535,-0.1410254205,-1.8414826057\H,-5.1174815352,-0.45916
38584,0.59831293\O,-3.0761338569,-0.7564626922,2.5284898941\O,-1.20919
17878,-0.1013182774,-1.5976087947\\Version=AM64L-G09RevB.01\HF=-906.52
91978\S2=0.758325\S2-1=0.\S2A=0.75005\RMSD=0.000e+00\RMSF=9.742e-05\Di
pole=-2.4342268,-1.0302265,-1.1046104\Quadrupole=6.8067467,-1.27356,-5
.5331867,-3.9068523,2.7497888,2.88193\PG=C01 [X(C7H7N2O7)]\\@

Sum of electronic and thermal Free Energies= - 906.434510 Hartrees

Reaction of NO_3• with Ac-Ala-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC171\FOpt\UM062X\6-31+G(d)\C6H11N2O6(2)\UWILLE\30-Au
g-2018\0\\#p M062X/6-31+G* scf=(qc,direct) freq=noraman nosymm geom=ch
eckpoint guess=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxc
ycle=500)\\geom&freq\\0,2\C,-1.9963790738,-0.8967422366,1.0411085597\O
,-1.6034102234,0.2336909913,1.2376277163\O,-3.8619006147,-4.7842986148
,2.3503367438\O,-4.3682479522,-5.3995072998,4.3779597072\O,-2.36577713
71,-4.8643111016,3.7074868831\N,-3.5960699295,-5.0367350124,3.54884605

36\C,-2.997723007,-1.6209750084,1.9212112035\H,-2.5611202316,-2.580950
7141,2.2190203471\H,-2.9229103837,0.1537115421,3.0743513064\C,-3.73019
54521,-1.3445677264,4.2367544222\C,-3.9163628084,-0.3999300421,5.39965
90233\H,-4.9871684555,-0.297867692,5.5969965654\H,-3.4519368188,-0.839
6302295,6.2854551701\H,-3.4919589727,0.5896379284,5.2207097984\N,-3.20
7959099,-0.8184990336,3.1090984974\O,-4.0487949691,-2.5349810462,4.311
5183566\C,-0.7102060616,-1.0396043131,-0.9170872573\H,0.2207283971,-0.
7873176746,-0.406312503\H,-0.5366063141,-1.7882772561,-1.6866677442\H,
-1.1548918084,-0.139970219,-1.3461233083\C,-4.3020301047,-1.8760846013
,1.1550658059\H,-5.0247126937,-2.3323215589,1.8351509898\H,-4.71127106
04,-0.9343587306,0.7765411286\H,-4.1277698876,-2.5555885301,0.31727465
23\O,-1.6265833377,-1.6458818206,0.0077978822\\Version=AM64L-G09RevB.0
1\HF=-795.6341026\S2=0.76113\S2-1=0.\S2A=0.750072\RMSD=0.000e+00\RMSF=
4.129e-06\Dipole=0.559718,1.0927336,-1.1492026\Quadrupole=0.9737926,-3
.8173009,2.8435083,-13.0656303,7.2520548,15.592044\PG=C01 [X(C6H11N2O6
)]\\@

Sum of electronic and thermal Free Energies= -795.492161 Hartrees

Entry 6: Reaction at $CH_3C(O)NH\text{-}C\mathit{H}(CH_3)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC121\FTS\UM062X\6-31+G(d)\C6H11N2O6(2)\UWILLE\05-Nov
-2017\0\\#p M062X/6-31+G* nosymm scrf=(cpcm,solvent=acetonitrile) opt=
(ts,noeigentest,calcfc,z-matrix,maxcycle=500) freq=noraman\\geom&freq\
\0,2\C,-0.2130788278,-0.642515992,0.4010362388\H,0.2226989032,-0.14829
12651,-0.5192272908\C,0.4466631245,-2.0107313099,0.491606749\H,-1.9352
128459,-1.7987240778,0.1448457095\C,-2.4411671466,0.1689284747,-0.2303
403479\O,-0.156243207,-3.0566948312,0.3873872578\O,1.7517273073,-1.907
8256515,0.6904326663\C,2.4765227069,-3.146265755,0.7818554646\H,2.3763
298889,-3.7021021755,-0.1519084114\H,3.5123243875,-2.8634998067,0.9542
560919\H,2.0928233957,-3.7389720904,1.613775272\O,-2.0276981294,1.3146
177517,-0.3789448961\N,-1.6101266252,-0.8358061594,0.1810621442\N,0.29
56077533,1.998719033,-1.7825624453\O,0.7245488809,2.3646610424,-0.7193
227244\O,0.2207589014,0.7093720508,-2.0286697893\O,-0.0761663757,2.684
0453147,-2.7142499152\C,0.1238581441,0.2657773674,1.5864055735\H,-0.39
21495263,1.2193304782,1.4721231219\H,-0.1994526372,-0.2166971129,2.514
6285204\H,1.1996776143,0.4431854213,1.6264129029\C,-3.8640645607,-0.24
04771865,-0.515447534\H,-3.9177295734,-0.6675843568,-1.5222517092\H,-4
.2211580501,-0.9898984388,0.19471481\H,-4.5021443733,0.6423452855,-0.4
75634821\\Version=AM64L-G09RevB.01\HF=-795.6285719\S2=0.75986\S2-1=0.\
S2A=0.750069\RMSD=5.570e-09\RMSF=9.019e-06\Dipole=0.0803726,-2.5564914
,1.2323777\Quadrupole=12.2094445,-9.207957,-3.0014875,-1.2115374,1.121
6783,3.6299458\PG=C01 [X(C6H11N2O6)]\\@

ν_{imag} = -437.9281 cm^{-1}

Sum of electronic and thermal Free Energies= -795.487140 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}C^{\bullet}(CH_3)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC007\FOpt\UM062X\6-31+G(d)\C6H11N2O6(2)\UWILLE\13-De
c-2018\0\\#p M062X/6-31+G* nosymm scf=(qc,direct) scrf=(cpcm,solvent=a
cetonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\C,
-0.2987765338,-1.4069148526,1.54859884\H,-0.6046959471,0.3667048236,-0
.9058489194\C,0.6826547317,-0.4376506971,1.9880664216\H,0.9997285893,-
1.7873391315,0.054545003\C,-0.585312074,-3.0648439416,-0.2816998042\C,
0.1420037696,-3.5499501735,-1.5117864073\H,1.1437823562,-3.9069062363,
-1.254947322\H,-0.4289178556,-4.3606267878,-1.9630627105\H,0.248586133
,-2.7346775774,-2.2348124703\O,1.736316247,-0.229761148,1.3823802221\O
,0.3347017196,0.2295585224,3.0977557722\C,1.2466333966,1.2530155843,3.

5090807593\H,1.3435788246,2.0044960863,2.7216186092\H,0.8093622947,1.6
901898949,4.4047577215\H,2.2250018624,0.8214161622,3.7303036694\O,-1.6
719358886,-3.4917144675,0.0758332908\N,0.0885676383,-2.0738404797,0.41
09344554\N,0.5855298903,1.6877739429,-0.3041994279\O,-0.2317821373,1.9
31829512,0.5539327534\O,0.2903299739,0.6717468369,-1.164098526\O,1.639
4951782,2.2200291748,-0.5075823288\C,-1.589932463,-1.6644155842,2.2433
827958\H,-1.679324782,-0.99291753,3.0966523879\H,-2.4371798385,-1.5088
938932,1.5681585337\H,-1.6489000855,-2.7009650399,2.5913726813\\Versio
n=AM64L-G09RevB.01\HF=-795.6761531\S2=0.756241\S2-1=0.\S2A=0.750031\RM
SD=0.000e+00\RMSF=3.811e-06\Dipole=-0.2921267,-0.4610935,-0.2980214\Qu
adrupole=-5.6538924,-3.2973083,8.9512007,-9.2984283,1.8488057,5.074124
8\PG=C01 [X(C6H11N2O6)]\\@

Sum of electronic and thermal Free Energies= -795.532333 Hartrees

Entry 7: Reaction at $CH_3C(O)NH\text{-}CH(C\mathit{H}_3)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC106\FTS\UM062X\6-31+G(d)\C6H11N2O6(2)\UWILLE\02-Nov
-2017\0\\#p M062X/6-31+G* nosymm scrf=(cpcm,solvent=acetonitrile) opt=
(ts,noeigentest,calcfc,z-matrix,maxcycle=500) freq=noraman\\geom&freq\
\0,2\C,-0.6154169186,-1.0681206851,-0.3694585295\H,-0.8464839283,-0.24
23279582,-1.0504295069\C,0.2894404742,-2.0574205704,-1.0853093206\H,-1
.8253394033,-2.7512629654,0.0392073966\C,-2.9863726394,-1.0439188273,0
.2251016793\C,-4.2096868406,-1.8626517105,0.5567776074\H,-4.6227003146
,-1.5036508389,1.5024707717\H,-4.9588966674,-1.698481451,-0.2223023431
\H,-4.0017836937,-2.9310473093,0.6349632732\O,0.025791448,-3.226767744
4,-1.2534408653\O,1.4132724761,-1.4694377253,-1.47270322\C,2.376897408
7,-2.3086198628,-2.130240297\H,1.950435685,-2.7098130992,-3.0510706377
\H,3.2241584108,-1.662105482,-2.3471420987\H,2.6711333964,-3.123999599
,-1.4674123736\O,-3.0145748754,0.183054226,0.1564538951\N,-1.840771202
8,-1.738121089,-0.004762817\N,2.6571704604,1.4122571281,1.3522119605\O
,1.975543837,1.3786301529,2.3463968796\O,2.2856626099,0.7049830074,0.2
769877264\O,3.6876601577,2.0232181023,1.1751059219\C,0.0543468963,-0.5
226978595,0.8885757454\H,-0.5387352101,0.252971554,1.3762854474\H,1.05
61082431,0.0294382186,0.5655144194\H,0.3820821914,-1.2973926119,1.5867
692825\\Version=AM64L-G09RevB.01\HF=-795.617357\S2=0.75936\S2-1=0.\S2A
=0.750053\RMSD=6.886e-09\RMSF=1.709e-05\Dipole=-0.4853553,-2.3168054,-
0.5700557\Quadrupole=-0.9633573,-0.8192012,1.7825585,-0.3141746,-7.841
3732,-4.3934006\PG=C01 [X(C6H11N2O6)]\\@

ν_{imag} = -635.7318 cm^{-1}

Sum of electronic and thermal Free Energies= -795.478862 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}CH(C^{\bullet}H_2)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC075\FOpt\UM062X\6-31+G(d)\C6H11N2O6(2)\UWILLE\05-No
v-2017\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent=a
cetonitrile) opt=(readfc,z-matrix,maxcycle=500) freq=noraman\\geom&fre
q\\0,2\C,-0.9474807708,-1.0195817078,-0.1823224027\H,-1.1836952625,-0.
2855581099,-0.9610337608\C,0.1167279007,-1.9615533583,-0.7319217571\H,
-2.0242111388,-2.8254596852,0.1230099139\C,-3.3333545711,-1.2317462176
,0.2957892893\C,-4.4933686898,-2.1579624331,0.5724622054\H,-4.92345345
61,-1.8969340366,1.5428914004\H,-5.2595299659,-1.9948053388,-0.1897980
225\H,-4.2095834393,-3.2118606812,0.5769747825\O,0.0380076884,-3.16489
21579,-0.743069856\O,1.180529916,-1.2861343059,-1.1934389999\C,2.28045
94985,-2.0861517735,-1.6773100279\H,1.9366563928,-2.7161642688,-2.4976
503847\H,3.0274257239,-1.3747868389,-2.0238831062\H,2.6711881483,-2.69
54148867,-0.861523544\O,-3.4684106069,-0.0084308635,0.2565933722\N,-2.
1310316058,-1.8180818703,0.0876541609\N,2.8616782031,0.7006545585,0.89

37994448\O,2.8534000095,-0.4801944509,1.1633254411\O,2.1667199701,1.09
23554962,-0.1998397277\O,3.4186669846,1.5903250072,1.481533423\C,-0.43
50804799,-0.3103652861,1.0385891885\H,-0.5404878615,0.762156227,1.1310
428768\H,1.7273111166,0.269604692,-0.5429534196\H,-0.1538467023,-0.897
850711,1.9048075219\\Version=AM64L-G09RevB.01\HF=-795.6480222\S2=0.755
256\S2-1=0.\S2A=0.750021\RMSD=0.000e+00\RMSF=5.265e-06\Dipole=0.010848
3,-1.8745279,-1.1929\Quadrupole=-3.7176833,0.1082137,3.6094696,3.56266
72,-9.0926052,-0.7337944\PG=C01 [X(C6H11N2O6)]\\@
Sum of electronic and thermal Free Energies= -795.505378 Hartrees

Reaction of $NO_3^\bullet$ with Ac-Aib-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC051\FOpt\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\29-Au
g-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm geom=checkpoint gues
s=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxcycle=500) fre
q=noraman\\geom&freq\\0,2\C,-1.8935661252,-0.774340641,0.9971452473\O,
-1.6571804618,0.3854886365,1.2670247328\O,-4.3191256025,-5.2312472706,
2.9176110736\O,-5.4649957547,-5.0605147856,4.7624709796\O,-3.308751144
2,-5.3530334167,4.6645915084\N,-4.4460239956,-5.195537011,4.1628271927
\C,-2.6528818414,-1.7380649147,1.9100964917\H,-2.6796212272,0.00695297
07,3.0957636658\C,-3.6720215312,-1.4255455204,4.1662567486\C,-3.910898
978,-0.4334764205,5.2830047691\H,-4.9894736927,-0.3053657814,5.4088684
145\H,-3.5127762358,-0.8522192551,6.2106959249\H,-3.4537933327,0.54124
91621,5.1032156401\N,-2.9894034452,-0.9588274647,3.099135796\O,-4.0890
571981,-2.5841965099,4.2478974055\C,-0.80243962,-0.5350661209,-1.06577
55757\H,0.1230164542,-0.1786492281,-0.6100675007\H,-0.5894226945,-1.17
65450061,-1.9180228733\H,-1.4201580151,0.3133670595,-1.3649560019\C,-3
.9202645974,-2.2136288506,1.1829985965\H,-4.4930693068,-2.8582169803,1
.851988611\H,-4.5391821279,-1.3585983203,0.8932028512\H,-3.6494797841,
-2.7769762264,0.2869290408\C,-1.732266176,-2.9162399434,2.2639372786\H
,-0.8104611938,-2.5557114101,2.7313622274\H,-2.2507308277,-3.575632348
3,2.9620719277\H,-1.4780411964,-3.4783776347,1.3622770208\O,-1.5205613
483,-1.359254768,-0.1349021928\\Version=AM64L-G09RevB.01\HF=-834.92854
83\S2=0.76108\S2-1=0.\S2A=0.750072\RMSD=0.000e+00\RMSF=3.799e-06\Dipol
e=0.9299364,0.9704836,-1.2276291\Quadrupole=-2.5714246,-1.4203532,3.99
17778,-12.2758053,9.6965391,13.1711474\PG=C01 [X(C7H13N2O6)]\\@
Sum of electronic and thermal Free Energies= -834.758253 Hartrees

Entry 8: Reaction at $CH_3C(O)N\mathit{H}$-$C(CH_3)_2$-CO_2CH_3

Charge transfer complex

1\1\GINC-SPARTAN-RC029\FOpt\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\21-Oc
t-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent=a
cetonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\C,
-0.579838613,-1.8075673155,0.8067731113\H,-2.5440838362,-2.3099084049,
1.423020668\C,-2.0963451094,-0.5195219672,2.3997345654\O,-1.3173411714
,0.4065799205,2.4604374129\N,-1.7728493878,-1.6717994009,1.6475238907\
O,-3.3818137671,-3.7018273451,2.7943498182\O,-2.0774545516,-4.94709360
48,3.9949800086\O,-1.3390779897,-3.0824974384,3.2438344425\N,-2.313785
6725,-3.9446022189,3.3408663336\C,-3.4210922492,-0.5733595421,3.103333
8483\H,-3.3324410828,-1.2224884495,3.9803024958\H,-3.6840978896,0.4340
424583,3.4247019027\H,-4.2029281743,-0.9836754264,2.460497325\C,-0.673
0676271,-3.2276404421,0.2351543865\O,0.4445145663,-3.5861172603,-0.374
0069457\O,-1.6703477316,-3.9140437088,0.303182573\C,0.4364793187,-4.88
38050609,-0.9945540369\H,1.4227321516,-4.9981495631,-1.4378945608\H,0.

25818914,-5.6513076473,-0.2400010123\H,-0.3405362043,-4.9224646316,-1.
7596648291\C,-0.6679838536,-0.8114425725,-0.3694563737\H,-1.6034412387
,-0.9368635488,-0.9216763012\H,-0.6067237801,0.2085630249,0.0147334135
\H,0.1721164897,-0.9899858586,-1.0451219\C,0.7199464773,-1.6151649392,
1.5974011288\H,0.7645314467,-0.6031369659,1.9994438286\H,0.7910834497,
-2.3370984878,2.4132883662\H,1.5604310603,-1.765197453,0.9173885401\\V
ersion=AM64L-G09RevB.01\HF=-834.9396206\S2=0.760421\S2-1=0.\S2A=0.7500
87\RMSD=0.000e+00\RMSF=5.144e-06\Dipole=0.7972988,1.4013438,-2.424255\
Quadrupole=10.7127605,-9.0822692,-1.6304912,-18.3488231,3.9580913,23.3
843885\PG=C01 [X(C7H13N2O6)]\\@
Sum of electronic and thermal Free Energies= -834.764871 Hartrees

Transition state

1\1\GINC-SPARTAN-RC097\FTS\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\16-Sep
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent=ac
etonitrile) opt=(ts,noeigentest,calcfc,maxcycle=500) freq=noraman\\geo
m&freq\\0,2\C,-1.2401845055,-1.7505526923,0.1948879362\C,-2.3297884778
,-2.0560493281,-0.8417253504\H,-3.0218324711,-2.0944925498,1.592139776
8\C,-1.3599894798,-1.016104703,2.6224825653\O,-3.4230593041,-1.5371694
962,-0.7977479208\O,-1.909866856,-2.8654402756,-1.7980981485\C,-2.8743
368468,-3.1776972379,-2.8170382945\H,-3.1858501557,-2.2647518452,-3.32
71556156\H,-2.3644198554,-3.850419732,-3.5027261713\H,-3.7367870349,-3
.6669485004,-2.3602221223\O,-0.1726941728,-0.7941413688,2.6608349982\N
,-1.9458473901,-1.60388701,1.4502292388\O,-4.1976760717,-2.6588205054,
1.8580461142\O,-5.2785031752,-4.4290865335,1.2559378939\O,-3.224994035
7,-4.1575192599,0.6302684074\N,-4.2410586415,-3.7852268488,1.229438593
8\C,-2.3493825466,-0.7692514388,3.7173508572\H,-2.7673485308,-1.723873
871,4.0541687849\H,-1.8511183481,-0.2667139897,4.545071545\H,-3.177270
8476,-0.1567438695,3.346123068\C,-0.5491524095,-0.4442206461,-0.254844
6614\H,0.306131175,-0.2225752969,0.3823688755\H,-0.2014713965,-0.58733
59935,-1.2815643447\H,-1.2556663991,0.3893418948,-0.2338785188\C,-0.21
2726348,-2.8974076896,0.3515886565\H,0.5315932598,-2.6204263251,1.0984
876482\H,-0.7133329906,-3.8235191369,0.6382912334\H,0.2737538563,-3.03
2424751,-0.616698045\\Version=AM64L-G09RevB.01\HF=-834.9324524\S2=0.76
2461\S2-1=0.\S2A=0.750074\RMSD=0.000e+00\RMSF=8.005e-06\Dipole=2.82819
76,1.9563256,-0.3642161\Quadrupole=-23.4178466,-5.450719,28.8685656,-2
7.4407762,4.3300264,9.2045782\PG=C01 [X(C7H13N2O6)]\\@

ν_{imag} = -531.4544 cm^{-1}

Sum of electronic and thermal Free Energies= -834.765224 Hartrees

Product association complex [HNO_3 - $CH_3C(O)N^{\bullet}$-$C(CH_3)_2$-CO_2CH_3]

1\1\GINC-SPARTAN-RC086\FOpt\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\13-Au
g-2018\0\\#p M062X/6-31+G* nosymm geom=checkpoint guess=check scrf=(cp
cm,solvent=acetonitrile) opt=(readfc,maxcycle=500) freq=noraman\\geom&
freq\\0,2\C,-0.3206340209,-0.9824191726,0.0111380283\C,0.1185182318,-2
.4394596841,0.1900568192\H,-2.044352442,-3.1626434416,1.283499682\C,-2
.193880793,0.4157218458,0.7511928547\O,-0.4878067052,-3.2875841555,0.8
24316628\O,1.2505929546,-2.6925938973,-0.4342522698\C,1.7506365109,-4.
0403047649,-0.3511269642\H,1.0152286771,-4.7316961991,-0.7643291162\H,
2.6632087783,-4.0442729865,-0.9418417571\H,1.9591997109,-4.2904692089,
0.6899011646\O,-1.77178581,1.4104661495,1.3119917405\N,-1.5061134605,-
0.802037444,0.8230936866\O,-3.0135320171,-3.2767954624,1.5522146461\O,
-4.9343945813,-3.0073930772,0.6192252915\O,-3.1412593359,-2.801883625,
-0.5798842349\N,-3.7379103059,-3.0135504709,0.4575865374\C,-3.52725274
75,0.3389232352,0.0553414337\H,-3.4339787121,-0.1491896909,-0.91925133
98\H,-4.2139539393,-0.2603775126,0.663731598\H,-3.9313652917,1.3455601

473,-0.0581039888\C,0.8047494675,-0.0517963465,0.517194997\H,0.5101560
687,0.9889927536,0.3740703072\H,1.0035222358,-0.2221253111,1.578516244
8\H,1.7107278596,-0.2546295116,-0.0576834149\C,-0.6001234112,-0.722876
8076,-1.4824975157\H,0.3044569048,-0.921059051,-2.06153074\H,-1.405835
565,-1.3655823437,-1.8478437191\H,-0.8803793413,0.325563514,-1.6254724
492\\Version=AM64L-G09RevB.01\HF=-834.9404962\S2=0.754429\S2-1=0.\S2A=
0.750017\RMSD=6.870e-09\RMSF=1.074e-05\Dipole=2.5938852,-0.3575425,-1.
5766608\Quadrupole=-5.6604127,3.7932061,1.8672066,-17.2725238,5.127055
6,1.7093864\PG=C01 [X(C7H13N2O6)]\\@
Sum of electronic and thermal Free Energies= -834.768421 Hartrees

Entry 9: Reaction at $CH_3C(O)NH\text{-}C(C\mathit{H}_3)_2\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC110\FTS\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\11-Aug
-2018\0\\#p M062X/6-31+G* nosymm scrf=(cpcm,solvent=acetonitrile) geom
=checkpoint guess=check opt=(ts,noeigentest,readfc,maxcycle=500) freq=
noraman\\geom&freq\\0,2\C,-0.6399220375,-0.973072194,-0.4491912149\C,0
.2209512118,-1.9907653511,-1.2039147086\H,-1.9577438394,-2.617371626,-
0.4738596177\C,-2.9384260295,-1.0968444376,0.5079578307\C,-4.161003808
9,-1.9721911113,0.6606112999\H,-4.9989436274,-1.4844495392,0.155592696
3\H,-4.0297221683,-2.9745314163,0.2498171314\H,-4.4053608997,-2.045372
2803,1.7232756581\O,-0.1676258032,-3.0800102114,-1.565184307\O,1.44036
74557,-1.5191011472,-1.4186829796\C,2.3410688522,-2.3798814422,-2.1350
580597\H,1.9510828337,-2.5695046226,-3.1364470566\H,3.2822872747,-1.83
71036804,-2.1830669419\H,2.4622719455,-3.3205429241,-1.5956217524\O,-2
.9070858673,0.0509860803,0.9486325618\N,-1.8894522253,-1.6589332619,-0
.1463371416\N,2.6040356458,1.4464813476,1.284099747\O,1.8682059243,1.5
127732316,2.2366549776\O,2.3215977311,0.5875084428,0.2952684747\O,3.61
93736144,2.0737906007,1.0797104123\C,0.0458738469,-0.5947363169,0.8622
677624\H,-0.4910276286,0.1821251199,1.4085205349\H,1.0846058697,-0.089
0598134,0.5924453698\H,0.308639591,-1.4491447806,1.4916635199\C,-0.856
1572624,0.2529347652,-1.3484500978\H,-1.3338659466,-0.0474973149,-2.28
548406\H,-1.4979215452,0.9692375248,-0.8326647937\H,0.1039306928,0.724
7018386,-1.5721038551\\Version=AM64L-G09RevB.01\HF=-834.9132925\S2=0.7
59349\S2-1=0.\S2A=0.750053\RMSD=6.118e-09\RMSF=2.110e-06\Dipole=-0.398
734,-2.1661413,-1.2855378\Quadrupole=-0.3613008,-0.0344406,0.3957414,-
1.9639482,-3.5543053,-5.2753882\PG=C01 [X(C7H13N2O6)]\\@
ν_{imag} = -603.2360 cm^{-1}
Sum of electronic and thermal Free Energies= -834.747995 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}C(CH_3)(C^{\bullet}H_2)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC078\FOpt\UM062X\6-31+G(d)\C7H13N2O6(2)\UWILLE\12-Au
g-2018\0\\#p M062X/6-31+G* nosymm freq=noraman geom=checkpoint guess=c
heck scrf=(cpcm,solvent=acetonitrile) opt=(z-matrix,readfc,maxcycle=50
0)\\geom&freq\\0,2\C,-1.0493588833,-0.8688492048,-0.4775805882\C,0.096
2301078,-1.8321374008,-0.8252116257\H,-2.004760156,-2.7280240719,-0.11
39381151\C,-3.2979422779,-1.2469407644,0.4933411716\C,-4.3241715666,-2
.2820849049,0.8945317938\H,-4.5416187561,-2.1605221621,1.9589056918\H,
-5.2460194634,-2.0902127838,0.3394515013\H,-4.0002690997,-3.3076706014
,0.7087719207\O,0.0378469686,-3.0350896013,-0.7459397706\O,1.182482645
3,-1.1802420449,-1.2698939085\C,2.2931390844,-2.0119182221,-1.66991122
36\H,1.9768238579,-2.6746481529,-2.4755860476\H,3.0608820492,-1.322959
4508,-2.0166767288\H,2.6422703646,-2.5897002721,-0.8131997567\O,-3.509
4622532,-0.0419904225,0.6341121178\N,-2.1420967244,-1.727505609,-0.023
1078444\N,2.8230210711,0.2989965847,1.2190116331\O,2.4591627445,-0.842
6788996,1.3993641006\O,2.4282564635,0.9021267751,0.0726394086\O,3.5034

770971,0.9815458582,1.9379823927\C,-0.6038771817,0.0658968918,0.611099
4832\H,-0.6485059928,1.1363680718,0.4570892541\H,1.8501839974,0.239812
3841,-0.3890210752\H,-0.443290473,-0.3266277736,1.6084800438\C,-1.4374
543199,-0.1059361964,-1.7563741993\H,-1.6936018989,-0.8074218591,-2.55
53108428\H,-2.304269825,0.5207643666,-1.5378765611\H,-0.6108029099,0.5
274564364,-2.0871520241\\Version=AM64L-G09RevB.01\HF=-834.9442623\S2=0
.755321\S2-1=0.\S2A=0.750022\RMSD=4.600e-09\RMSF=8.928e-06\Dipole=0.07
88537,-1.3983773,-1.8950272\Quadrupole=-4.2328627,1.443822,2.7890407,4
.0663953,-7.3848359,-0.0577904\PG=C01 [X(C7H13N2O6)]\\@
Sum of electronic and thermal Free Energies= -834.774388 Hartrees

Reaction of $NO_3^\bullet$ with Mal-Aib-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC154\FOpt\UM062X\6-31+G(d)\C9H11N2O7(2)\UWILLE\29-Au
g-2018\0\\#p M062X/6-31+G* scf=(qc,direct) freq=noraman nosymm geom=ch
eckpoint guess=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxc
ycle=500)\\geom&freq\\0,2\C,0.266062214,-1.1152399221,-0.005030567\C,0
.1636322262,-2.6371555676,-0.2091982249\O,-0.5415923606,-3.3531542767,
0.4658692049\O,1.0102728211,-3.0801787369,-1.1320175327\C,1.0567292215
,-4.5041242809,-1.3066403987\H,0.0786027657,-4.8710557307,-1.623110580
8\H,1.8022006328,-4.6783327491,-2.079400969\H,1.3493549486,-4.98489554
26,-0.3713636169\N,-1.7725866406,2.9346680212,-1.3479282745\O,-1.77859
77048,3.163164779,-0.0452553668\O,-1.4263121748,1.8294400597,-1.671965
7249\O,-2.1178221906,3.8712866139,-2.016515864\C,1.1186586448,-0.94836
92188,1.2630297651\H,1.3310607116,0.1041610884,1.4500059235\H,0.614678
9819,-1.3740965871,2.1341028949\H,2.064718352,-1.4758967842,1.10798432
98\C,0.9031941124,-0.3803936426,-1.1841962288\H,0.3288653639,-0.531540
172,-2.1020863646\H,0.9314278461,0.6883912885,-0.9537460108\H,1.923923
039,-0.7341081516,-1.3438051622\N,-1.1092998073,-0.6395193378,0.195081
0371\C,-1.5804682538,0.3410179523,1.0656121768\C,-2.1354173803,-1.0128
81312,-0.6765546912\C,-3.0515776461,0.4899394542,0.7879327515\C,-3.374
9925321,-0.2962054247,-0.2363376377\H,-3.6705865591,1.1646624,1.365429
1556\H,-4.3299337779,-0.4402394684,-0.7245273962\O,-0.9330196886,0.979
1466652,1.8661462476\O,-2.0003241651,-1.7718234165,-1.6112098749\\Vers
ion=AM64L-G09RevB.01\HF=-985.0993769\S2=0.756233\S2-1=0.\S2A=0.75003\R
MSD=0.000e+00\RMSF=4.182e-06\Dipole=0.1330768,0.3761395,-0.1986659\Qua
drupole=11.2882452,-2.4060782,-8.882167,-9.4239371,-3.5111015,2.656101
8\PG=C01 [X(C9H11N2O7)]\\@
Sum of electronic and thermal Free Energies= -984.937552 Hartrees

Entry 10: Reaction at $(CHC(O))_2N$-$C(C\mathit{H}_3)_2$-CO_2CH_3

Transition state

1\1\GINC-SPARTAN-RC040\FTS\UM062X\6-31+G(d)\C9H11N2O7(2)\UWILLE\26-Aug
-2018\0\\#p M062X/6-31+G* nosymm geom=checkpoint guess=check scrf=(cpc
m,solvent=acetonitrile) opt=(ts,noeigentest,readfc,maxcycle=500) freq=
noraman\\geom&freq\\0,2\C,-0.5500304704,-1.001835584,-0.4230048304\C,0
.3186668758,-2.0199409425,-1.1817250946\O,-0.0956577803,-2.7012621558,
-2.0896205036\O,1.5826710706,-1.9784386332,-0.7801510267\C,2.499465118
,-2.8204946158,-1.4975665137\H,2.5190254807,-2.5335669816,-2.550244383
5\H,3.4692097094,-2.6563424674,-1.0334463128\H,2.1938461991,-3.8636540
784,-1.3999513875\N,2.0913876727,1.9472580961,0.5678234518\O,0.9821601
781,2.3609735242,0.7957351891\O,2.2869397942,0.6250423414,0.452982519\
O,3.1089185294,2.5824469725,0.4090352782\C,0.0330315366,-0.6128033514,
0.9289903953\H,-0.5722011156,0.13978051,1.4399304769\H,1.0640736785,-0

.0555774107,0.7220629509\H,0.302851973,-1.4555054244,1.5683751168\C,-0
.6645017536,0.2175582398,-1.3496981185\H,-1.203827459,-0.0408267934,-2
.2631330411\H,-1.1745371823,1.0400668111,-0.8481066274\H,0.3475982419,
0.5375889524,-1.6159748376\N,-1.841548133,-1.6437891923,-0.1733377275\
C,-1.9091274762,-2.9312721031,0.3752651367\C,-3.1374155366,-1.12285826
07,-0.3019590745\C,-3.35789256,-3.2659222055,0.533001905\C,-4.07200085
07,-2.213691906,0.1387438484\H,-3.6902166619,-4.2210522075,0.917760674
8\H,-5.1455135975,-2.0787717097,0.1151512015\O,-0.9418113655,-3.601257
963,0.6597241271\O,-3.4398305352,-0.0160918411,-0.6783922421\\Version=
AM64L-G09RevB.01\HF=-985.0847141\S2=0.759409\S2-1=0.\S2A=0.750053\RMSD
=6.348e-09\RMSF=3.198e-06\Dipole=-1.0285093,-1.4529484,0.6593416\Quadr
upole=8.7077674,-7.8066811,-0.9010863,2.5716875,-7.9792225,-4.132556\P
G=C01 [X(C9H11N2O7)]\\@

ν_{imag} = -711.0227 cm^{-1}

Sum of electronic and thermal Free Energies= - 984.925009 Hartrees

Product association complex [HNO_3 – $(CHC(O))_2N$-$C(CH_3)(C^{\bullet}H_2)$-CO_2CH_3]

1\1\GINC-SPARTAN-RC019\FOpt\UM062X\6-31+G(d)\C9H11N2O7(2)\UWILLE\27-Au
g-2018\0\\#p M062X/6-31+G* scf=(qc,direct) scrf=(cpcm,solvent=acetonit
rile) geom=checkpoint guess=check nosymm opt=(readfc,maxcycle=500) fre
q=noraman\\geom&freq\\0,2\C,-0.9926321749,-0.9893064415,-0.5224094127\
C,0.0803742346,-2.0182096368,-0.927817571\O,-0.148175102,-3.0758804291
,-1.4544107473\O,1.3239195126,-1.5578248436,-0.709237847\C,2.400757605
6,-2.4258790466,-1.1228813711\H,2.3475065527,-2.5864441781,-2.20003295
68\H,3.3158280952,-1.9037245015,-0.8506582135\H,2.3228218154,-3.374051
5522,-0.590681621\N,2.6816704327,1.1619390373,0.8842152589\O,2.6339002
154,0.1825904508,1.5945375789\O,2.1951373185,1.0378587365,-0.372700757
2\O,3.1149315074,2.2519120217,1.1521368084\C,-0.5724326788,-0.13974285
33,0.6417473338\H,-0.9488903351,0.8749160305,0.6984797077\H,1.82721024
4,0.1162081296,-0.4253017725\H,-0.1487529076,-0.6014118361,1.524560308
7\C,-1.2265028787,-0.1244761421,-1.7723015033\H,-1.5684621015,-0.72961
0128,-2.6141700322\H,-1.96596257,0.649708249,-1.5656652548\H,-0.280111
1606,0.3560413904,-2.03689282\N,-2.1862749373,-1.7718861552,-0.1581441
405\C,-2.0955493473,-2.7670149562,0.8196282304\C,-3.5096282465,-1.6721
171879,-0.6021826839\C,-3.4620393591,-3.3541903675,0.9820391559\C,-4.2
851266916,-2.7167450778,0.1526876927\H,-3.6656024549,-4.1563856043,1.6
790970172\H,-5.3454669271,-2.8583436882,-0.0118555909\O,-1.0739267337,
-3.0606255253,1.4003092466\O,-3.9512833972,-0.8989538845,-1.4197483435
\\Version=AM64L-G09RevB.01\HF=-985.1146514\S2=0.755051\S2-1=0.\S2A=0.7
50019\RMSD=0.000e+00\RMSF=7.507e-06\Dipole=-0.9457572,-0.9376983,-0.45
71124\Quadrupole=7.2957586,-3.2369433,-4.0588153,0.3411719,-14.3925716
,-2.5805333\PG=C01 [X(C9H11N2O7)]\\@

Sum of electronic and thermal Free Energies= - 984.952960 Hartrees

Reaction of $NO_3^{\bullet}$ with Ac(F_3)-Aib-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC005\FOpt\UM062X\6-31+G(d)\C7H10F3N2O6(2)\UWILLE\08-
Dec-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent
=acetonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\
C,-1.8888167984,-0.7728606896,0.9927376571\O,-1.6627764142,0.383987242
,1.2814178053\O,-4.3616032105,-5.2307251123,2.9451724983\O,-5.47873656
52,-5.0777842335,4.8099471652\O,-3.322825319,-5.3589244978,4.674799379
7\N,-4.4697924185,-5.2068876725,4.1923317801\C,-2.6513314069,-1.747140
0253,1.8923756546\H,-2.6863133949,0.0043495296,3.0874898569\C,-3.66621
4156,-1.4503266592,4.127255567\C,-3.9027688086,-0.44150344,5.274061041

3\N,-2.9924726136,-0.9660623546,3.0852716223\O,-4.1079541441,-2.580956
0253,4.276629253\C,-0.7859239915,-0.5139504557,-1.0598400017\H,0.13546
18547,-0.1598887342,-0.5945690092\H,-0.5669152028,-1.1492443333,-1.914
9016477\H,-1.4051350155,0.3344731808,-1.3552064748\C,-3.921538282,-2.2
160024871,1.1687126224\H,-4.4957598073,-2.8679209373,1.8300775852\H,-4
.5404979429,-1.360171448,0.883372638\H,-3.6484297927,-2.7740412669,0.2
704594974\C,-1.7369984557,-2.9248770116,2.2575195672\H,-0.8214015918,-
2.5679484013,2.7385253299\H,-2.2626796287,-3.5929427831,2.9426661055\H
,-1.4730692257,-3.4800143832,1.3547591684\O,-1.5082342419,-1.347498876
,-0.138167416\F,-3.322928681,-0.8828786552,6.3954775782\F,-5.212698874
1,-0.318742378,5.5114745127\F,-3.4154569613,0.7806867985,5.0129750439\
\Version=AM64L-G09RevB.01\HF=-1132.5755877\S2=0.761086\S2-1=0.\S2A=0.7
50072\RMSD=0.000e+00\RMSF=5.434e-06\Dipole=1.2786053,0.1977343,-2.4218
639\Quadrupole=-0.6049997,8.7316685,-8.1266688,-7.2618305,20.0046729,8
.8068127\PG=C01 [X(C7H10F3N2O6)]\\@
Sum of electronic and thermal Free Energies= -1132.431623 Hartrees

Entry 11: Reaction at $CF_3C(O)N\mathit{H}\text{-}C(C\mathit{H}_3)_2\text{-}CO_2CH_3$

Charge transfer complex

1\1\GINC-SPARTAN-RC040\FOpt\UM062X\6-31+G(d)\C7H10F3N2O6(2)\UWILLE\08-
Dec-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent
=acetonitrile) freq=noraman opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\
C,-0.5683503034,-1.8097702312,0.8068662894\H,-2.5405249731,-2.30925563
13,1.4184117591\C,-2.0546262669,-0.5502995846,2.379956289\O,-1.3134971
802,0.387548842,2.549139797\N,-1.7667320207,-1.6785022264,1.653887359\
O,-3.341365701,-3.774650137,2.8323512244\O,-2.0502360249,-4.9075080064
,4.1608423765\O,-1.3067428124,-3.1202892084,3.2819520863\N,-2.30137705
06,-3.9784865347,3.4213157833\C,-3.4494163315,-0.5639621979,3.05401533
5\C,-0.6745282987,-3.2234760325,0.2245694343\O,0.4248488094,-3.5697586
371,-0.4209935591\O,-1.6662525186,-3.9145616364,0.3222191465\C,0.40825
85073,-4.8645728748,-1.0486088581\H,1.3831778086,-4.9698392921,-1.5183
885658\H,0.2549808616,-5.6367949865,-0.2935355818\H,-0.3888125984,-4.9
030436619,-1.7926710228\C,-0.6462854409,-0.7953732174,-0.3499128993\H,
-1.5768652649,-0.9109621864,-0.9121898734\H,-0.5890782763,0.2191842882
,0.0510773455\H,0.1998112811,-0.9573087179,-1.0219034281\C,0.730051356
,-1.6407509423,1.6041154409\H,0.7827333295,-0.6364401362,2.0254982808\
H,0.801898247,-2.3778457938,2.4060463262\H,1.5700569279,-1.7810234128,
0.9214123022\F,-4.3626869535,-1.1610384463,2.2800836563\F,-3.395087904
9,-1.2176216819,4.2203770248\F,-3.8453031875,0.6812765139,3.2922343119
\\Version=AM64L-G09RevB.01\HF=-1132.5760034\S2=0.759516\S2-1=0.\S2A=0.
750075\RMSD=0.000e+00\RMSF=1.304e-05\Dipole=1.7747251,0.4488522,-2.758
3122\Quadrupole=2.5030938,1.3053151,-3.8084089,-15.3837176,10.7008029,
18.0546109\PG=C01 [X(C7H10F3N2O6)]\\@
Sum of electronic and thermal Free Energies= -1132.428750 Hartrees

Transition state

1\1\GINC-SPARTAN-RC174\FTS\UM062X\6-31+G(d)\C7H10F3N2O6(2)\UWILLE\09-D
ec-2018\0\\#p M062X/6-31+G* nosymm scrf=(cpcm,solvent=acetonitrile) op
t=(ts,noeigentest,calcfc,maxcycle=500) freq=noraman scf=(qc,direct)\\g
eom&freq\\0,2\C,-1.2660560996,-1.7239321097,0.1275652163\C,-2.31554213
38,-2.1074285131,-0.927059461\H,-3.1415476953,-1.6050547817,1.35287558
79\C,-1.4022297284,-1.1590768206,2.5577257327\O,-3.4140023198,-1.60344
42974,-0.9519492516\O,-1.8473558969,-2.970503227,-1.8080222736\C,-2.74
74786613,-3.327171973,-2.8743372077\H,-3.0200057759,-2.4346404137,-3.4
393636984\H,-2.1940543991,-4.0272678659,-3.4952670759\H,-3.6398030049,
-3.7960703204,-2.4566875054\O,-0.2346661478,-0.8982722862,2.6729764886

\N,-2.0215360827,-1.4762944421,1.3377899577\O,-4.471834308,-2.15351026
61,1.6313933405\O,-5.2033511254,-4.1888836003,1.6230248455\O,-3.101462
1873,-3.7945973513,1.293864287\N,-4.2720362764,-3.4065428729,1.5161939
049\C,-2.3646976194,-1.149728899,3.7697122888\C,-0.6355670873,-0.37773
14312,-0.3495683205\H,0.1406826485,-0.0647469869,0.3490253819\H,-0.189
7921154,-0.5661947785,-1.3297578594\H,-1.4013404122,0.3948426852,-0.43
95099645\C,-0.1829976438,-2.79408976,0.3383407499\H,0.5623172339,-2.45
13589717,1.0551421143\H,-0.6399936787,-3.7241291491,0.6826693943\H,0.3
041022469,-2.9658128337,-0.6223243701\F,-3.3948821549,-0.33509005,3.53
62776907\F,-1.7173131818,-0.7348274247,4.8471636472\F,-2.8313358631,-2
.3807203589,3.9882886502\\Version=AM64L-G09RevB.01\HF=-1132.5612886\S2
=0.765968\S2-1=0.\S2A=0.750109\RMSD=0.000e+00\RMSF=6.235e-06\Dipole=3.
5803796,1.614512,-2.0666147\Quadrupole=-24.1525229,2.8608853,21.291637
6,-24.9917936,15.1402466,15.2922824\PG=C01 [X(C7H10F3N2O6)]\\@

ν_{imag} = -257.2253 cm^{-1}

Sum of electronic and thermal Free Energies= -1132.417237 Hartrees

Product association complex [HNO_3 - $CF_3C(O)N^{\bullet}$-$C(CH_3)_2$-CO_2CH_3]

1\1\GINC-SPARTAN-RC092\FOpt\UM062X\6-31+G(d)\C7H10F3N2O6(2)\UWILLE\09-
Dec-2018\0\\#p M062X/6-31+G* nosymm scf=(qc,direct) scrf=(cpcm,solvent
=acetonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\
C,-0.2984635207,-0.9800706448,-0.045500369\C,0.1289317904,-2.435032007
4,0.1767591127\H,-2.1208659943,-3.1469999455,1.1808429468\C,-2.1225872
136,0.4354962354,0.815911722\O,-0.5223549668,-3.2724083839,0.776767470
9\O,1.2965622126,-2.6899533231,-0.3737838264\C,1.7910224624,-4.0377912
92,-0.2511010787\H,1.0783599466,-4.7322378674,-0.6974091592\H,2.734700
3614,-4.0464016712,-0.7906431983\H,1.9418496371,-4.2766014214,0.802476
8264\O,-1.6801132682,1.4018252987,1.3888001876\N,-1.5219783686,-0.7976
678961,0.7038704542\O,-3.0843597987,-3.3223110335,1.4198848826\O,-4.97
46335796,-3.3415870284,0.3910334503\O,-3.1858916059,-2.7864470057,-0.6
99308575\N,-3.7902603559,-3.1396497383,0.2937925918\C,-3.5685922672,0.
4416067697,0.2561231139\C,0.8065013811,-0.0252733957,0.4553996329\H,0.
5236500196,1.0094386513,0.2571021522\H,0.9796266454,-0.1512198843,1.52
69486769\H,1.7246013653,-0.2548953823,-0.0887709712\C,-0.5616587322,-0
.7605572787,-1.5514194982\H,0.3563883499,-0.9705049727,-2.1045829391\H
,-1.3599187547,-1.4166973437,-1.9061910375\H,-0.8454819449,0.281564727
7,-1.7235189737\F,-4.1410211939,1.6241747932,0.4428102355\F,-4.3122595
798,-0.4923862717,0.8641567437\F,-3.5588540565,0.1732286317,-1.0551498
142\\Version=AM64L-G09RevB.01\HF=-1132.5803763\S2=0.755029\S2-1=0.\S2A
=0.750021\RMSD=0.000e+00\RMSF=7.053e-06\Dipole=3.8384155,-0.2516882,-1
.0296754\Quadrupole=-13.7467878,6.9479898,6.798798,-18.7507044,2.90697
11,1.5036795\PG=C01 [X(C7H10F3N2O6)]\\@

Sum of electronic and thermal Free Energies= -1132.435250 Hartrees

Reaction of $NO_3^{\bullet}$ with Ac-Val-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC053\FOpt\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\23-Se
p-2018\0\\#p M062X/6-31+G* scrf=(cpcm,solvent=acetonitrile) scf=(qc,di
rect) nosymm freq=noraman opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\C,
-0.5839394047,-1.2803005431,0.0566723321\H,-0.1027367561,-1.0381841821
,-0.8975278663\C,-0.6522001484,-2.7889412429,0.2084743727\H,-2.6670902
867,-1.3922042093,0.4147016998\C,-2.2765725399,0.3885423189,-0.5391843
147\O,-1.6353441017,-3.3871586706,0.5943160155\O,0.5009094351,-3.37381
33127,-0.0952816479\C,0.5551978979,-4.7971818452,0.0905645347\H,-0.175

2634987,-5.2843471138,-0.5577205648\H,1.5672518939,-5.0874737275,-0.18
26711135\H,0.3499804754,-5.0441450866,1.1337841037\O,-1.4300367732,1.1
249515124,-1.0531237346\N,-1.9487788208,-0.7924290499,0.0254277881\C,0
.2330561815,-0.6240749855,1.2025179176\H,0.0558284659,0.4502242137,1.0
684180657\C,-0.2856247143,-1.045841311,2.5768839897\H,-0.0878551134,-2
.1094986577,2.762597324\H,0.2187328134,-0.4746064671,3.3625095098\H,-1
.3632882063,-0.8760966604,2.6739932624\C,1.7353096365,-0.8778039925,1.
0743527105\H,1.9895763696,-1.9128290613,1.3227173028\H,2.0924776222,-0
.6740847031,0.0595699014\H,2.2777458813,-0.2209503107,1.7624679926\N,0
.9756147744,1.8866102366,-1.8212490349\O,1.3659742769,1.9299793587,-0.
628782278\O,1.2368602612,0.7487889808,-2.2666069629\O,0.4707998136,2.7
820911997,-2.4208052565\C,-3.739681326,0.7620440615,-0.5159505458\H,-3
.8445968395,1.7357992465,-0.0310265163\H,-4.0900703746,0.8611754384,-1
.5468550154\H,-4.3607809847,0.0311572357,0.0050560583\\Version=AM64L-G
09RevB.01\HF=-874.2232349\S2=0.761185\S2-1=0.\S2A=0.750073\RMSD=0.000e
+00\RMSF=2.490e-06\Dipole=-0.0568935,-1.4128887,0.8322783\Quadrupole=4
.2355084,0.7344478,-4.9699561,-4.6733347,0.061408,4.1529773\PG=C01 [X(
C8H15N2O6)]\\@

Sum of electronic and thermal Free Energies= -874.027375 Hartrees

Entry 12: Reaction at $CH_3C(O)NH\text{-}C\mathit{H}(CH(CH_3)_2)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC053\FTS\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\08-Nov
-2017\0\\#p M062X/6-31+G* nosymm scf=(qc,direct) freq=noraman geom=che
ckpoint guess=check scrf=(cpcm,solvent=acetonitrile) opt=(ts,noeigente
st,calcfc,z-matrix,maxcycle=500)\\geom&freq\\0,2\C,-0.1695987309,-0.64
37174057,0.439944317\H,0.2857911423,-0.1236105592,-0.4568412797\C,0.47
77053611,-2.0159701667,0.5599302665\H,-1.8942828496,-1.8062137968,0.23
30211762\C,-2.3985919161,0.1325202919,-0.2666876771\O,-0.1479465984,-3
.032011595,0.7758437986\O,1.79152156,-1.9597315848,0.4105594419\C,2.49
95941518,-3.2020898248,0.56261447\H,2.147634671,-3.9212801762,-0.17861
6971\H,3.5471275389,-2.9597900878,0.3997827747\H,2.3435240596,-3.59550
88347,1.5685130528\O,-1.9992651891,1.2768619786,-0.4544417269\N,-1.558
1134798,-0.8475934289,0.1850066043\N,0.3028668304,1.9564023079,-1.8408
386818\O,0.6559938356,2.4236378677,-0.7885961379\O,0.3142672542,0.6515
364677,-1.9954784019\O,-0.0601475432,2.5547209758,-2.835369963\C,0.115
1862788,0.2866554299,1.6543744083\H,-0.4364412634,1.2051199186,1.43064
18167\C,-0.4415413218,-0.3403718467,2.9357795958\H,0.0972299607,-1.261
2517084,3.1911724057\H,-0.3177502916,0.3595077676,3.7682510878\H,-1.50
57343585,-0.5770884972,2.8405794503\C,1.5954360734,0.6350024658,1.8039
635983\H,2.1743625959,-0.2186273467,2.1688192578\H,2.0306426229,0.9672
652631,0.8567728892\H,1.6979968288,1.4486752914,2.5289601764\C,-3.8128
004951,-0.30674522,-0.5511894826\H,-4.4620405329,0.5687324388,-0.55005
0102\H,-3.8498269267,-0.7711112319,-1.5421359869\H,-4.1687662693,-1.03
44211558,0.18190882\\Version=AM64L-G09RevB.01\HF=-874.2170714\S2=0.760
046\S2-1=0.\S2A=0.750071\RMSD=0.000e+00\RMSF=8.318e-06\Dipole=0.063129
3,-2.5625286,1.3671904\Quadrupole=12.3198351,-8.5923377,-3.7274974,-1.
2925954,1.2424335,6.5156191\PG=C01 [X(C8H15N2O6)]\\@

ν_{imag} = -501.6113 cm^{-1}

Sum of electronic and thermal Free Energies= -874.021163 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}C^{\bullet}(CH(CH_3)_2)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC080\FOpt\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\25-No
v-2017\0\\#p M062X/6-31+G* nosymm scf=(qc,direct) geom=checkpoint gues
s=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,z-matrix,maxcycle
=500) freq=noraman\\geom&freq\\0,2\C,-0.0776959405,-1.6790907905,0.121

3561685\C,-0.0588184239,-2.721237935,-0.8992974363\H,-1.7782277923,-1.
1019033373,-0.7983848293\C,-1.2868690381,0.4679238817,0.3943179761\C,-
2.5554125611,1.1720917587,0.0053603178\H,-3.0839363242,1.4677700442,0.
9164912909\H,-2.2939090916,2.0808509942,-0.5444559752\H,-3.2154644695,
0.5546938078,-0.6046239686\O,-0.9105143604,-2.8073859433,-1.7793385203
\O,0.9750187856,-3.5618732293,-0.7923425449\C,1.0380289138,-4.60036627
06,-1.777149243\H,1.1364729766,-4.1661585588,-2.7740808236\H,1.9201328
427,-5.1854245459,-1.525253216\H,0.1398335992,-5.2191279338,-1.7311529
274\O,-0.4312395777,1.0089320778,1.114611878\N,-1.1145719839,-0.789167
8242,-0.0909016562\C,0.9255695341,-2.9232186765,2.0829712596\H,1.47804
69531,-3.6859142333,1.5311948022\H,1.4415454659,-2.7405841008,3.031001
6068\H,-0.0731826473,-3.3104412337,2.3097776275\H,-0.6134255341,2.4043
853385,1.6155222461\O,-0.7231186372,5.4074491445,1.4850256382\O,-0.696
1760316,3.335817873,2.0700731604\O,-0.5554455476,3.8795468598,-0.04189
44872\N,-0.6546764595,4.2610280361,1.1090044331\C,2.2450017735,-1.1156
802866,0.8852128829\H,2.1917923566,-0.1329585845,0.4071306111\H,2.8864
275701,-1.0367817021,1.7691922667\H,2.7037168914,-1.8236131139,0.18788
22811\C,0.8474155208,-1.5989748583,1.3044975324\H,0.4253332366,-0.8520
82658,1.9776456485\\Version=AM64L-G09RevB.01\HF=-874.271378\S2=0.75764
2\S2-1=0.\S2A=0.750045\RMSD=0.000e+00\RMSF=6.880e-06\Dipole=-0.3850542
,-3.2867636,-0.3536132\Quadrupole=13.0242807,-13.987155,0.9628743,-4.0
219451,-1.4817161,-6.3926602\PG=C01 [X(C8H15N2O6)]\\@

Sum of electronic and thermal Free Energies= -874.073866 Hartrees

Entry 13: Reaction at $CH_3C(O)NH\text{-}CH(C\mathit{H}(CH_3)_2)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC214\FTS\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\21-Nov
-2017\0\\#p M062X/6-31+G* freq=noraman nosymm scf=(qc,direct) scrf=(cp
cm,solvent=acetonitrile) geom=checkpoint guess=check opt=(ts,noeigente
st,readfc,z-matrix,maxcycle=500)\\geom&freq\\0,2\C,-0.6659297695,-1.04
56771775,-0.2514157192\H,-0.9176438106,-0.2636200868,-0.9751509177\C,0
.2726850133,-2.0455881694,-0.9079114736\H,-1.814143598,-2.7062685505,0
.3627199629\C,-3.1118076665,-1.1666203107,-0.0525173548\O,0.0927721087
,-3.2433260673,-0.9295410521\O,1.3395340166,-1.4352960097,-1.410038049
9\C,2.3478058481,-2.281681079,-1.9839154771\H,1.9291215026,-2.84288706
21,-2.821084585\H,3.1319553144,-1.6088928386,-2.3243310579\H,2.7293962
077,-2.9689942926,-1.2266287002\O,-3.2545117143,-0.0147935464,-0.45796
19431\N,-1.8848575216,-1.7253690199,0.1177067486\N,2.9320288205,1.3931
993275,1.1737837159\O,2.2890527768,1.4959160875,2.1864956667\O,2.43828
05956,0.6860562999,0.1635466401\O,4.0223259477,1.857816376,0.923653416
\C,0.0216914789,-0.3900017483,0.9718408014\H,0.943534479,0.1270125158,
0.5310641607\C,-0.8352833889,0.69834132,1.5928341157\H,-1.1477864149,1
.4373352109,0.850090367\H,-1.7385290097,0.2521172594,2.0285874706\H,-0
.2864732743,1.2009312634,2.3938628937\C,0.5397750102,-1.396620362,1.98
66609633\H,1.2439325033,-2.1092134002,1.5431245905\H,1.047346044,-0.87
8453535,2.8051650212\H,-0.2976338112,-1.9617282871,2.4155693204\C,-4.2
909775354,-2.047192009,0.2871528316\H,-4.9431094904,-1.5063519964,0.97
68515706\H,-4.8556059315,-2.2404026593,-0.6291937287\H,-4.0002365592,-
2.9990697508,0.7352286972\\Version=AM64L-G09RevB.01\HF=-874.2163716\S2
=0.758639\S2-1=0.\S2A=0.750052\RMSD=0.000e+00\RMSF=5.844e-06\Dipole=-0
.0768033,-1.7577317,0.2142518\Quadrupole=-3.0223119,-0.2883501,3.31066
2,0.6046735,-9.262321,-2.8259089\PG=C01 [X(C8H15N2O6)]\\@

ν_{imag} = -264.5643 cm^{-1}

Sum of electronic and thermal Free Energies= -874.022990 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}CH(C^{\bullet}(CH_3)_2)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC161\FOpt\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\17-No
v-2017\0\\#p M062X/6-31+G* nosymm geom=checkpoint guess=check scrf=(cp
cm,solvent=acetonitrile) opt=(readfc,z-matrix,maxcycle=500) freq=noram
an\\geom&freq\\0,2\C,0.3223526858,-1.3509477797,-0.2503671727\H,0.9571
830369,-0.6631941436,-0.8200251354\C,0.1322334831,-2.6265579848,-1.065
5326063\H,-1.7999441642,-1.321043691,-0.3490794087\C,-1.1629692558,0.4
637384909,0.4142254918\C,-2.56139419,0.9866499667,0.5754538922\H,-2.72
68820897,1.2293250542,1.6290046341\H,-2.6597948605,1.907259498,-0.0062
883349\H,-3.3147569687,0.268576457,0.2501006155\O,-0.9434247531,-3.147
6254169,-1.2705441705\O,1.2985340612,-3.1217366305,-1.4580627196\C,1.2
491223699,-4.3761818348,-2.1582735686\H,0.6644479437,-4.2681742733,-3.
0733764523\H,2.2835974261,-4.6201812985,-2.388532803\H,0.8043445718,-5
.1399143159,-1.5179745729\O,-0.1626023634,1.1503459501,0.7321858915\N,
-0.9973323622,-0.7599331841,-0.0803581594\C,0.9879426435,-1.6898587432
,1.0612598585\C,0.2692112233,-2.6061665313,1.9968614398\H,0.4909535112
,-3.6640177503,1.7730473757\H,0.5841238372,-2.4303291369,3.0310593148\
H,-0.817207033,-2.4795978665,1.9416618409\H,-0.348974789,2.4361103113,
1.2400848428\O,-0.8900685673,5.4347826525,1.2537257078\O,-0.4025629334
,3.3924286538,1.7283374293\O,-1.0988761532,3.9139833209,-0.273394898\N
,-0.8191487572,4.2943131423,0.8507144949\C,2.4680457133,-1.544649161,1
.1655671989\H,2.8285387904,-0.6743315805,0.6068469497\H,2.7728690112,-
1.4381193725,2.2125412467\H,2.9906409261,-2.4300948045,0.7674977714\\V
ersion=AM64L-G09RevB.01\HF=-874.2598791\S2=0.755148\S2-1=0.\S2A=0.7500
22\RMSD=3.142e-09\RMSF=1.117e-05\Dipole=0.1068522,-4.3387569,-0.621224
3\Quadrupole=12.1047981,-17.6655487,5.5607506,-3.2140047,-3.7859017,-3
.2024859\PG=C01 [X(C8H15N2O6)]\\@

Sum of electronic and thermal Free Energies= -874.067722 Hartrees

Entry 14: Reaction at $CH_3C(O)NH\text{-}CH(CH(C\mathit{H}_3)_2)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC151\FTS\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\18-Nov
-2017\0\\#p M062X/6-31+G* freq=noraman nosymm scrf=(cpcm,solvent=aceto
nitrile) scf=(qc,direct) geom=checkpoint guess=check opt=(ts,noeigente
st,readfc,z-matrix,maxcycle=500)\\geom&freq\\0,2\C,-0.8132095298,-1.54
40512725,-0.5562468404\H,-1.3998059933,-0.9560443404,-1.2702889851\C,-
0.1421858579,-2.6918419732,-1.2964563373\H,-1.511402205,-3.0392278292,
0.7569943151\C,-2.9055684823,-1.529428273,0.721369035\C,-3.7904056621,
-2.2893385516,1.6802754577\H,-4.0332284553,-1.6378173478,2.5234568631\
H,-4.7238834286,-2.5377407881,1.1679085362\H,-3.3309453564,-3.20624915
95,2.0538852307\O,-0.2307350425,-3.8559130156,-0.9714291691\O,0.584780
5497,-2.2522950914,-2.316845769\C,1.3111987014,-3.2513585397,-3.051338
914\H,0.6172147052,-3.9791669191,-3.4746547132\H,1.8323370419,-2.71196
78123,-3.8388338804\H,2.020444023,-3.7527381983,-2.3901479285\O,-3.243
7631535,-0.4464442544,0.2434816402\N,-1.7143525234,-2.104403874,0.4221
448537\N,1.3979418047,2.9770721381,2.4153546751\O,0.2537649712,3.04681
12583,2.7821526429\O,1.7423813634,2.0476044219,1.5214484\O,2.339593373
,3.6605340358,2.7522986487\C,0.2855215349,-0.6256101133,0.06226874\H,0
.8634168285,-0.2407107071,-0.7862359556\C,-0.3715791999,0.5383320492,0
.7725546423\H,-0.9986673733,1.1673284732,0.1362531343\H,-0.8960823248,
0.2614156867,1.6927077212\H,0.4813132586,1.2324170994,1.1290000299\C,1
.206807174,-1.4066619637,0.9985946995\H,1.7407436719,-2.204728052,0.47
11310687\H,1.9559642484,-0.7399382963,1.4361148253\H,0.6349993347,-1.8
563777877,1.8182493381\\Version=AM64L-G09RevB.01\HF=-874.2107856\S2=0.
758996\S2-1=0.\S2A=0.750053\RMSD=0.000e+00\RMSF=1.453e-05\Dipole=0.537
5932,-2.3320551,-0.9868716\Quadrupole=-2.6295555,-3.9851171,6.6146726,
-1.2088017,-9.2502895,-6.903847\PG=C01 [X(C8H15N2O6)]\\@

ν_{imag} = -408.3455 cm^{-1}
Sum of electronic and thermal Free Energies= -874.017566 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH$-$CH(CH(CH_3)(C^{\bullet}H_2)$-$CO_2CH_3$]

1\1\GINC-SPARTAN-RC145\FOpt\UM062X\6-31+G(d)\C8H15N2O6(2)\UWILLE\30-No
v-2017\0\\#p M062X/6-31+G* nosymm opt=(calcfc,z-matrix,maxcycle=500) s
crf=(cpcm,solvent=acetonitrile) freq=noraman\\geom&freq\\0,2\C,-0.1908
509122,-1.0471299516,-0.3871373853\H,0.09925473,-0.6227155349,-1.35462
57257\C,0.3114243572,-2.4765087823,-0.303448366\H,-2.0637088403,-1.926
4378774,0.0467252617\C,-2.3919009449,-0.0560721914,-0.7603281072\O,-0.
3833773409,-3.425230266,-0.006037774\O,1.6132439714,-2.5542194159,-0.5
531150156\C,2.2037951546,-3.858981776,-0.4374029336\H,1.7494710137,-4.
5379234928,-1.1610208425\H,3.2608812332,-3.7185464221,-0.6515453061\H,
2.0615004799,-4.2430311241,0.5743716023\O,-1.8879036943,0.9591816519,-
1.2442458882\N,-1.6323701651,-1.0784491169,-0.3027876193\C,0.437673069
1,-0.1716839921,0.7479604474\H,-0.1578987915,0.7505937977,0.7190837143
\C,0.2798655572,-0.8311908976,2.1172085764\H,0.9001846329,-1.733185345
8,2.1932282013\H,0.5930137078,-0.1413039175,2.9060343536\H,-0.76188786
94,-1.111202588,2.304041107\C,1.8581630045,0.1733317699,0.4453267402\H
,2.6660528477,-0.3694413684,0.9278121876\H,2.0944585708,0.7091421085,-
0.4706527885\H,2.376369024,1.8916103087,1.5258911993\O,0.5814677119,2.
5381176197,2.5808652885\O,2.6776844516,2.7291989626,1.967135879\O,1.78
44673939,4.2205482322,3.2391799769\N,1.5976146045,3.1927997152,2.64425
93212\C,-3.8874329132,-0.2242878495,-0.6374343221\H,-4.273242311,0.558
8704428,0.0207489612\H,-4.3370040802,-0.0860266108,-1.6237749483\H,-4.
178383652,-1.1989560886,-0.2412367963\\Version=AM64L-G09RevB.01\HF=-87
4.2358845\S2=0.754718\S2-1=0.\S2A=0.750018\RMSD=5.790e-09\RMSF=2.538e-
05\Dipole=0.7398404,-2.6815834,-0.3292895\Quadrupole=13.7896794,-9.294
4311,-4.4952483,-6.6292915,-3.7705189,-5.9470631\PG=C01 [X(C8H15N2O6)]\\@
Sum of electronic and thermal Free Energies= -874.040547 Hartrees

Reaction of $NO_3^{\bullet}$ with Ac-Leu-OMe

Reactant association complex

1\1\GINC-SPARTAN-RC102\FOpt\UM062X\6-31+G(d)\C9H17N2O6(2)\UWILLE\22-Se
p-2018\0\\#p M062X/6-31+G* scf=(qc,direct) scrf=(cpcm,solvent=acetonit
rile) nosymm opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\C,
-1.9260196507,-1.4883226788,0.0779497828\O,-1.3261103172,-0.43397344,0
.0979265314\O,-4.0989615104,-4.9785271998,2.0073689654\O,-6.0078377096
,-4.8430360117,2.9207722371\O,-4.1943269559,-4.8718258683,4.133125875\
N,-4.8225763584,-4.882703418,3.1105847546\C,-2.6199991701,-2.082082711
2,1.2876146247\H,-3.450930995,-2.7075533866,0.9502628899\H,-2.67069200
6,-0.0932102672,2.018802052\C,-4.1935460315,-1.1516354245,2.9145274237
\N,-3.1597761863,-0.9799764667,2.0623067614\O,-4.7633842448,-2.2396092
089,3.0339122945\C,-1.339728103,-1.8169724806,-2.170163392\H,-0.273534
3587,-1.7053781776,-1.9603969265\H,-1.5052335287,-2.5879403286,-2.9191
616733\H,-1.7530469175,-0.8620270989,-2.4994132441\C,-1.6606697414,-2.
9355473589,2.1427475754\H,-2.2622728373,-3.2785246793,2.9935012856\C,-
1.0210278947,-4.1345352912,1.422880589\H,-1.7231829036,-4.4977486537,0
.6582646517\C,-4.6212584864,0.0590747607,3.7085232443\H,-4.0484464471,
0.9555299255,3.4644736595\H,-5.6812204427,0.2448835549,3.5181614662\H,
-4.5047544751,-0.1636852428,4.7724963323\H,-0.8796093391,-2.2781766182
,2.5495348758\C,-0.7669931409,-5.268887706,2.4187089192\H,-1.700252652
9,-5.6250385632,2.8686971086\H,-0.2778337694,-6.1185787388,1.930664879
6\H,-0.1141495735,-4.9234128124,3.2301673445\C,0.294646331,-3.75194804

33,0.737436875\H,0.1895173138,-2.9038294275,0.0537185127\H,1.041692167
3,-3.4766305588,1.4923564175\H,0.691555017,-4.5943319608,0.1610766048\
O,-2.0283640812,-2.2685274189,-0.9935642985\\Version=AM64L-G09RevB.01\
HF=-913.5156644\S2=0.756156\S2-1=0.\S2A=0.750029\RMSD=0.000e+00\RMSF=8
.846e-06\Dipole=1.2485068,0.6467086,-1.1649494\Quadrupole=-12.145132,5
.3096715,6.8354605,-16.7778694,11.3720276,13.8266194\PG=C01 [X(C9H17N2
O6)]\\@
Sum of electronic and thermal Free Energies= -913.290296 Hartrees

Entry 15: Reaction at $CH_3C(O)NH\text{-}CH(CH_2C\mathit{H}(CH_3)_2)\text{-}CO_2CH_3$

Transition state

1\1\GINC-SPARTAN-RC168\FTS\UM062X\6-31+G(d)\C9H17N2O6(2)\UWILLE\12-Oct
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cp
cm,solvent=acetonitrile) opt=(ts,noeigentest,calcfc,maxcycle=500)\\geo
m&freq\\0,2\C,-1.8394155447,0.8872656128,2.6824523549\H,-1.258091634,1
.6270657377,2.1182600028\C,-3.253096606,0.9018104764,2.1334712437\H,-2
.796377374,1.3087454002,4.5212372903\C,-0.8002496781,1.7798178514,4.70
9204219\C,-0.9949574551,2.1735896024,6.1540897082\H,-0.4373746791,1.47
22718385,6.7819055871\H,-0.5736247972,3.1697352644,6.3071039797\H,-2.0
419692454,2.1693470836,6.4628691344\O,-4.2476371448,1.0669307802,2.808
442725\C,-4.5602351319,0.6203507638,0.2060993683\H,-5.1390686197,-0.19
33001086,0.6488972951\H,-5.086606528,1.5669336754,0.3392193267\H,-4.37
06875174,0.4277864402,-0.8474282974\O,0.2972926111,1.8633762085,4.1537
420478\N,-1.8916689898,1.3100534513,4.0645319093\C,-1.1381374304,-0.47
66661726,2.4696598408\H,-0.1665488444,-0.3873897867,2.9681190405\H,-0.
9430423859,-0.5809474525,1.3966151637\O,0.2463938569,-3.6593076713,3.3
842518382\O,0.5421816188,-5.5986721694,4.1460320057\O,-1.4740361509,-4
.9673304267,3.6181940121\N,-0.2878717974,-4.8191869618,3.7359176496\C,
-1.8513238931,-1.7370874469,2.9719901463\H,-1.074755662,-2.5334454065,
2.8214058129\C,-2.1764340102,-1.7384488889,4.4621108063\H,-1.320496836
5,-1.4152420138,5.0640551107\H,-3.0175698323,-1.0689656612,4.677735263
8\H,-2.471091657,-2.7440611584,4.7818927441\C,-3.049851849,-2.18173606
25,2.1362249503\H,-3.9271138164,-1.5547490142,2.340610947\H,-2.8343404
933,-2.1389497793,1.0639498137\H,-3.3264239635,-3.2088590376,2.3966172
047\O,-3.2655715193,0.6885690322,0.8228937546\\Version=AM64L-G09RevB.0
1\HF=-913.5085514\S2=0.757817\S2-1=0.\S2A=0.750046\RMSD=0.000e+00\RMSF
=5.805e-06\Dipole=-1.7566737,0.4077206,-0.642164\Quadrupole=4.6987969,
-11.5379215,6.8391246,-2.2618984,-5.4642779,6.6501814\PG=C01 [X(C9H17N
2O6)]\\@

ν_{imag} = --163.3396 cm^{-1}

Sum of electronic and thermal Free Energies= -913.286833 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH\text{-}CH(CH_2C^{\bullet}(CH_3)_2)\text{-}CO_2CH_3$]

1\1\GINC-SPARTAN-RC102\FOpt\UM062X\6-31+G(d)\C9H17N2O6(2)\UWILLE\11-Se
p-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm geom=checkpoint gues
s=check scrf=(cpcm,solvent=acetonitrile) opt=(readfc,maxcycle=500) fre
q=noraman\\geom&freq\\0,2\C,0.621205929,-1.2708055275,0.1240402441\H,1
.2724244105,-0.7655640793,-0.5976613145\C,0.7019685374,-2.7694877356,-
0.0978229745\H,-1.4457214454,-1.6225293387,-0.1609073714\C,-1.13190687
37,0.3897729544,-0.2449798813\O,-0.2634519254,-3.4807246692,-0.2815077
672\O,1.9513541706,-3.2054426658,-0.0171520465\C,2.1334105364,-4.62632
00808,-0.1308392168\H,1.7613934978,-4.9738157309,-1.0958435823\H,3.205
5573307,-4.7888220578,-0.0476153378\H,1.6019432521,-5.1320336757,0.678
0784387\O,-0.2728321084,1.3048210143,-0.2208145803\N,-0.7556477454,-0.
8785614918,-0.116080556\C,0.3292095353,-1.6272348813,2.641066911\H,-0.
7226833711,2.6246040683,-0.1774027783\O,-1.7083859657,3.0497720848,1.7

848017736\O,-1.0076801459,3.6607337052,-0.190745316\O,-1.9297832136,5.
0988355645,1.1208225226\N,-1.5785304484,3.9491068714,0.9717127883\C,-2
.5972215247,0.6838306508,-0.3902342112\H,-2.9526662748,1.1241295899,0.
5472384243\H,-2.7425699938,1.4153446616,-1.1886349145\H,-3.1801426518,
-0.2122761746,-0.6065515852\C,0.7362859691,-3.0060959137,3.0582998526\
H,0.1694298953,-3.7867752225,2.5211585709\H,1.8002771992,-3.1897896368
,2.872488849\H,0.5385331615,-3.1687998979,4.125380271\C,-1.0221691464,
-1.1335524122,3.0498109598\H,-1.1393913299,-0.0607381799,2.8573151138\
H,-1.8321132261,-1.6516928061,2.5080984931\H,-1.2014896203,-1.31402975
91,4.1176983281\C,1.0897734636,-0.917520028,1.5671675338\H,0.98477389,
0.1687992391,1.6657558148\H,2.1576812326,-1.1578214389,1.6209675443\\V
ersion=AM64L-G09RevB.01\HF=-913.5556958\S2=0.755271\S2-1=0.\S2A=0.7500
23\RMSD=0.000e+00\RMSF=4.192e-06\Dipole=0.7230018,-4.0231037,-0.688650
2\Quadrupole=11.061328,-14.1748249,3.1134969,0.7896984,3.9681163,-7.90
42595\PG=C01 [X(C9H17N2O6)]\\@
Sum of electronic and thermal Free Energies= -913.332366 Hartrees

Reaction of $NO_3^\bullet$ with Ac-Ser(OMe)-CO_2CH_3

Reactant association complex

1\1\GINC-SPARTAN-RC087\FOpt\UM062X\6-31+G(d)\C7H13N2O7(2)\UWILLE\12-Au
g-2018\0\\#p M062X/6-31+G* nosymm geom=checkpoint guess=check scrf=(cp
cm,solvent=acetonitrile) opt=(readfc,maxcycle=500) freq=noraman scf=(q
c,direct)\\geom&freq\\0,2\C,-0.0704139828,-0.6835228144,-0.1815109561\
H,0.1750583837,-0.1929170834,-1.129082448\C,0.4322323168,-2.1187067044
,-0.2066513115\H,-1.9170707508,-1.4850293561,0.4634327579\C,-2.2953902
814,0.2481588813,-0.5731495416\C,-3.7802075109,0.1063493772,-0.3453213
071\H,-4.1177873804,0.9484042855,0.2653285474\H,-4.2888915947,0.165460
3989,-1.3103147236\H,-4.0495443783,-0.8270260918,0.1520262209\O,-0.264
9829314,-3.089660586,-0.0138572713\O,1.7383791067,-2.1662393169,-0.444
48411\C,2.3314256358,-3.4742954859,-0.4431540779\H,1.8761733005,-4.089
9175466,-1.220849944\H,3.3877463537,-3.3147161363,-0.6473462502\H,2.19
26657066,-3.9429965586,0.5325940671\O,-1.8318433467,1.1920449617,-1.22
52476125\N,-1.5066487251,-0.7013484578,-0.0306146592\N,0.7758146771,3.
0686882746,-0.8807619743\O,-0.4831426892,3.092030815,-0.4616175146\O,1
.1916837938,2.0578331291,-1.3756539378\O,1.3414220462,4.1185244948,-0.
6814822605\C,0.6198154725,0.0755393658,0.9510265906\H,0.147294136,1.06
29525575,1.0571089578\H,1.6858812601,0.2101275508,0.7196163719\O,0.454
6219604,-0.6808265898,2.1275248106\C,1.0232138587,-0.0482141541,3.2585
789257\H,0.5515692522,0.9269772598,3.4364825817\H,0.8465983621,-0.6994
93616,4.1154299256\H,2.1032882789,0.0942789865,3.1232614829\\Version=A
M64L-G09RevB.01\HF=-910.1077668\S2=0.756774\S2-1=0.\S2A=0.750036\RMSD=
0.000e+00\RMSF=4.235e-06\Dipole=0.2226721,-0.9697759,1.1363058\Quadrup
ole=8.4191138,-8.9501382,0.5310244,-5.7960184,-0.0319418,6.1566422\PG=
C01 [X(C7H13N2O7)]\\@
Sum of electronic and thermal Free Energies= -909.934857 Hartrees

Entry 16: Reaction at $CH_3C(O)NH$-CH(CH_2OCH_3)-CO_2CH_3

Transition state

1\1\GINC-SPARTAN-RC005\FTS\UM062X\6-31+G(d)\C7H13N2O7(2)\UWILLE\22-Oct
-2018\0\\#p M062X/6-31+G* scrf=(cpcm,solvent=acetonitrile) scf=(qc,dir
ect) nosymm geom=checkpoint guess=check opt=(ts,noeigentest,readfc,max
cycle=500) freq=noraman\\geom&freq\\0,2\C,-0.580858039,-1.0499388906,1
.8859980592\H,0.1517908628,-1.5395648123,2.5376301967\C,-1.1396132601,
0.1725235536,2.5985943894\H,-2.6075797363,-1.5956558735,1.650140268\C,

-1.4326106567,-3.2706232924,1.3487935418\C,-2.6476199118,-4.1251114135
,1.0817398665\H,-2.5945223221,-5.0160960666,1.711488675\H,-3.587417127
6,-3.6027073461,1.2695316055\H,-2.6222201487,-4.4487912566,0.037379445
4\O,-2.3138318361,0.3468903085,2.8340983498\O,-0.1712831484,1.02729706
18,2.9035510589\C,-0.5824145943,2.2491913145,3.5376214639\H,-1.0839139
953,2.0278214044,4.4812528824\H,0.3340879592,2.8089586478,3.7098854129
\H,-1.2544803868,2.7997798691,2.8770640111\O,-0.2895635982,-3.72816567
89,1.3200086642\N,-1.6635531301,-1.9616552948,1.6097026796\N,3.3189057
828,0.8764683729,0.0222265793\O,3.1304864863,-0.0752018724,-0.68896173
88\O,2.3900531,1.2037405862,0.9053124283\O,4.2721496153,1.6227043453,0
.0548485968\C,0.1382405056,-0.6012969618,0.61518406\H,0.5027537478,-1.
4638909562,0.0384971913\H,1.0471390083,-0.0099748204,0.9105318953\O,-0
.7204680101,0.2009651722,-0.1233736031\C,-0.1658618086,0.6356882684,-1
.3519388878\H,0.1709188202,-0.2197129926,-1.9495191085\H,-0.946902106,
1.179741488,-1.8820256165\H,0.685354298,1.3038161257,-1.1633345166\\Ve
rsion=AM64L-G09RevB.01\HF=-910.1030825\S2=0.758535\S2-1=0.\S2A=0.75005
3\RMSD=0.000e+00\RMSF=3.762e-06\Dipole=-1.3351382,0.9813889,0.0278381\
Quadrupole=-6.1537581,-1.7300501,7.8838083,6.4194402,2.4190779,4.46411
23\PG=C01 [X(C7H13N2O7)]\\@

ν_{imag} = -258.7698 cm^{-1}

Sum of electronic and thermal Free Energies= -909.932115 Hartrees

Product association complex [HNO_3 - $CH_3C(O)NH$-$CH(C^{\bullet}HOCH_3)$-CO_2CH_3]

1\1\GINC-SPARTAN-RC110\FOpt\UM062X\6-31+G(d)\C7H13N2O7(2)\UWILLE\16-Au
g-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(c
pcm,solvent=acetonitrile) opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\C,
-1.0875387933,-1.002075169,-0.1834520665\H,-1.3912122494,-0.4208804701
,-1.0675497026\C,0.1181648605,-1.8552016971,-0.564110071\H,-1.91716168
44,-2.8520029975,0.4049671524\C,-3.4425246112,-1.4584976711,0.24340636
02\O,0.2101016857,-3.0419589225,-0.3497410601\O,1.0625002299,-1.114651
1429,-1.1359696192\C,2.2770136133,-1.8023192078,-1.4855417864\H,2.0638
362181,-2.5810322241,-2.219294956\H,2.9292537635,-1.0419525328,-1.9114
531007\H,2.7206803712,-2.2461119122,-0.5911609722\O,-3.737945674,-0.29
54783613,-0.0297143049\N,-2.1558044266,-1.8854216328,0.2141977389\N,3.
1488789881,0.2449717483,0.7436504845\O,2.4364237834,-0.4372999282,1.45
0047968\O,2.5435594308,1.0690985599,-0.1360449983\O,4.3524563658,0.274
5669966,0.7229308444\C,-0.7237018691,-0.0614072852,0.9324919081\H,1.56
15697254,0.9060405557,-0.0273865082\H,-1.3612444048,0.0011030609,1.809
6974619\O,-0.0699184149,1.0709350477,0.5501311047\C,0.0451730661,2.059
6237656,1.5841500561\H,0.6173077891,1.6510321107,2.4227346596\H,0.5654
260151,2.9101383911,1.1455683244\H,-0.9528328569,2.3592964796,1.913368
7605\C,-4.4824813534,-2.4884170738,0.6110763367\H,-5.1499741032,-2.060
687087,1.3621615696\H,-5.0766253675,-2.7141929456,-0.2792359302\H,-4.0
509980973,-3.4153214551,0.993148346\\Version=AM64L-G09RevB.01\HF=-910.
1382228\S2=0.754154\S2-1=0.\S2A=0.750014\RMSD=0.000e+00\RMSF=3.372e-06
\Dipole=-0.6814274,-0.2084316,0.1845604\Quadrupole=-12.0749899,6.39785
31,5.6771369,8.0981459,-8.6400299,5.3282544\PG=C01 [X(C7H13N2O7)]\\@

Sum of electronic and thermal Free Energies= -909.962781 Hartrees

Reaction of $NO_3^\bullet$ with *N*-ethylacetamide

Reactant association complex

1\1\GINC-SPARTAN-RC103\FOpt\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\13-Sep
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm scrf=(cpcm,solvent=ac
etonitrile) opt=(calcfc,maxcycle=500) freq=noraman\\geom&freq\\0,2\N,1
.8289845464,1.9350055585,-1.6929007086\O,2.2668104822,1.1505229784,-0.
894362703\O,1.8587475819,1.5233127489,-2.9477646646\O,1.3685598364,3.0
336519139,-1.535067228\C,2.797427332,-0.2730619131,1.9240380179\H,2.39
26500511,0.7436624829,1.9558474219\H,2.7145696177,-0.7147490346,2.9225
121044\C,4.2415489329,-0.2704030682,1.4489397588\H,4.6391273478,-1.290
301378,1.4059233\H,4.317700784,0.1506961069,0.4402305183\H,4.624200922
4,0.9895056055,3.1299983058\C,6.4002039278,0.6393281473,2.1822026299\C
,7.141333264,1.4871308242,3.1910186519\H,7.9080079068,0.8725411407,3.6
700795703\H,7.6455916705,2.3006181163,2.6629368733\H,6.4894639841,1.90
88919624,3.9587129915\N,5.0694713708,0.5155625858,2.3549891992\O,6.987
3945242,0.0861951905,1.2468093731\H,2.179736867,-0.8594172283,1.239123
398\\Version=AM64L-G09RevB.01\HF=-567.8265613\S2=0.756147\S2-1=0.\S2A=
0.750029\RMSD=0.000e+00\RMSF=6.308e-06\Dipole=-1.5282295,0.5711618,0.9
837863\Quadrupole=-25.2901544,6.5903839,18.6997705,8.9704109,10.088614
6,3.7487956\PG=C01 [X(C4H9N2O4)]\\@

Entry 17: Reaction at $CH_3C(O)N\mathit{H}\text{-}CH_2CH_3$

Charge transfer complex

1\1\GINC-SPARTAN-RC058\FOpt\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\09-Oct
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cp
cm,solvent=acetonitrile) opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\C,-
0.4636361337,-1.7806207455,0.9099410189\H,-2.2099030556,-2.3911616315,
2.0293193685\C,-2.0605589462,-0.2202788474,2.0717896263\O,-1.422264098
,0.7257346507,1.6795410931\N,-1.636879914,-1.5381030359,1.680662241\O,
-3.0097007641,-3.6184236986,2.5562498773\O,-2.4796929984,-5.4242092953
,3.6235751731\O,-0.9435236608,-4.0040595852,3.0662600014\N,-2.13150855
48,-4.3701325603,3.0965901494\C,-3.2699221446,-0.202513733,2.948878879
2\H,-3.0610750033,-0.7480314783,3.8748707998\H,-3.5340700464,0.8293615
356,3.1747881557\H,-4.1027686047,-0.7044926031,2.4465591186\H,0.286106
7356,-2.1296015689,1.6457196435\C,-0.7027424726,-2.8967680273,-0.11118
34905\H,0.2386123748,-3.0950532488,-0.6269586938\H,-1.0311531941,-3.81
23102476,0.3867148011\H,-1.4519375956,-2.5882359647,-0.8431929304\H,-0
.1120159236,-0.8418419148,0.4796611679\\Version=AM64L-G09RevB.01\HF=-5
67.8429063\S2=0.763407\S2-1=0.\S2A=0.750089\RMSD=0.000e+00\RMSF=4.412e
-06\Dipole=0.4441924,3.3128764,-1.9002472\Quadrupole=18.7245304,-25.59
15127,6.8669822,-19.6291904,5.3642994,31.7222326\PG=C01 [X(C4H9N2O4)]\\@

Transition state

1\1\GINC-SPARTAN-RC167\FTS\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\05-Oct-
2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cpc
m,solvent=acetonitrile) opt=(ts,noeigentest,calcfc,maxcycle=500)\\geom
&freq\\0,2\C,-0.67905922,-1.6186280509,0.5019522835\H,-2.4542450057,-2
.5526354567,1.7081264126\C,-1.9848419094,-0.3150348941,2.1439244461\O,
-1.3497157827,0.6722826439,1.8453692814\N,-1.7540116901,-1.5391719563,
1.4443525274\O,-3.087392716,-3.5653219293,2.0176389536\O,-2.9164925026
,-5.1429990492,3.4823657604\O,-1.507199238,-3.5004573188,3.5073920952\
N,-2.4809003328,-4.0946671858,3.043984573\C,-3.0351836264,-0.406719675
9,3.2055745218\H,-2.7316427687,-1.151368314,3.9495496926\H,-3.16370586
03,0.565351306,3.6796901034\H,-3.9810417562,-0.7396664887,2.766766871\
H,0.1816645375,-1.0606864475,0.890775161\C,-0.3525887609,-3.0597868394

,0.1424047415\H,0.4627283015,-3.0749351082,-0.5840551363\H,-0.03492625
3,-3.6144461134,1.0297971849\H,-1.2176995254,-3.561000444,-0.299423457
6\H,-1.0123774208,-1.0508532278,-0.3864003556\\Version=AM64L-G09RevB.0
1\HF=-567.8418884\S2=0.762044\S2-1=0.\S2A=0.750073\RMSD=0.000e+00\RMSF
=7.030e-06\Dipole=0.5031748,1.9890362,-1.7017808\Quadrupole=12.327261,
-16.0244874,3.6972264,-16.3816294,4.4683997,21.4796268\PG=C01 [X(C4H9N
2O4)]\\@

Product association complex [HNO_3 – $CH_3C(O)N^{\bullet}$-CH_2CH_3]

1\1\GINC-SPARTAN-RC097\FOpt\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\16-Sep
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cp
cm,solvent=acetonitrile) opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\H,-
2.7597243312,-3.8230939265,0.5624081387\C,-2.3027702237,-1.3518406382,
1.3536436475\O,-2.0282663705,-0.2138067847,1.0188242858\N,-1.990193119
4,-2.4391238121,0.5087793114\O,-3.2434597992,-4.7480834996,0.578663948
1\O,-3.0991078497,-6.5892329503,1.6873952999\O,-1.7663938659,-4.952379
8071,2.1771657598\N,-2.6643478411,-5.4721260312,1.5411146994\C,-3.0345
209767,-1.7298852401,2.6063731146\H,-2.5010290804,-2.5336893268,3.1227
598425\H,-4.0288821577,-2.1074038348,2.3436688933\C,-0.9380953811,-2.2
601887883,-0.4548374098\H,-0.3509975531,-3.1862289673,-0.4563926751\H,
-0.2993738979,-1.4115811022,-0.1981440447\H,-3.1326118673,-0.859996665
8,3.255416252\C,-1.5686561306,-2.0684241796,-1.847576877\H,-0.76633203
64,-1.9920779302,-2.5856899559\H,-2.2072398875,-2.9170864624,-2.105000
1904\H,-2.1624083107,-1.1517631329,-1.8709299\\Version=AM64L-G09RevB.0
1\HF=-567.8435238\S2=0.761662\S2-1=0.\S2A=0.750065\RMSD=0.000e+00\RMSF
=6.708e-06\Dipole=0.4414664,1.260587,-0.7649433\Quadrupole=8.9242999,-
21.7671349,12.842835,-11.2025836,-0.0831344,10.1287948\PG=C01 [X(C4H9N2O4)]\\@

Entry 18: Reaction at $CH_3C(O)NH$-C*H_2*CH_3

Transition state

1\1\GINC-SPARTAN-RC102\FTS\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\13-Sep-
2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cpc
m,solvent=acetonitrile) opt=(ts,noeigentest,calcfc,maxcycle=500)\\geom
&freq\\0,2\N,3.9026292738,2.9018814537,0.3279143072\O,4.6508567605,1.9
576844815,-0.0239431393\O,3.3962132093,2.8821342314,1.4770953438\O,3.6
712801041,3.8341886691,-0.438312474\C,2.9229648646,-1.190268843,1.3342
401788\H,2.0657854766,-0.6595989144,0.9101890108\H,2.5774304996,-1.839
4532878,2.1430594145\C,3.9747777646,-0.218272763,1.8356627714\H,4.8912
504504,-0.6758255111,2.2250725333\H,4.3386736278,0.4485160491,0.990780
5922\H,2.4919538317,0.9242176254,2.7970334463\C,4.3588834503,1.4992884
217,3.5966814251\C,3.6633866578,2.4632760557,4.5028443164\H,4.41038221
35,3.0350317792,5.0516720801\H,3.0328536991,3.1358753264,3.9145518829\
H,3.0267782038,1.9166535062,5.2061416811\N,3.4899670578,0.7040045629,2
.7829279833\O,5.5474986011,1.3158379069,3.500626312\H,3.3600868436,-1.
8121872499,0.5511065841\\Version=AM64L-G09RevB.01\HF=-567.8383307\S2=0
.759426\S2-1=0.\S2A=0.750066\RMSD=0.000e+00\RMSF=1.002e-05\Dipole=-1.7
878374,-3.0624855,3.1222441\Quadrupole=-18.9127167,-9.1645217,28.07723
83,-24.7256956,13.4005049,7.0777055\PG=C01 [X(C4H9N2O4)]\\@

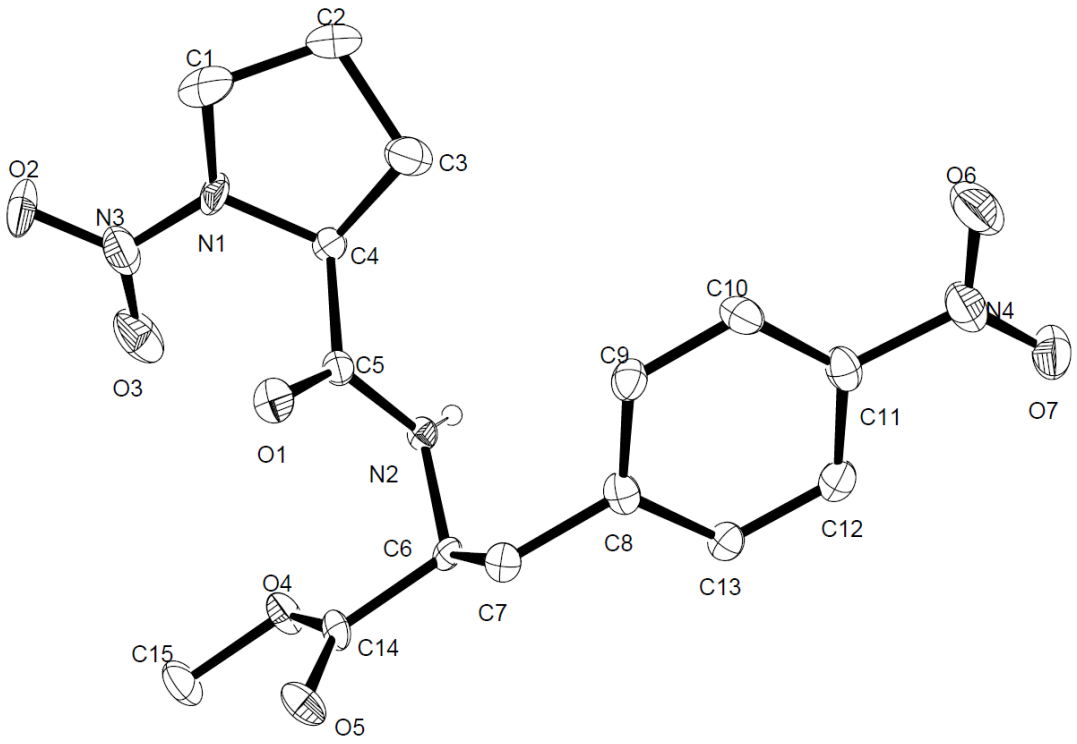
Product association complex [HNO_3 – $CH_3C(O)NH$-$CH^{\bullet}CH_3$]

1\1\GINC-SPARTAN-RC172\FOpt\UM062X\6-31+G(d)\C4H9N2O4(2)\UWILLE\20-Sep
-2018\0\\#p M062X/6-31+G* scf=(qc,direct) nosymm freq=noraman scrf=(cp
cm,solvent=acetonitrile) opt=(calcfc,maxcycle=500)\\geom&freq\\0,2\N,2
.2287591232,3.5658982834,5.9067741954\O,3.5340270437,3.3585156811,5.83
4033695\O,1.4939799818,2.8327034168,5.265994292\O,1.8724367479,4.48515

52239,6.6124879026\C,4.7128825924,-0.8338208875,2.9917169081\H,4.87212
46269,-0.8607493751,4.0592103272\H,3.7170059695,2.495252481,5.19989016
01\H,4.1227838627,0.4291342075,1.4766223817\C,3.9337188486,1.441076927
5,3.2190602123\C,3.4096294024,2.6585825647,2.5194889708\H,2.3797092264
,2.8380721737,2.8435023959\H,4.0105459948,3.525997142,2.8076304026\H,3
.4292628237,2.5570363091,1.4335463961\N,4.2489425658,0.3600006365,2.48
20762691\O,4.0795168915,1.395227055,4.468647781\C,5.1107099813,-1.9168
277044,2.0564771693\H,5.2611121724,-2.8482960711,2.6051282507\H,4.3453
838916,-2.0923518089,1.2902513593\H,6.0489716133,-1.6834402054,1.53174
30309\\Version=AM64L-G09RevB.01\HF=-567.874609\S2=0.760523\S2-1=0.\S2A
=0.750059\RMSD=0.000e+00\RMSF=9.195e-06\Dipole=1.5585402,-2.2233678,-3
.7614578\Quadrupole=33.5034376,3.5114953,-37.014933,-8.8240012,-11.434
2283,-32.7561761\PG=C01 [X(C4H9N2O4)]\\@

Crystallographic Data

Intensity data for **127B** were collected with an Oxford Diffraction SuperNova CCD diffractometer using Mo-Kα microsource radiation (graphite crystal monochromator λ = 1.54814). The temperature during the data collection was maintained at 130.0(1). Structure solution and refinement were implemented within the WingX suite of programs.



Thermal ellipsoid for **127B**, atoms are at the 20% probability level

Crystal data for **127B** $C_{15}H_{18}N_4O_7$ M = 366.33, T = 130.0(1) K, λ = 1.54184 Å, Monoclinic, space group $P2_1$ a = 4.7625(2), b = 13.0451(8), c = 13.8262(10) Å, β = 95.126(5)°, V = 855.55(9) Å^3, Z = 2, D_c = 1.422 Mg/m^3, μ(Mo-Kα) = 0.977 mm^{-1}, F(000) = 384. Crystal dimensions 0.508 x 0.062 x 0.027 mm^3. 4959 reflections measured, 2723 independent reflections (R_{int} = 0.04) the final R was 0.0447 [I > 2σ(I)] and wR(F^2) was 0.1183 (all data).

Crystal data and structure refinement for **127B**

Identification code	shelx	
Empirical formula	C15 H18 N4 O7	
Formula weight	366.33	
Temperature	130.0(1) K	
Wavelength	1.54184 Å	
Crystal system	Monoclinic	
Space group	P 21	
Unit cell dimensions	a = 4.7625(2) Å	α= 90°.
	b = 13.0451(8) Å	β= 95.126(5)°.
	c = 13.8262(10) Å	γ = 90°.
Volume	855.55(9) $Å^3$	
Z	2	
Density (calculated)	1.422 Mg/m^3	
Absorption coefficient	0.977 mm^{-1}	
F(000)	384	
Crystal size	0.508 x 0.062 x 0.027 mm^3	
Theta range for data collection	4.669 to 74.735°.	
Index ranges	-5<=h<=5, -13<=k<=16, -16<=l<=16	
Reflections collected	4959	
Independent reflections	2723 [R(int) = 0.0400]	
Completeness to theta = 67.684°	99.6 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.00000 and 0.83605	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	2723 / 2 / 239	
Goodness-of-fit on F^2	1.070	
Final R indices [I>2sigma(I)]	R1 = 0.0447, wR2 = 0.1079	
R indices (all data)	R1 = 0.0555, wR2 = 0.1183	
Absolute structure parameter	0.3(3)	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.180 and -0.240 $e.Å^{-3}$	